TEXT FLY SOME PGES AND TEXT CROSS WITHIN THE BOOK ONLY

UNIVERSAL
LIBRARY
OU_158631
UNIVERSAL
LIBRARY

OUP—408—16-6-64—5,000.

OSMANIA UNIVERSITY LIBRARY

Call No. 535.2/W84P Accession No. 44565

Author Wood, Robert W.

Title Physical optics. 3d. ed. 10

This book should be returned on or before the date last marked below.

PHYSICAL OPTICS



THE MACMILLAN COMPANY
NEW YORK • CHICAGO
DALLAS • ATLANTA • SAN FRANCISCO
LONDON • MANILA
IN CANADA
BRETT-MACMILLAN LTD.
GALT, ONTARIO

Optical Phenomena of Sodium Vapor

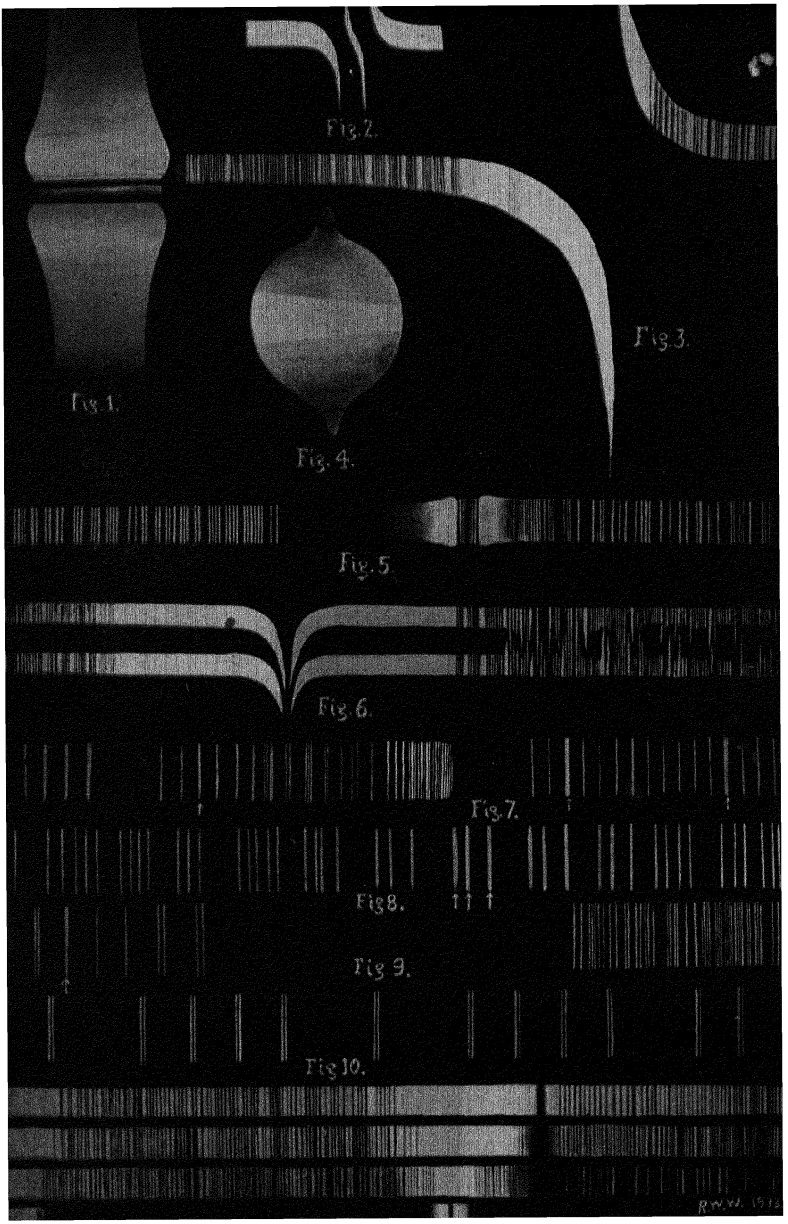


FIG. 1. Anomalous Spectrum by Vapor Prism.
FIG. 2. Dispersion at D Lines.
FIG. 3. Dispersion by Dense Vapor.
FIG. 4. Primary and Secondary Resonance Radiation.
FIG. 5. Magnetic Rotation Spectrum.
FIG. 6. Positive and Negative Magnetic Rotation.
FIG. 7. Resonance Spectrum: Lithium Arc Excitation. Exciting Lines Indicated by Arrows.
FIG. 8. Resonance Spectrum: Magnesium Arc Excitation.
FIG. 9. Resonance Spectrum: Bismuth Arc Excitation.
FIG. 10. Resonance Spectrum of Iodine Vapor. Hg 5461 Excitation.
FIG. 11. Absorption Spectrum of Sodium Vapor of Increasing Density.

PHYSICAL OPTICS

BY

ROBERT W. WOOD

HON. PH. D. (UN. OF BERLIN); LLD. (EDINBURGH); FOREIGN MEM. ROYAL SOC. (LONDON); ACCAD. DEI LINCEI (ROME)
PROFESSOR OF EXPERIMENTAL PHYSICS IN THE JOHNS HOPKINS UNIVERSITY

THIRD EDITION

NEW YORK
THE MACMILLAN COMPANY

THIRD EDITION COPYRIGHTED, 1934,

BY ROBERT W. WOOD

ALL RIGHTS RESERVED—NO PART OF THIS BOOK MAY BE REPRODUCED IN ANY FORM WITHOUT PERMISSION IN WRITING FROM THE PUBLISHER, EXCEPT BY A REVIEWER WHO WISHES TO QUOTE BRIEF PASSAGES IN CONNECTION WITH A REVIEW WRITTEN FOR INCLUSION IN MAGAZINE OR NEWSPAPER.

Printed in the United States of America.

Set up and electrotyped. Published Jan., 1934.

Fourteenth Printing, 1962

Revised edition copyrighted and published 1911,
By The Macmillan Company.
Copyright renewed 1939 by Robert W. Wood.

INTRODUCTION

The preparation of the present edition of *Physical Optics* has been governed by the generally expressed opinion that the most appreciated feature of the book is the large amount of space devoted to the experimental side of the subject, and the inclusion of somewhat lengthy descriptions of the experimental technique employed in the investigations with which the author has been more or less associated.

Since the last revision (1911) the theory of optical phenomena has developed to such an extent that nearly one half of the old edition has become obsolete. This has necessitated the rewriting of many of the old chapters and the inclusion of several new ones. The earlier chapters show little change as judged by the Table of Contents, yet some new material has been added, such as a revised list of light filters, Michelson's newer experiments on light velocity, Hubble's photographs showing the enormous velocities of the extra-gallactic nebulae and a number of other matters of minor importance.

The introduction of a new Chapter on the Origin of Spectra has seemed necessary, as the subjects of Magneto- and Electro-Optics, Resonance Radiation and Fluorescence, which certainly belong to physical optics, cannot be treated without an elementary knowledge of the present theories of the radiation of light by matter. The scope of the other Chapters is indicated in the Table of Contents. The Chapters on Relativity and the Nature of White Light have been omitted entirely, and a Chapter on the Raman Effect introduced. A good deal of new material will be found in the Chapters on Diffraction and Interference. The newer refractometers and interference spectroscopes are described in considerable detail, ten pages being devoted to the Lummer Gehrcke plate and the technique employed in its use. The Chapters on the Theory of Dispersion, Magneto- and Electro-Optics have been practically rewritten, and three new Chapters on Resonance Radiation and Fluorescence of Atoms and Molecules, and of Liquids and Solids replace the old Chapter on Transformation of Absorbed Radiations which dealt for the most part with phenomena for which no satisfactory theory had been developed at the time.

Nine new full page plates (two in color) and over one hundred and fifty new illustrations have been prepared; ninety of the old

ones having been discarded. Of the 827 pages of the new edition (an increase of 132 pages), only 332 have been taken from the old edition, 495 pages being new material.

In the work of revision I have attempted to give, in as many instances as possible, a physical picture of the processes usually described by equations. In this effort I have been greatly assisted by my colleagues, Professors Herzfeld and Dieke, with whom I have spent many hours in conference over the interpretation of the equations employed in the theories of scattering, refraction, dispersion, magneto- and electro-optics, thermal radiation and the Raman effect. Any small success achieved along these lines is very largely due to their patient coöperation.

R. W. W.

BALTIMORE, MD.
December, 1933.

CONTENTS

CHAPTER PAGE

I. THE NATURE AND RECTILINEAR PROPAGATION OF LIGHT. 1

The Nature of Light: Older Theories — The Elastic Solid Theory — The Electromagnetic Theory — Lorentz's Theory of Light Sources — The Quantum Theory of Planck — Simple Periodic Motion — Absence of Back-Wave — Wave-Front — Frequency and Wave-Length — Sources of Light for Experimental Purposes — Velocity of Light — Fizeau's Method — Foucault's Method — Michelson's Experiments — Newcomb's Experiments — Michelson's Later Measurements — Group-Velocity — The Döppler-Fizeau Principle — Döppler Effect in the Laboratory — Moving Source and Moving Observer — Rectilinear Propagation of Light: Huygens's Principle — Rectilinear Propagation — Effect of a Plane-Wave on an Exterior Point — Zone-Plate — Reflection and Refraction by Unpolished Surfaces.

II. THE REFLECTION OF LIGHT FROM PLANE AND CURVED SURFACES 48

Reflection of Waves and the Formation of Images — Reflection of a Plane-Wave from a Plane-Mirror — Reflection of Sound-Waves — Mirrors at Right Angles: Return of Ray to Source — Reflection by Ellipsoidal Mirror — Reflection from a Parabolic Mirror — Projection of Narrow Beams to Great Distances — Formation of Images by Parabolic Mirrors — Reflection by Hyperboloid — Reflecting Telescopes — Reflection from Spherical Surfaces — Reflection from Convex Spherical Surfaces — Reflection of Plane-Waves from Concave Spherical Mirrors — Formation of Images — Focal Lines — Fermat's Principle.

III. REFRACTION OF LIGHT 64

Refraction of Plane-Wave at Plane Surface — Total Reflection — Total Reflecting Prism — Total Reflection within Rods — Total-Reflection Refractometer — Refraction of a Wave by a Plane-Parallel Plate — The Ophthalmometer — Refraction of a Spherical Wave at a Plane Surface — Refractive Index of a Plate Measured by the Microscope — Fermat's Law — Refraction by a Prism — Refractive Index of a Prism — Effect of Refraction

CHAPTER PAGE

on the Width of the Beam — Magnifying Power of Prisms — Refraction by a Lens — Refraction by Sphere — Microscope Objective — Refraction of Light in Non-Homogeneous Media — Astronomical Refraction — The "Green Ray" — Schmidt's Theory of the Sun — Refraction by Planetary Atmospheres — Mirage — Non-Homogeneous Cylinders as Pseudo-Lenses — Curved Light Rays — Scintillation — The Method of Striae — Invisibility of Objects.

IV. THE ABSORPTION AND DISPERSION OF LIGHT . 99

The Absorption of Light — Laws of Absorption: Lambert's and Beer's — Determination of Extinction Coefficient — Body Color and Surface Color — Coefficient of Transmission: Dichromatism — Absorption by Porous Surfaces — Dispersion of Light — Achromatism — Partial Achromatization by a Grating — Direct-Vision Prisms — Achromatic Lenses — Resolving Power of Prisms — Christiansen's Experiment — Determination of the Dispersion of a Substance in the Form of a Powder — Anomalous Dispersion — Anomalous Dispersion in Its Bearing on Solar Phenomena.

V. THE ORIGIN OF SPECTRA 124

Older Theories — Production of the Hydrogen Spectra — Early Studies of Spectral Series — Bohr's Theory of Atomic Radiation — The Spectrum of Ionized Helium and Doubly Ionized Lithium — Elliptical Orbits and Two Quantum Numbers — Spectra of Alkali Metals — Penetrating Orbits — The Effective Quantum Number — Absorption of Light on Bohr's Theory — Absorption Spectra of Sodium — Molecular Spectra — Electronic Bands.

VI. INTERFERENCE OF LIGHT 157

Average Kinetic Energy of a Vibrating Particle — Composition of Vibrations — Distribution of Illumination — Resultant of a Large Number of Disturbances of Arbitrary Phase — Interference of Light — Lloyd's Single Mirror — Fresnel's Bi-Prism — The Corresponding Points of the Sources — Limit to the Number of Fringes — Shift of the Fringes by Introduction of Thin Transparent Plate — Interference Fringes with White Light — The Flow of Energy in a System of Interference Fringes — A Simple Interference Refractometer — Application of Interference Phenomena to Astronomical Observations — Williams Refractometer — Interference of Waves of Different Lengths. Light-Beats — Achromatic Inter-

CHAPTER PAGE

ference Fringes — Introduction of Thin Transparent Plate — Distribution of Phase over Small Area Illuminated by Source of Finite Size — Interference Fringes When the Sources Are in Line — Interference Fringes along Caustics — The Colors of Thin Plates — Newton's Rings — Newton's Black Spot — Refractive Index and Dispersion of a Thin Plate — Influence of Multiple Reflections — Silvered Films — Interference with Thick Plates — Testing Glass Plates for Flatness and Plane-Parallelism — Colors of Iridescent Crystals and Opals — Effect of a Prism upon Newton's Rings — Achromatization of the Fringes Formed by a Thin Reflecting Lamina — Preparation of Films for the Exhibition of Newton's Colors — Colors of Frilled Transparent Films on Metallic Surfaces — Curves of Equal Thickness and Equal Inclination. Haidinger's Fringes — Influence of Multiple Reflections on Width of Fringes — Stationary Light-Waves — Transition from Stationary Waves to Lloyd Mirror Fringes — Lippmann's Color Photographs.

VII. DIFFRACTION 218

Fresnel Class of Diffraction Phenomena — Diffraction by a Straight Edge — Circular Disk and Aperture — Graphical Solution of Diffraction Problems — Cornu's Spiral — Straight Edge — Narrow Slit — Narrow Wire — Two Parallel Slits — Fraunhofer Class of Diffraction Phenomena — Narrow Slit — Two Parallel Slits — Transition from Fresnel to Fraunhofer Class — Mathematical Treatment of Fraunhofer Diffraction Phenomena — Diffraction by a Rectangular Aperture — Diffraction by a Single Slit — Diffraction by Two Parallel Slits — Diffraction by Any Number of Parallel Equidistant Slits (Diffraction Grating) — Number of Spectra and Intensity Distribution: Absent Spectra — Verification of Results — The Laminary Grating — The Plane Grating — Overlapping Spectra — Resolving Power of Gratings — Focal Properties of Gratings: Zone-Plates — Intensifying Glass Gratings — False Lines or "Ghosts" Due to Errors of Ruling — Concave Gratings — Comparison Spectrum with Concave Grating — Testing of Resolving-Power and Focus — Paschen's Mounting of Concave Gratings — Eagle's Mounting — Concentration of Light into a Single Spectrum — The "Echelette" Grating for the Infra-Red — Diffraction by a Circular Aperture — Resolving-Power of Telescope — Babinet's Principle — Diffraction by Two Small Apertures — Halos — Young's Eriometer — Diffraction by Small Particles on the Surface of a Mirror — Diffraction in Relation to the Forma-

CHAPTER PAGE

tion of Optical Images — Image of Single and Double Apertures: The Edge-Wave — Abbe's Diffraction Theory of Microscopic Vision — The Colors of Mixed Plates — Talbot's Bands — The Echelon Grating — Transformation of Echelon into Prism — Anomalous Propagation of Waves in the Vicinity of Foci.

VIII. INTERFERENCE SPECTROSCOPES 292

The Michelson Interferometer — Adjustment of Interferometer — Effects of Surface Films on the Mirrors — Twyman and Green's Application of the Interferometer to the Correction of Imperfections in Prisms and Lenses — Determination of Refractive Index and Dispersion with the Interferometer — Light-Waves as Standards of Length — The Visibility Curves — The Echelon Grating — The Reflecting Echelon — The Interferometer of Fabry and Perot — Determination of Wave-Lengths with the Fabry and Perot Interferometer — Determination of Wave-Lengths — Absent Orders: The Compound Interferometer — The Interferometer of Lummer and Gehrcke — Interference Points of Crossed Plates.

IX. THE POLARIZATION OF LIGHT 329

Plane of Polarization — Discovery of Polarization — Polarization by Reflection — Angle of Polarization. Brewster's Law — Polarization by Refraction — Law of Malus — Polarization by Double Refraction — The Nicol Prism — Angular Aperture of Polarizing Prisms — Double Image and Polarizing Prisms for the Ultra-Violet — Detection of Polarized Light — Degree of Polarization — The Eye as an Analyzer: Haidinger's Brush — Polarization by Oblique Emission — Stationary Polarized Waves — Landolt's Fringe — Interference of Polarized Light — The Fresnel-Arago Laws — On the Polarized Fringes Produced by the Interference of Two Streams of Light Polarized at Right Angles — Elliptical and Circular Polarization — Polarizing Monochromator for the Separation of Close Spectrum Lines — Polarization of Emergent Light — Production and Properties of Circular Light — Circular Polarization by Total Reflection. Fresnel's Rhomb — Production and Properties of Elliptical Light — Babinet Compensator — Determination of the Constants of Elliptical Polarization — Position of the Axes — Ratio of the Axes — Elliptical Polarization by Reflection — Direction of Revolution in Circularly Polarized Light — Direction of Revolution in the Case of Fresnel's Rhomb — Natural and Partially Polarized

CHAPTER PAGE

Light — Rotation of the Planes in Unpolarized Light — Light-Beats — Relation between Amplitude and Intensity — Intensity of Circular and Elliptical Light.

X. DOUBLE REFRACTION. 365

Physical Explanation of Double Refraction — Wave-Surface in Uniaxial Crystals — Huygens's Construction — Verification of Huygens's Construction — (1) The refracting face is parallel to the optic axis, and the plane of incidence perpendicular to it — (2) Optic axis parallel to the surface of the crystal and to the plane of incidence — Fresnel's Theory of Double Refraction — The Normal Velocity Surface — The Wave-Surface — The Optic Axes or Axes of Single Wave-Velocity — Internal Conical Refraction — Axes of Single Ray Velocity. External Conical Refraction — Colors of Thin Crystal Plates — Colors of Crystal Plates in Convergent and Divergent Polarized Light — Uniaxial Crystal Cut Perpendicular to Axis — Isochromatic Surfaces — Isochromatic Surfaces in Biaxial Crystals — Position of Axes as a Function of Temperature — Phenomena Exhibited by Twin Crystals — Convergent Circular Light — Double Refraction in Non-Crystalline Media.

XI. METEOROLOGICAL OPTICS 388

The Rainbow — Polarization of Light in the Rainbow — Halos, Mock Suns and Related Phenomena.

XII. THEORY OF REFLECTION AND REFRACTION . . 397

Maxwell's Electromagnetic Theory — Derivation of Maxwell's Equations — Displacement Currents in Free Ether — Isotropic Dielectrics — Boundary Conditions — Velocity of the Wave — Deduction of the Laws of Reflection and Refraction for Transparent Media — Perpendicular Incidence — Oblique Incidence — Polarizing Power of Oblique Plates — Change of Phase — Change of Phase by Passage through the Polarizing Angle — Elliptical Polarization. Surface Films or Strains? — Total Reflection — Penetration of the Disturbance into the Second Medium.

XIII. THE SCATTERING OF LIGHT 423

Scattering of Light by Small Particles — The Scattering of Light by Gases: Blue-Sky — Scattering by Liquids, Solids and Crystals — The Residual Blue — Abnormal Polarization and Color of Light Scattered by Small Absorbing Particles — Density Fluctuations and Light

CHAPTER PAGE

Scattering — Measurement of Polarization — Depolarization of Scattered Light — Molecular Structure and Depolarization — Scattering and Transmission Colors of Granular Metal Films — Colors of Granular Films of Gold and Silver — The Scattering of Light by Sodium and Potassium Fogs — Scattering by Metal Colloids — Colors in Metal Glasses and Metallic Films — Polarization Effects of Ultra-Microscopic Particles.

XIV. THE RAMAN EFFECT 444

Smekal's Prediction — Raman's Experiment — Relation of Raman Lines to Infra-Red Absorption — Theory of the Raman Effect: Diatomic Gases — Selection Rule and Virtual Transitions — Non-Polar Molecules — Triatomic Molecules — Tetrachlorides — Relation between Raman and Resonance Spectra — Frequency Changes in the Unmodified Line — Influence of Density on Intensity Ratio of Modified and Unmodified Scattering — Polarization of Raman Radiations.

XV. THEORY OF DISPERSION AND SELECTIVE REFLECTION 469

Theory of Refraction and Dispersion — Free and Forced Vibrations — Molecular Scattering — Refraction and Dispersion — Experimental Proof of Dispersion Formula — The Damping Factor — The Complex Refractive Index — Experimental Proof of Dispersion Formula — Dispersion of Hydrogen — Dispersion of Metal Vapors — The Refraction and Dispersion of Dense Sodium Vapor — The "Hacken-Methode" — Absorption and Dispersion of Electrically Excited Hydrogen — Dispersion of Nitroso-Dimethyl-Aniline — Dispersion within the Absorption Band — Selective Reflection by Absorbing Media: Surface Color — Selective Reflection in Ultra-Violet and Visible Spectra — Absorption and Selective Reflection by Crystals — Residual Rays — Interferometer Study of Residual Rays — Absorption and Selective Reflection of Quartz — Dispersion of Quartz — Fine Structure of Reflection Bands — The Longest Heat-Waves: Focal Isolation — Absorption of Fluorite, Rock Salt and Sylvite — Dispersion of Fluorite, Sylvite and KBr — Effect of Density and Pressure on Absorption Lines and Bands — Reflection and Dispersion by Rock Salt — Selective Reflection by an Absorbing Gas.

XVI. THE OPTICAL PROPERTIES OF METALS . . . 542

Electromagnetic Theory of Metallic Absorption — Metallic Reflection — Determination of R from Emission

CHAPTER PAGE

— Determination of n and κ — Inhomogeneous Waves and Rigorous Formulae — Determination of Principal Incidence and Azimuth — Optical Properties of the Alkali Metals — Optical Constants of Silver and Gold — Effect of Thickness of Film on Reflection and Transmission — Relation between the Emission Coefficients of Metals and the Temperature.

XVII. ROTARY POLARIZATION 570

Rotatory Dispersion — Fresnel's Explanation of the Rotation — Unequal Absorption of the Circular Components — Other Rotatory Crystals — Rotatory Power of Biaxial Crystals — Rotatory Polarization of Liquids — Rotatory Power and Change of State — Rotatory Dispersion of Liquids — Molecular Rotation and the Influence of the Solvent on the Rotatory Power — Anomalous Rotatory Dispersion — Double Refraction of Active Liquids — Rotatory Polarization in Convergent Light — Physical Explanation of the Rotation — Theory of Rotatory Polarization — The Secondary Wave from the Compound Resonator.

XVIII. RESONANCE RADIATION AND FLUORESCENCE OF ATOMS 587

Resonance Radiation of Mercury Vapor — Absorption by a Resonating Gas — Polarization of the Resonance Radiation — Resonance Radiation of Sodium Vapor — Bohr's Theory of Resonance Radiation — Separate Excitation of Centres of Emission of the D Lines of Sodium — The Imprisonment of Resonance Radiation — Step-Up Excitation of Spectra — Optically Controlled Radiating States — Power Relation of Intensities — Formation of Metastable Atoms by Nitrogen — The Forbidden Line 2655.8 — Resonance Radiation of Electrically Excited Atoms — Resonance Radiation of Neon — Duration of the Metastable States — Fluorescence of Atomic Vapors — Sensitized Fluorescence — Effects of Foreign Gas on Sensitized Fluorescence.

XIX. THE RESONANCE AND FLUORESCENCE SPECTRA OF MOLECULES 616

Absorption Spectrum of Sodium Molecules — Discovery of Resonance Spectra — Resonance Spectra of Sodium Vapor — Resonance Spectra of Iodine — Theory of Resonance Spectra — Emission of Polarized Light by Fluorescent Vapors — Collision Transformation of Resonance to Band Spectrum — The Fluorescence of Diatomic Mercury — Time Interval between Excitation and

CHAPTER PAGE

Fluorescence of Mercury — Effects of Temperature on Mercury Fluorescence — Theory of Sodium and Mercury Fluorescence. The Franck and Condon Principle.

XX. THE FLUORESCENCE AND PHOSPHORESCENCE OF SOLIDS AND LIQUIDS 648

Resonance Radiation of Solids and Liquids — Absorption of Light in Solids — Fluorescence of Organic Compounds — Resonance Spectrum of Benzene — Fluorescent Efficiency and Stokes's Law — Methods of Investigating Fluorescence — Variation of the Intensity with the Angle of Emission — Effect of the Solvent on the Intensity and Color of Fluorescence — Fluorescence of Inorganic Compounds — Influence of Concentration — Polarized Fluorescence — Fluorescence of Crystals — Fluorescence of the Rare Earths — Phosphorescence — Duration of the Phosphorescence. Phosphoroscope — Nature of Phosphorescence: Solid Solutions — Fluorite Phosphorescence — Quenching of Phosphorescence by Red and Infra-Red Rays — Imprisonment of Fluorescent Radiation.

XXI. MAGNETO-OPTICS 668

The Zeeman Effect — The Zeeman Effect and Spectral Series — Study of the Zeeman Effect without a Spectroscope — Magnitude of the Separation — The Inverse Zeeman Effect — The Zeeman Effect on the Sun — Classical Theory of the Zeeman Effect — The Larmor Precession — The Quantum Theory of the Normal Zeeman Effect — Magnetic Quantum Number "m" and Space-Quantization — The Anomalous Zeeman Effect — Landé's Formula — Magnetic Moment of Atom: Bohr's Magneton — The Spinning Electron — Magnetic Levels in the Anomalous Zeeman Effect — The Paschen-Back Effect — Theory of Paschen-Back Effect — Larmor Precession of the Orbit in Strong Fields — The Faraday Effect: Magnetic Rotation of the Plane of Polarization — Explanation of the Magnetic Rotation — Resolution into Circular Components — Relation between the Field Strength and Rotation — Direction of the Magnetic Rotation — Time Required for the Faraday Effect — The Kerr Magneto-Optic Effect — Theories of Magnetic Rotation — Magnetic Rotation of Absorbing Gases and Vapors — Ladenburg's Method of Determining Magnetic Rotation — Magnetic Rotation within an Absorption Band: Experiments of Zeeman — Relation between the Faraday and Zeeman Effect — Paramagnetic Rotation — Zeeman Effect and Magnetic Rotation of Crystals at Low Temperatures — Voigt Effect: Magnetic Double

CHAPTER PAGE

Refraction — Cotton-Mouton Effect — Magneto-Optics of Resonance Radiation — Theory of the Effects — Hanle's Observation of the Larmor Precession — Effects of Rapidly Alternating Fields — Magneto-Optics of Sodium Resonance Radiation — Magneto-Polarization Effects in Vanishing Fields — Polarization in Step-Up Optical Excitation — Magneto-Optics of Band Spectra — Zeeman Effect for Band Spectra — Positive and Negative Rotation of Band Spectra Lines — Examination for the Zeeman Effect.

XXII. ELECTRO-OPTICS 736

The Kerr Electro-Optic Effect in Liquids — The Kerr Constant — The Method of Abraham and Lemoine — Determination of Velocity of Light with Kerr Cell — Determination of Duration of Fluorescence with the Kerr Cell — Modulation of Light with Kerr Cell: Light-Beats — The Stark Effect — Theory of the Stark Effect — The Electron's Motion in Space in the Stark Effect: Stereoscopic Models — Energy Levels and Stark Components — The Quadratic Stark Effect — Quadratic Effect of the Alkali Metals — Electrical Double Refraction in Vicinity of Absorption Lines — The Photo-Electric Effect — Electron Velocity and Frequency of Light — Relation between Electron Current and Wave-Length — Photo-Electric Effects by Polarized Light — Directional Distribution — Photo-Electric Effects of X-Rays — Ionization of Gases and Vapors — Photo-Conductivity: The Internal Photo-Electric Effect — Photo-Voltaic Effects.

XXIII. THERMAL RADIATION 772

The Relation between Emission and Absorption — Ritchie's Experiment — Kirchhoff's Law — Absorption by Flames — Temperature Radiation of Gases — Temperature Excitation of Line Spectra — Temperature Radiation of Solids and Liquids — Hydraulic Analogy of a Radiating Body — Lambert's Cosine Law — Variation of Intensity and Polarization with Angle of Emission — Emission of Polarized Infra-Red Rays — Relation between Emission and Reflection — Emission of Substances Having Absorption Bands — Emission of Polarized Light by Crystals — Proof of Kirchhoff's Law and Black-Body Radiation of a Hollow Enclosure — The Perfect Black Radiator — Equilibrium between Radiation and Material Bodies — Pressure of Radiation — Experiments of Nichols and Hull — Tangential Component of Radiation Pressure — Pressure of Radiation on an Absorbing Gas — The Stefan-Boltzmann Law — Proof of

CHAPTER PAGE

Stefan's Experimental Law — Optical Pyrometers — Temperature of the Sun — Change in the Spectrum of a Black Body with the Temperature. Wien's Laws — Complete Radiation Formulae — The Rayleigh-Jeans Formula for Spectral Distribution — The Nature of White Light.

XXIV. THE RELATIVE MOTION OF ETHER AND MATTER 812

Aberration of Light — The Klinkerfuess Experiment — Airy's Experiment — The Fresnel Entrainment Coefficient and Fizeau's Experiment — Lorentz's Treatment — Zeeman's Experiments — The Michelson-Morley Experiment — Miller's Experiments — Lodge's Experiment — Influence of the Earth's Motion on Rotatory Polarization — Effect of the Earth's Motion upon the Intensity of Terrestrial Sources — The Principle of Relativity.

APPENDIX 829

INDEX 833

PHYSICAL OPTICS

CHAPTER I

THE NATURE AND RECTILINEAR PROPAGATION OF LIGHT

The Nature of Light: Older Theories. — The foundations of our present knowledge respecting the nature of light were laid during the latter part of the 17th century, although the modern wave-theory did not take definite form until over a century later. The important discoveries which may be said to mark the beginning of the science of optics may be summed up in a few words.

In 1666 Sir Isaac Newton effected the prismatic decomposition of white light into its component colors, and proved that no further color change resulted from subsequent refractions. He moreover recombined the spectrum colors, and formed from them white light. This was a great step in advance in one way, for it had been thought previously that color was produced by refraction, manufactured by the prism so to speak, whereas Newton showed that the colors were originally present in the white light, the function of the prism being merely to separate them or sort them out, which it accomplished in virtue of its power of deviating rays of different colors through different angles. The present theory, however, is that white light is constituted of irregular pulses, the wave-trains giving rise to colors being manufactured within the prism, by its action on the pulses.

The importance of Newton's discovery is not to be underestimated on this account, and his conception of the nature of white light will be held throughout the greater part of this book, for it represents perfectly all of the experimental facts with which we are acquainted, and the treatments of nearly all of the optical phenomena which we are to study are greatly simplified by its use.

Newton elaborated what is known as the corpuscular theory of light, and clung to it tenaciously to the last, the weight of his opinion retarding in no small degree the development of the wave-theory, which was first clearly expressed in 1678. On the corpuscular theory light was regarded as a flight of material particles emitted by the source, the sensation of sight being produced by

their mechanical action upon the retina. The rectilinear propagation followed at once from the second law of motion, whereas the early supporters of the wave-theory were unable to account for it, as every known form of wave-motion bent freely around the edges of obstacles.

Reflection and refraction he explained as due to forces of repulsion and attraction exerted at the surface of the medium, the corpuscle being supposed to arrive at the surface in different "phases" in some of which it was repelled or reflected by the surface, in others attracted or refracted. He had studied carefully the colors of thin films and had established the relation between the color of the reflected light and the thickness of the film. The intensity of the transmitted light is a minimum, — he stated — if the corpuscles that have traversed the front surface of the film, having reached that surface in a phase of "easy transmission" have passed to the opposite phase the moment they arrive at the back surface. The results of his experiments with thin films, if handled from the point of view of the wave-theory, would have established its validity beyond any doubt, but he clung tenaciously to the idea of corpuscles.

As to the "phases" Newton expressed himself as follows: "Nothing more is requisite for putting the Rays of Light into Fits of easy Reflection and easy Transmission than that they be small Bodies, which, by their attractive Powers or some other Force, stir up *Vibrations* in what they act upon, which Vibrations being swifter than the Rays, overtake successively, and agitate them so as by turns to increase and decrease their Velocities and thereby put them into those Fits."

We thus see that Newton had a dim notion of a *dual* nature of light, Corpuscles and Waves (vibrations in a medium) acting together. This is interesting in view of the most recent theory of light corpuscles or "light quanta" (sometimes called "photons") which travel along paths marked out by a wave-field. This matter will be more fully discussed when we come to the quantum theory of radiation.

In the early part of the 19th century the final blow was given to the corpuscular theory by the experiments of Foucault, which showed that the velocity of light in water was less than in air, as required by the wave-theory whereas the theory of Newton required a higher velocity.

In 1676 it was demonstrated by Römer, a Danish astronomer, that light required a finite time for its propagation, travelling across space with a velocity which he estimated at 192,000 miles per second. Now the impact of corpuscles moving at such a speed

might well be expected to exert a pressure, and attempts were at once made to establish the materiality of light by detecting this pressure, all of which were failures however. At the present time we know that light does exert a pressure, though a very small one, but this pressure can be shown to be the necessary consequence of the impact of waves, so that it is as strong evidence of the truth of the wave-theory, as it would have been of the emission-theory had it been discovered in the days of Newton.

A wave-theory of light was first expressed in definite form by Huygens in 1678, and twelve years later he explained satisfactorily reflection, refraction, and the phenomenon of double refraction in uniaxal crystals, which was discovered by Bartholinus in 1670. Although he discovered the phenomenon of polarization, which would have practically been the death-blow to the emission-theory, had its nature been understood he was wholly unable to account for it. We must remember, however, that he had longitudinal waves in mind, *i.e.* waves in which the direction of the vibration was parallel to the direction of propagation, and polarization would be as difficult to account for by such a theory as by the corpuscular one. He was moreover unable to offer any satisfactory explanation of the rectilinear propagation of light, or the formation of shadows, and his theory fell into disrepute.

Grimaldi in 1665 was engaged with the study of diffraction, or the bending of light around the edges of obstacles. Admitting sunlight through two small apertures into a darkened room, he observed what he thought to be a darker region at the point where the two diverging beams overlapped. As he was merely looking for evidence of the non-materiality of light, he regarded his experiment as conclusive and pursued the subject no further. The apparent destructive interference of light, which Grimaldi thought that he had observed, was without doubt an effect due to contrast.

True interference was first observed by Thomas Young at the beginning of the 19th century nearly 150 years later, whose justly celebrated experiments established almost beyond question the validity of the wave-theory.

He regarded the waves as longitudinal, however, which assumption, though erroneous, did not affect the validity of his reasoning concerning the formation of interference fringes and the colors of thin plates.

Fresnel commenced his optical studies in 1814 and introduced, for the first time, the conception of transverse waves, a conception which he found necessary for an explanation of polarization. Rectilinear propagation he accounted for by a most ingenious method of dividing the wave-front up into zones, often wrongly

attributed to Huygens, and showing that the disturbances coming from the collective zones, produced zero illumination within the shadow according to the well-known principles of interference.

This was a very bold hypothesis for it appeared to be necessary to consider space filled with a medium capable of transmitting transverse waves such as are propagated in elastic solids. The idea of the so-called luminiferous ether thus came into existence and the elastic solid theory of light replaced the corpuscular.

The Elastic Solid Theory. — This theory, which held the field until Maxwell's development of the Electromagnetic theory, assumed all space filled with a medium having very remarkable properties. To transmit waves having the frequency and velocity of light, it must be millions of times more rigid than steel, while at the same time offering no resistance to the passage of the earth and planets around the sun. This ether was assumed to be composed of extremely minute particles, far smaller than atoms, held together by forces of attraction, and when one or more was displaced from its normal position, its neighbor was pulled aside, the displacement being handed on from one to another in the form of a wave. The theory had one great advantage: it gave a clear picture of the nature of light, and practically all of the optical phenomena known at the time were very satisfactorily explained by it. On this theory a source of light was considered as a collection of molecules, the atoms of which were in a state of vibration. These communicated their motion to the ether particles, and the light-wave advanced out into space.

There were many objections to this theory, one of which was the difficulty regarding the longitudinal disturbance which always accompanies the transverse one in the case of wave-motion in an elastic solid. No evidence of any such longitudinal disturbance in the ether has ever been found.

The Electromagnetic Theory. — In 1860 Clerk Maxwell showed that the propagation of light could be regarded as an electromagnetic phenomenon, the wave consisting of an advance of coupled electric and magnetic forces. If an electric field is varied periodically, a periodically varying magnetic field is obtained, which in its turn generates a varying electric field, and so the disturbance is passed on in the form of a wave, electric force generating magnetic and magnetic generating electric. Maxwell's theory predicted the speed at which these electromagnetic waves would travel from measurements of the magnetic fields of electric currents, the velocity of propagation being the ratio of the electromagnetic to the electrostatic unit, or the number of electrostatic units of quantity which are contained in an electromagnetic unit.

This ratio, determined from electrical and magnetic measurements, turned out to be the velocity of light, and indicated that light was essentially an electromagnetic phenomenon. The waves predicted by Maxwell were discovered by Hertz in 1892, who produced electrical oscillations in a pair of conductors between which sparks were passed. The conductors radiated waves which could be detected by another pair of similar conductors, which oscillated in resonance with the first pair, causing the passage of minute sparks between them.

Maxwell's theory, like the elastic solid theory, required an ether, but not a mechanical one, in which material displacements took place, but rather an electromagnetic one, in which displacement currents and magnetic fields could occur. The periodic disturbances, which are supposed to constitute these waves, were called displacement currents by Maxwell, and these displacement currents can occur in the free ether or in a dielectric, *i.e.* in a non-conductor of electricity.

Maxwell's theory told us nothing about the nature of this electric displacement, so that in one sense one's ideas about the real nature of the luminous disturbances were much vaguer than they were fifty years earlier, when the elastic solid theory was generally accepted, for in the motion of a solid we are dealing with perfectly definite physical processes. As Schuster remarks in the preface of his work on Optics, "So long as the character of the displacements which constitute the waves remains undefined, we cannot pretend to have established a theory of light."

Lorentz's Theory of Light Sources. — Maxwell's theory having identified light with an electromagnetic disturbance it was next necessary to postulate some sort of an electrical oscillator of atomic proportions to account for these very high frequency electric-waves.

The theory was expanded by H. A. Lorentz, who advanced the idea that the atoms and molecules contained electrons, minute spheres of negative electricity held in equilibrium positions, but capable of vibrating under the influence of a restoring force, when displaced. The bound electron, executing damped vibrations, was supposed to be the source of the luminous disturbance, or it might act as an absorber of radiant energy, when its natural frequency of vibration was in agreement with that of the radiation. This theory accounted for many of the newly discovered effects of magnetism on light, as well as the older phenomena of refraction, dispersion, etc.

The Quantum Theory of Planck. — In 1900 it was shown by Max Planck that if oscillations of the kind imagined by Lorentz, occurred in the molecules of the walls of an enclosure, and the electrons radiated continuously while they vibrated, the character

of the radiation within the enclosure would be totally different from what is actually observed, but that the experimentally observed facts would be completely accounted for if the oscillators emitted the radiation in outbursts of definite energy value. This theory will be more fully discussed in the Chapter on Radiation. There is, moreover, much other evidence which points to the emission of light in definite energy packets or quanta, and this evidence seems to show that these light quanta traverse space without losing energy or increasing in size. For example, when light of the proper frequency falls upon a metal surface electrons are ejected at a definite velocity which depends only on the frequency of the light and not on its intensity. Reducing the intensity merely reduces the number of emitted electrons, and no matter how much we reduce it, we still have electrons emitted, though in continually diminishing numbers. This suggested that light had a corpuscular nature, and a vast amount of effort has been spent in an attempt to reconcile these new facts with the numerous experiments that clearly indicate a spreading wave of continually decreasing intensity. The present trend seems to be a sort of fusion of the two theories, an energy quantum directed in its motion by a wave-field, for only in this way has it been possible to reconcile the facts of interference with the laws of the photoemission of electrons. These matters will be more fully discussed later on.

Simple Periodic Motion. — Many of the optical problems which we shall consider can be treated from the standpoint of the old elastic solid theory, for the propagation of light in many cases is governed by the same laws which hold in the case of acoustical phenomena and the transverse vibrations in elastic media. The source of light we may consider as a quasi-elastic oscillator, that is a vibrating particle which is urged back towards its equilibrium position by a force proportional to its displacement.

Such a particle vibrates in harmonic motion or as Schuster prefers to call it "Simple Periodic Motion," if no other forces, such as friction for example, come into play.

Let p be the force corresponding to unit displacement and m equal the mass of the particle, and we have

$$m\frac{d^2y}{dt^2} = -py = -mk^2y \quad \text{(writing } k^2 \text{ for } p/m\text{)}$$

$$\frac{d^2y}{dt^2} + k^2y = 0$$

of which the solution is $y = a \cos(kt + e)$.

This motion can be represented by the projection on the diameter of a circle of a point P moving around the circumference with a

uniform velocity. When $t=0$ the point is at P_0, the angle e being the initial phase or "epoch." The angle $kt+e$ is the phase after time t when the particle is a P. The radius of the circle is a.

The projection of P on the diameter is at distance y from the centre at time t

$$y = a \cos (kt+e).$$

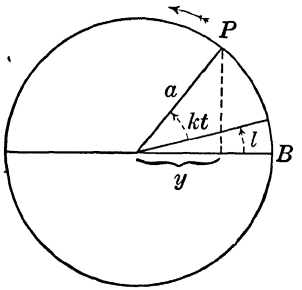


FIG. 1

The frequency f or number of to-and-fro excursions of the projection of the point P along the diameter is $k/2\pi$.

If the harmonic oscillator is embedded in an elastic medium, capable also of quasi-elastic vibration, a wave will be propagated with a velocity v, depending upon its elasticity and density. The wave-length $\lambda = vT$ in which $T = 1/f$ is the period or time of one oscillation. Substituting, we have for the displacement y along the x axis, if the spherical wave is transformed into a plane-wave, as by a lens or mirror,

$$y = a \cos 2\pi\left(\frac{t}{T} - \frac{x}{\lambda}\right).$$

We may get the form of the wave by giving to t any fixed value, for example $t = T/4$, when our equation becomes

$$y = a \sin 2\pi\frac{x}{\lambda}.$$

We can plot the curve in the following way.

We will plot the ordinates (y) for values of x equal to multiples of $\lambda/12$ (Fig. 2). Divide the circumference of a circle into 12

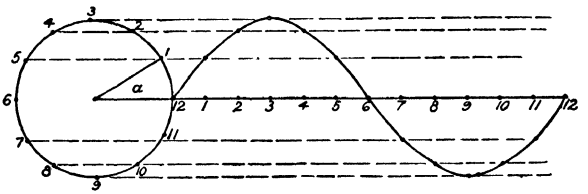


FIG. 2

equal parts, and call the radius unity. For $x = 1 = \lambda/12$, the cos of α is the ordinate of point 1 on the circumference of the circle. The same holds for the other points, therefore we have only to draw lines parallel to x through the points on the circle and mark their intersection with ordinates erected at 1, 2, 3, etc. The points thus determined lie on the wave.

Absence of Back-Wave. — If a point in a medium is made to vibrate in simple periodic motion, it sends out waves in both the positive and negative direction. Now when a wave meets a point in a medium, the point is made to execute periodic motion, and the wave beyond the point can be regarded as due to its motion. In this case, however, the moving point only sends out a disturbance in one direction, though its motion is identical with that of the point sending out waves in both directions. As we shall in the next chapter make use of this conception of a point thrown into vibration by a wave as a source of other waves, it is of some importance to distinguish between a secondary source of this nature and an actual source of light.

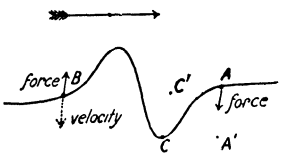


FIG. 3

Let the curved line in Fig. 3 represent a wave travelling towards the right. We know that this wave will be propagated with its type unchanged, and that the medium behind it will come to rest the moment the wave has passed. If, however, we distort the medium into the shape figured, and then release it, we shall have a wave travelling in both directions. The difference between the two cases will become at once apparent if we consider the velocities as well as the displacements of the particles. Consider the first case, that of the moving wave: the particle at A is acted on by a force drawing it downward, and being at rest initially it moves in consequence. The particle at B is acted upon in the opposite direction by an equal force. It, however, is *not at rest*, for it is moving in a downward direction with a velocity represented by the dotted arrow, for the wave has just passed by it, and it is returning to its position of equilibrium: this velocity just compensates the force due to the distortion of the medium and the particle comes to rest. In the second case both A and B are at rest initially, and both move the moment the restraint is removed, and we have a wave moving in both directions. We can in the same way see how the vibration of A by the passage of the wave through it fails to give a back-wave. It moves let us say to A', which it will do in time $T/4$. In the meantime the point C has returned to C', and its velocity just compensates the force due to the displacement of A, which in a medium initially at rest would result in a back-wave.

Wave-Front. — We may define the wave-front as the continuous locus of the points of the medium which are about to be disturbed. Thus defined the wave-front marks the limit which the disturbance has reached at the instant considered. A more general definition,

however, and one which we shall find more useful is the following. *The wave-front is the continuous locus of points which are in the same phase of vibration, or a surface of equal phase.* If this surface is plane, we speak of the waves as plane-waves, and since in isotropic media the rays are perpendicular to the wave-front, the rays are in this case parallel. The waves coming from sources of light situated at infinity (*e.g.* the stars) are plane.

If the source is at a finite distance, the wave-fronts are spherical, if the velocity of propagation is independent of the direction, as is the case in isotropic media. By means of mirrors or lenses it is possible to transform a spherical wave-front into a plane one, but we possess no means of starting a plane-wave directly. We can perhaps get a better case of what this would involve in the following way.

Consider a vibrating particle attached to an elastic string: waves will run along the string and the wave-front will be a point (Fig. 4*a*). Attach a number of strings to a rod vibrating in a direction parallel to its length (Fig. 4*b*), and the wave-front will be a straight line if we regard the strings as forming a continuous medium (Fig. 4*b*).

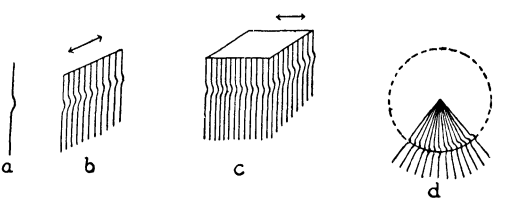


FIG. 4

There is no such thing in nature as a linear wave of light, for the reason that such waves can occur only in a medium of two dimensions. The conception of such a wave is often made use of in elementary treatments of diffraction, as the problems are much simplified by restricting the disturbance to two dimensions.

If now our strings are attached to a vibrating plane, the continuous locus of equal phase is obviously a plane, parallel to the moving plane, since the waves all start at the same instant, and travel with equal velocities. To realize this condition in optics it would be necessary to arrange a plane source of light, over the surface of which the vibration was uniform, *i.e.* the phases of all the vibrating particles would have to be the same, a condition which obviously cannot be realized. By attaching the strings to a vibrating point and arranging them so that they stretch out in all directions, we represent roughly the conditions under which we obtain a spherical wave. It should be observed, however, that in the case of a to-and-fro motion of the point, there are two directions in which transverse waves will not be given out, these directions coinciding with the direction of motion of the point. We have this circumstance

occurring in certain optical phenomena, as we shall see later on (certain facts connected with the Zeeman effect, for example).

Frequency and Wave-Length. — The length of the light-wave depends as we have seen upon two factors, the velocity and the frequency or time of vibration. Since the velocity in refracting media is usually less than the velocity in free space, the wave-length is reduced when the disturbance enters such a medium, for the frequency remains the same. The wave-length and frequency obviously depend upon the nature of the source. Flames colored by metallic salts may emit light of definite frequencies, such as the sodium flame, the light of which consists chiefly of two yellow radiations, commonly designated as the *D* lines.

Until quite recently the frequency of light was seldom used or mentioned in connection with optical problems. Light was defined in terms of intensity and wave-length. With the modern developments of the theory of spectra it has turned out that the frequency is the thing that matters, and very definite relations between the various lines of the spectrum of a given substance appear as soon as we substitute frequencies for wave-lengths. Practically all researches in spectral theory are now reported with no mention at all of wave-lengths, and it is accordingly now important for the student of modern optics to be able to think of light both in terms of frequency and wave-length.

The former is really the more logical way to define light, as the frequency remains the same while the wave-length changes with every passage of the light across a boundary separating media of different optical densities. In practice we employ what is known as the wave-number, which is the number of waves (in vacuum) per centimetre. If λ is the wave-length in vacuum measured in Ångström units (abbreviated to A.U.), 1 A.U. being .0000001 mm., the wave-number is

$$\nu = \frac{10^8}{\lambda}.$$

The wave-length of the *D* lines of sodium are .0005896 mm. and .0005890 mm., or 5896 A.U. and 5890 A.U.

In the following table are given the wave-lengths and wave-numbers of a number of typical radiations.

The "residual rays" are obtained by reflecting the radiations of a Welsbach lamp from a number of surfaces of quartz or other material. To get a comparative idea of these wave-lengths we may take a metre stick as our scale. Calling the Schumann waves 1 mm., green light will be 5 mms., red light 7 mms., and the long heat-waves from KI 96 cms., or practically the entire metre. Now

Radiation	Wave-Length in A.U.	Wave-Number
γ-rays of radioactive	.0056 to	$1.8 \cdot 10^{10}$
substances	.27	$3.4 \cdot 10^{8}$
Hard X-rays	.057 to	$2 \cdot 10^{9}$
	1.2	$0.8 \cdot 10^{8}$
Soft X-rays	1.2 to	$0.8 \cdot 10^{8}$
	13.	$7.6 \cdot 10^{6}$
Ultra-violet	40 to	$2.5 \cdot 10^{6}$
(Siegbahn)	140	700000
Ultra-violet	140 to	700000
(Millikan)	510	200000
Ultra-violet	510 to	200000
(Lyman)	1220	80000
Ultra-violet	1220 to	80000
(Schumann)	1800	55000
Ultra-violet	1800 to	27769
trans. by air	3600	
Violet	4000	24992
Blue	4500	22216
Green	5000	19994
Yellow	5800	17236
Red	6500	15380
Extreme red	7200	13885
Infra-red	10000 A.U. or	10000
1 μ	.001 mm.	
Residual rays from quartz	8.5 μ	1200
Residual rays from sylvite	60 μ or .06 mm.	166
Residual rays from KI	96 μ or	100
	nearly 0.1 mm.	
Focal isolation	110 μ — from	90
Rubens & Wood	Welsbach lamp	
Rubens & von Bayer	300 μ from Hg arc	33
Shortest electromag. waves	0.1 mm.	

Hertz waves	1.8 metres to 300 metres
Radio waves	300 metres to 30 kilometres or more.

consider this scale reduced in length to $\frac{1}{10}$ of a millimetre, and we have our waves as they actually are.

Light in which we have but a single wave-length is said to be monochromatic. It must be remembered, however, that strictly monochromatic light involves an infinite train of waves, such as would emanate from a particle the vibrations of which were subject to no sudden or gradual changes of phase. Absolutely homogeneous or monochromatic light is something that has no actual existence, though we are accustomed to speak of light which the spectroscope shows as a single narrow line, as monochromatic.

The color depends upon the wave-length, but the color cannot always be taken as an indication of wave-length, as certain colors

can be imitated by the simultaneous action upon the retina of two trains of waves, either of which acting alone would give rise to a totally different color from that perceived when both act together.

For example, a yellow scarcely distinguishable from the yellow of the sodium flame can be produced by a mixture of red and green light in the proper proportions. A screen can be easily prepared which transmits red and green only and in about the right proportions to produce the sensation of "subjective yellow," as it is called.

Lord Rayleigh recommended a mixture of an alkaline solution of litmus with chromate of potash. If a window, backed by well-lighted clouds, is viewed through such a solution and a prism, it presents a most splendid appearance, for the red and green images are widely separated, the region where they overlap being colored with the compound yellow. A screen capable of transmitting only the yellow region is difficult to prepare. A mixture of bichromate and permanganate of potash answers fairly well, and can be made to match the color of the first screen. A sodium flame is invisible through the first and easily visible through the second. Both together are practically opaque even with very intense white light.

The different radiations present in a source may be separated by a prism or diffraction grating, as we shall see, and we obtain in this way what is known as a spectrum of the source.

Not all of the radiations in the spectrum affect the eye, for, as we know by experiment, there are regions beyond the red and violet which we cannot see. The longer waves in the infra-red spectrum can be recognized by their heating power, or by their action on phosphorescent substances; the ultra-violet or short waves can be detected by photography or by their action in causing fluorescence.

The length of the light-wave can be measured with great precision by methods which will be described later on.

The spectrum has been gradually extended during the past quarter of a century, the gaps between the spectrum of electromagnetic, optical, X- and γ- (gamma) rays having been practically filled up. The optical spectrum was extended into the more remote ultra-violet by Schumann, who reached a wave-length of 1000 A.U. by employing a prism of fluorite and exhausting the air from his spectrograph, for these rays are powerfully absorbed by air. Lyman made more accurate determinations of the wave-lengths in this region and still further extended the spectrum. More recently Siegbahn, employing a grating ruled on glass, as recommended by the author, has reached a wave-length of 40 A.U.

Sources of Light for Experimental Purposes. — It will perhaps be well in the introductory chapter to describe briefly a number of sources of light, which will be found useful in experimental work. As a source of white light, the sun is to be preferred when great intensity is required. Next to this comes the electric arc, the type most suitable for experimental work being a lamp in which the positive carbon is horizontal. If great intensity is not necessary, a 100-watt Mazda lamp will be found very serviceable. It ranks next to the arc in intrinsic intensity, and requires no attention. If an electric current is not available, a Welsbach lamp, surrounded by a sheet-iron chimney furnished with a small vertical slit, will be found an excellent substitute. If an ultra-violet continuous spectrum is required, by far the best source is an end-on hydrogen tube, with a window of crystal quartz cemented on with sealing wax. It is best to have the tube in communication with an oil pump, with a stopcock between, and provided with a palladium tube for the introduction of dry hydrogen. The narrow portion of the tube should be silvered on the inside to catalyze the atomic hydrogen. The proper pressure is best determined by trial. When first started the water-bands are strong, but these disappear with continued operation. The pump should be operated occasionally and fresh hydrogen admitted. A cadmium spark can also be used.

As sources of monochromatic light we possess various colored flames and vacuum-tubes, from the spectrum of which we can pick out a monochromatic radiation by screening off the wave-lengths which are not desired. A simple form of apparatus for accomplishing this is described in Mann's *Manual of Optics*. It is easily constructed, not expensive, and can be made without the services of a skilled mechanician (Fig. 5). Light from a slit S, made parallel by a lens L, traverses a glass prism, after which it is reflected back through the prism and collimating lens, the convergent beam being then deviated to one side by a small right-angled prism, the focussed spectrum falling upon a screen provided with a vertical slit S'. By turning the mirror M by means of adjusting screws, any desired portion of the spectrum may be passed out through the side slit. It is possible with this instrument to obtain fairly monochromatic light from a source giving a continuous spectrum, or to pick out the highly homogeneous radiations which are emitted by metallic vapors, brought to a state of luminescence by the electrical discharge in vacuum-tubes, or in the arc or spark discharge.

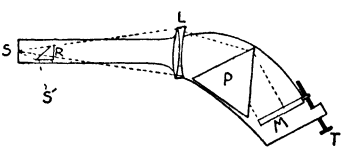


FIG. 5

The sodium flame is the most generally used source of monochromatic light. Its radiations, however, consist of two wavelengths, consequently its spectrum consists of two lines in the yellow very close together. To separate one of these from the other is a matter of considerable difficulty if much intensity is required and for most purposes the complete radiation will be found equally satisfactory. The most satisfactory flame can be obtained by winding a piece of asbestos paper around the top of a Bunsen burner (fastening it with wire) and saturating it with strong brine.

Monochromatic red light can be obtained by saturating an asbestos cylinder with a solution of chloride of lithium, and a satisfactory green by means of a small fragment of metallic thallium, fused to a loop of platinum wire. For long-continued work, however, the most satisfactory light is the mercury arc, from the radiation of which we can pick out by means of color screens, or the simple spectroscope described above, any one of the numerous bright lines.

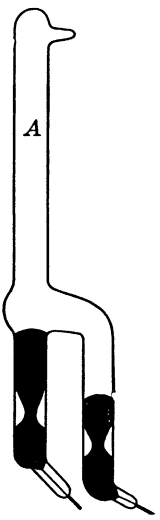


FIG. 6

A commercial mercury arc is the easiest to operate, and gives no trouble. It will be found immensely useful and should be installed in every laboratory. Its light is not as intense as that emitted by the lamps of fused quartz, and it does not give us much of the ultra-violet, but for most purposes it is most satisfactory.

Small mercury arcs, suitable for laboratory use, are easily constructed from pyrex glass, with tungsten electrodes. A convenient form is shown in Fig. 6. It is lighted by tilting and the vapor condenses in the cooling chamber A and running back keeps the mercury in the two branches at the constant levels shown in the figure. It is operated on a 110-volt D. C. current with a suitable resistance, in series, which may be increased after starting the lamp. During exhaustion the condensation chamber should be heated with a Bunsen flame for some time to drive off occluded water, and the lamp should be operated for half an hour or so before sealing from the pump. It operates continuously on about 2.5 amperes.

Quartz mercury lamps are now easily obtainable and have the advantage of supplying radiations in the ultra-violet. Used with suitable color filters such a lamp furnishes at will a number of approximately monochromatic radiations.

Suitable filters are supplied by the Eastman Kodak Co. or they can be made of solutions in glass cells.

Bichromate of potash transmits the green and the two yellow lines. Addition of a neodymium salt removes the yellow lines, without reducing the intensity of the green line in the least. No other substance is as satisfactory as this. Cobalt glass + aesculin solution transmits the 4359 line or still better a dilute solution of iodine in carbon tetrachloride combined with a sodium nitrite solution. Nickel sulphate is also useful. Chrysoidine + eosine transmits the yellow lines 5790. The chrysoidine should be dilute, and the eosine added until the green line disappears. A very thick cell with a saturated solution of bichromate of potash also transmits the two yellow lines. Glass very strongly colored with nickel oxide transmits the 3650 group of lines, and is opaque to visible light. This is most suitable for experiments in fluorescence and phosphorescence. It is supplied by the Corning Co., Corning, N. Y., and by Chance Bros. in England.

A film of silver chemically deposited on a quartz lens or plate is practically opaque to all radiations except the ultra-violet region 3160–3260. The silver film should be of such thickness that a window backed by a brilliantly lighted sky is barely visible through it. Lenses prepared in this way have been used by the author for photographing the moon, landscapes, and various objects in ultra-violet light. Bromine vapor gives equally good results with a much shorter exposure. Very dense cobalt glass combined with a layer a centimetre or more in thickness of a saturated solution of bichromate of potash cuts off everything except the extreme red above wave-length 69. This screen was used by the author in making infra-red landscape photographs. A clear blue sky is nearly black through it, while sunlit foliage comes out very bright.

A saturated solution of iodine in carbon tetrachloride is opaque to all visible radiations, and transmits freely the infra-red. A saturated solution of sodium or potassium nitrite is very useful in some cases as it absorbs the mercury 4046 line and everything below, while transmitting 4358 and everything above.

Gelatine filters, stained with aniline dyes, can be obtained in great variety from the Eastman Kodak Company, Rochester, N. Y., which supplies a catalogue giving photographs of the absorption curves. Some are, however, liable to fade under strong illumination.

Because of their permanence colored glasses are preferable in some respects, though the curves of absorption are usually less steep and the cut-off less sharp in consequence. A useful set of absorption curves of various substances has been published by K. S. Gibson [1] and the Corning Glass Company has recently issued a very complete catalogue of its colored glass filters.

[1] *Jour. of Opt. Soc. of Am.*, Sept., 1926.

The absorption curves of a number of useful filters are given on Plate I as follows:

1. Cobalt chloride in alcohol. 2. Potassium film (see Chapter on Optical Properties of Metals). 3. Praseodymium chloride. 4. Neodymium chloride. 5. Uranine (Na salt of fluorescein). 6. Cuprammonium. 7. Cobalt glass. 8. Corning signal red glass. 9. Corning G 34 glass. 10. Corning sextant green. 11 Signal green glass. 12. Sodium nitrite (saturated solution 2 cms.). 13. Corning glass 984 B. 14. 985 B. 15. G 986 A. 16. Cyanosine. 17. Potassium chromate. 18. Copper nitrate. 19. Corning glass G 586 J. 20. Cobalt chloride in acetone. 21. Cobalt sulphate in water. 22. Nickel chloride. 23. Iodine in CCl_4. 24. Nitroso-dimethylaniline in water. 25. Bromine vapor. 26 Chlorine.

Velocity of Light. — The first determination of the velocity of light was made by a Danish astronomer Römer in 1676. From observations made on the eclipses of Jupiter's satellites he showed that the inequalities noted in their times could be explained by the finite velocity of propagation of light. Since the time of rotation of the satellites around the planet is constant for each satellite, they will enter the shadow of the planet at regular intervals, and the times of the eclipses can be predicted with the greatest accuracy. Römer found, however, that the intervals between successive eclipses of a given satellite varied gradually if the observations extended over a year. The eclipses were found to occur earlier or later than the calculated time, according as the earth and Jupiter were on the same, or opposite sides, of the sun. The discrepancy was obviously due to the time taken by light to travel across the earth's orbit. Calculation showed that the velocity of light was about 192,000 miles per second.

The second determination was made in 1728 by Bradley, who discovered the phenomenon known as the aberration of light. He observed that the apparent position of the stars shifted slightly from time to time, and finally came to the conclusion that this small apparent motion could be explained by taking into account the earth's motion in its orbit, together with the fact that light is propagated with a finite velocity. The phenomenon of aberration will be more fully discussed in the chapter on the relative motion of matter and ether.

Fizeau's Method. — Galileo had made an unsuccessful attempt to determine the velocity of light, by placing two observers at a great distance apart, each furnished with a lamp. One observer uncovered his lamp and the second observer watched for the flash and removed the screen from his lamp at the moment it appeared. The first observer was to determine the velocity by noting the

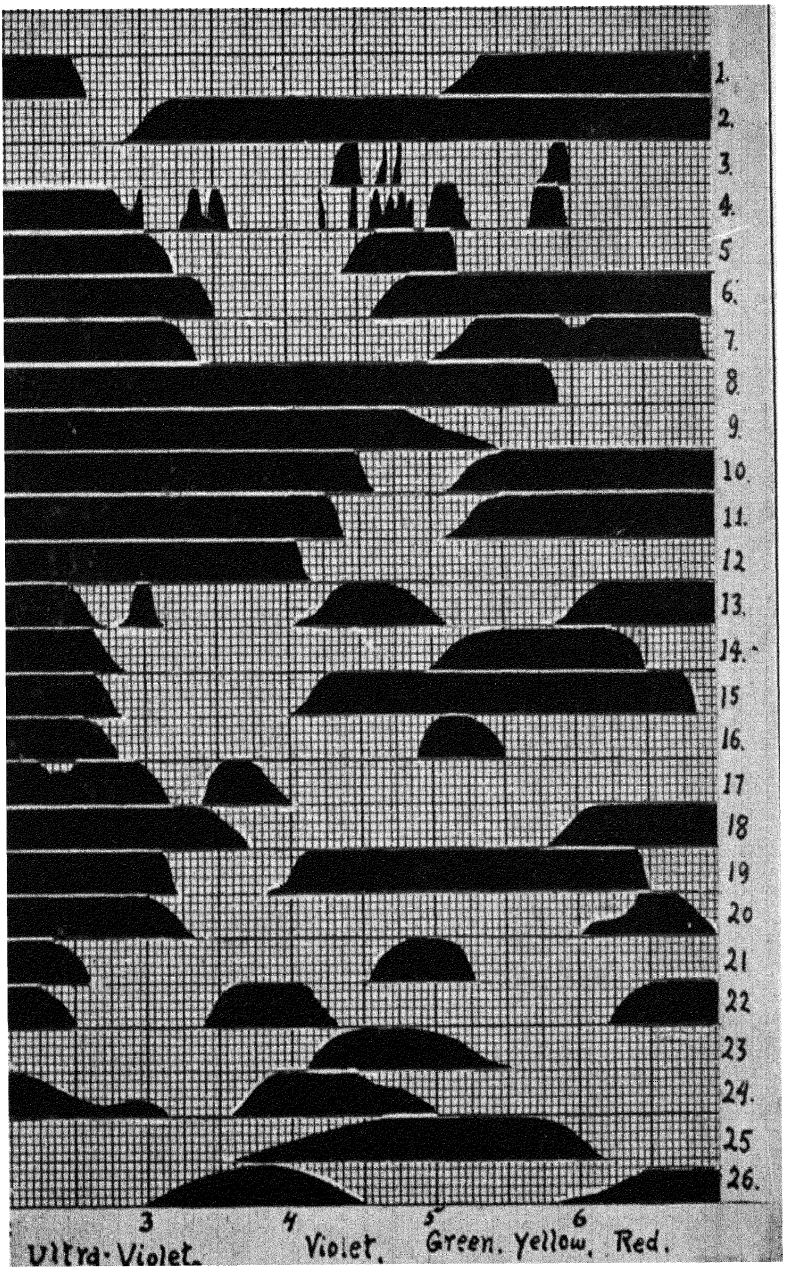


PLATE 1. ABSORPTION OF SUBSTANCES SUITABLE FOR FILTERS.

time elapsing between the uncovering of his own lamp and the appearance of the distant light.

This method failed obviously, owing to the enormous velocity of light. In 1849 Fizeau made an experimental determination of the velocity of light by means of a revolving disk furnished with a toothed rim. The method is essentially as follows: A beam of light was introduced into the tube of a telescope by means of a collimator fitted into its side, and was focussed by means of a reflecting plate upon the rim of the toothed wheel (Fig. 7). This point was at the principal focus of the object-glass of the telescope; consequently the light, after passing between the teeth of the wheel, was made parallel by the objective.

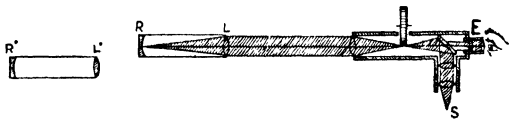


FIG. 7

After traversing a distance of three or four miles, it fell upon a second lens, which brought it to a focus upon a concave spherical mirror, the centre of curvature of which coincided with the centre of the lens. The light was thus returned as a parallel beam over the same path, and entered the eye-piece at *E*, passing through the reflecting plate. If the toothed wheel is rotated the beam of light will be made intermittent, and if the speed be great enough the light which passes through the space between two teeth will, upon its return, be cut off by the adjacent tooth, which in the meantime has advanced into the position previously occupied by the space. On looking into the telescope the observer sees at first a bright star, which diminishes in intensity as the speed of rotation is increased, finally disappearing entirely. Further increase in speed causes the reappearance of the star, the light passing through a given space, falling upon the next adjacent space upon its return. Fizeau experienced great difficulty in determining accurately the speed at the moment when the eclipses occurred. The image of the distant star was never bright, and the light reflected from the teeth of the wheel caused a general illumination of the whole field. To obviate this difficulty Young and Forbes, in repeating the experiment, bevelled the teeth so that the light reflected from them fell upon the blackened sides of the telescope. The teeth were also blackened so as to diminish their reflecting power as much as possible. In 1874 Cornu repeated the experiment with certain modifications. To avoid the difficulty of determining the exact moment at which the star was eclipsed, he made use of an electrical chronograph, arranged so as to record every hundred revolutions. Seconds were marked by a clock, and tenths of a

second by means of a vibrating spring. By means of a key the observer could record any instant at which he wished to know the velocity. The speed and its rate of change could be determined at every instant from the record of the chronograph. Instead of attempting to determine the moment of complete extinction, Cornu compared the brilliancy of the image with a light of fixed intensity. On increasing the speed the intensity of the image sank, and the speed of the wheel was recorded at the moment at which it was equal to the intensity of the standard light. After extinction the star reappeared and the speed was recorded at the moment when it regained its former brightness. The speed corresponding to complete extinction was the mean of these two. Cornu's final result for the velocity was 300,330 kms. per sec. in air, or 300,400 in vacuo.

Foucault's Method. — Wheatstone had suggested that a revolving mirror might be employed in the determination of the velocity of light, and his suggestion was taken up by Arago, but it remained for Foucault to carry out the experiment in a form capable of giving accurate results. The arrangement of his apparatus is shown in Fig. 8. Sunlight after transmission through an aperture at S and an achromatic lens I falls upon a mirror R, which can be rotated at high speed. A concave mirror M fixed at a distance of several metres returns the light to the revolving mirror. If the mirror R is at rest, the light returned by it after reflection from the inclined plane mirror comes to a focus at a.

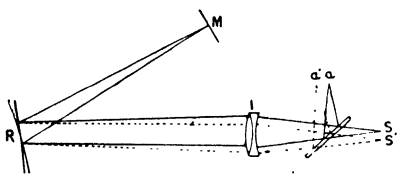


FIG. 8

The axis of the mirror R is at the centre of curvature of the mirror M, consequently the cone of rays, which converges upon M, is returned over the same path, and the rotation of R will not affect the position of the image at a. This, however, is only true if the mirror is in the same position when the rays meet it a second time, as will be readily seen by considering the passage of a ray from S to a. If the mirror turns through an appreciable angle while the light is traversing the distance $2RM$, the image will be shifted to a point a'.

The revolving mirror was driven by an air turbine, the speed being determined by a stroboscopic method. The displacement of the image amounted to only .7 mm., which gave for the velocity of light 298,000,000 metres per second.

Michelson's Experiments. — Foucault's method was improved by Michelson, who placed the lens between the two mirrors (Fig. 9)

The lens was 8 inches in diameter and had a focal length of 150 feet. The revolving mirror was placed 15 feet inside the principal focus, and the mirror M at a distance of 2000 feet. Deflections of the image amounting to 133 mms. were obtained, which made it possible to dispense with the oblique reflecting plate, and observe the image directly, with an eye-piece placed to one side of the slit. The speed of the mirror was determined by means of a tuning-fork, one of the prongs of which carried a light mirror, which reflected the light from the revolving mirror into the eye-piece. When the fork vibrated, the spot of light was drawn out into a band, which broke up into a number of moving images as soon as the mirror was set in rotation. A single stationary image was obtained only when the mirror made as many turns per second as the frequency of the fork; this condition was easily secured by regulating the air pressure at the turbine. The mean result for the velocity of light (reduced to the velocity in vacuo) was $299{,}910 \pm 50$ kilometres per second.

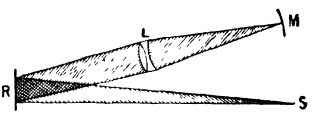


FIG. 9

Experiments were also made on the velocity of light in bisulphide of carbon, a tube three metres in length being interposed between the mirrors. The ratio of the velocity in air to the velocity in this fluid was found to be 1.758, while the ratio indicated by the refractive index is 1.64. This discrepancy will be explained presently.

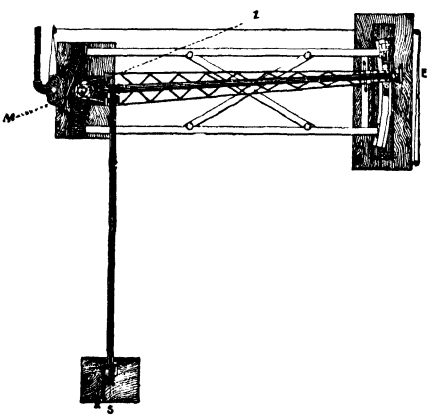


FIG. 10

Professor Michelson also experimented with lights of different colors, and found that red light travelled 1 or 2% faster than green light in the carbon bisulphide.

Newcomb's Experiments. — A series of experiments were made by Newcomb at Washington in 1880–82, with an apparatus of slightly different type. Sunlight entered the slit at S (Fig. 10), and, after reflection from a mirror at the elbow joint, passed through the telescope lens and fell upon the revolving mirror m, from which it was reflected along the line z to the distant mirror. The object-glass of the receiving telescope was immediately below that of the sending telescope, the light entering it being received from the lower part

of the revolving mirror. This consisted of a rectangular prism of steel (Fig. 11), the surfaces of which were nickel plated, driven by means of an air blast. The speed was regulated by means of a slight counter blast directed against the lower fan wheel.

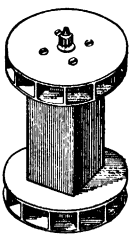


FIG. 11

By employing two lenses in the manner indicated, the diffused light from the strongly illuminated upper portion of the mirror did not enter the receiving telescope. The mirror could be driven in either direction, by interchanging the direct and counter blasts; the displacement measured could thus be doubled.

The quantity measured directly was the angular deviation of the return image, and not its linear displacement; this was accomplished by swinging the observing telescope, the eye-piece end moving along a graduated arc, the divisions of which were read by means of a pair of microscopes.

Newcomb's final result was, for the velocity in vacuo,

$$v = 299{,}860 \pm 30 \text{ kilometres.}$$

Michelson's Later Measurements.—A new determination of the velocity of light was made by Michelson [1] in 1924 at the Mt. Wilson Observatory in California. The light was reflected from an octagonal mirror rotating at high speed to a reflector on Mt. San Antonio, distant 22 miles, which returned it to Mt. Wilson, where the displacement of the image was measured with a micrometer. The apparatus is shown in Fig. 12.

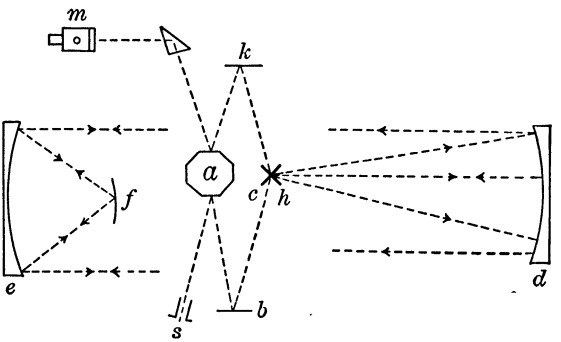


FIG. 12

Light from a slit S, illuminated with a Sperry arc, fell on a face of the octagonal mirror a, from which it was reflected by mirrors b and c to mirror d, of 24 in. aperture and 30 ft. focus. This mirror sent a parallel beam to the distant mirror e, which focussed it upon a small concave mirror f by which combination it was returned over the same path to the mirror d which reflected the light as a convergent beam to mirrors

[1] *Astrophysical Journal*, *60*, 246.

h and k and on to the opposite face of the rotating mirror. From here it passed to the eye-piece micrometer m the image appearing as a minute star whose position could be very accurately determined. The mirror rotated 530 turns per second and at this rate the returning ray arrived after one-eighth of a turn had been completed, so that the light was reflected from the next face of the mirror.

The time required for the light to make the round trip was .00023 sec., and the velocity determined was 299,728 kil. per sec. in air or reduced to vacuum, 299,796 kil. per sec. in vacuum. The result was considered correct to 1 part in 10,000.

An attempt to still further increase the accuracy by reflecting the light 88 miles to Mt. San Jacinto was a failure owing to poor atmospheric conditions. He then determined to measure the velocity in a vacuum-tube a mile long, the light being reflected back and forth five times, giving a total distance of 10 miles. A very good account of the engineering aspects of this gigantic undertaking will be found in *The Scientific American* for February, 1931.

Group-Velocity. — An important distinction exists between the velocity of a group of waves and the velocity of a single wave. We can get a very good idea of what is meant by group-velocity by throwing a stone into a quiet pond, and watching the circular waves which spread out. If the attention be fixed on a single wave-crest at the centre of the group, it will be seen presently to lead the group, the waves ahead of it appearing to die out, and in a few seconds its amplitude will become so small that the eye can no longer be kept on it. There are just as many waves in the group, however, as there were before, and a little further observation will reveal the fact that, as the waves in front die out, new ones appear in the rear. The group is obviously moving forward with a velocity less than that of the individual waves.

The explanation of the phenomenon was first given by Stokes, who regarded the group as formed by the superposition of two infinite trains of waves, of slightly different wave-length, which advanced in the same direction but with different velocities.

Lord Rayleigh was the first to draw attention to the bearing of group-velocity on optical problems. In his article on "The Velocity of Light"[1] he called attention to the fact that, in all experiments made for the purpose of determining the velocity of light, it is the group-velocity, and not the wave-velocity, which is actually measured. What is actually determined is the velocity with which some peculiarity impressed upon the wave-train moves forward. Since it is impossible in the case of light to pick out and watch a single wave, the best that we can do is to measure the

[1] *Nature*, 1881.

speed with which a block, cut out of a wave-train, advances. If the medium is free from dispersion, *i.e.* if waves of all possible lengths are propagated with the same velocity, the group-velocity and wave-velocity will be the same, the group being propagated without alteration.

This will be made clear by reference to Fig. 13. In the lower diagram we have two superposed trains of waves, moving in the direction of the arrow. The resultant disturbance is indicated in the upper diagram. The longer waves (dotted line) are out of step with the shorter (solid line) at A and C, and the resultant is zero at these points. At B, where there is agreement of phase, the resultant amplitude is double that of the single waves. If now the velocities of the two sets of waves are equal, it is evident that the group shown in the upper diagram will move forward without alteration with the wave-velocity. If, however, the shorter waves move at the higher speed, it is evident that they will presently get out of step at B, and into step at C, which now becomes the centre of the group. The group thus advances with a velocity greater than that of the individual waves. If the reverse is the case, the amplitude to the left of B increases as the group advances, the amplitude to the right of B diminishing, A becoming eventually the centre of the group. In this case the group-velocity is less than the wave-velocity.

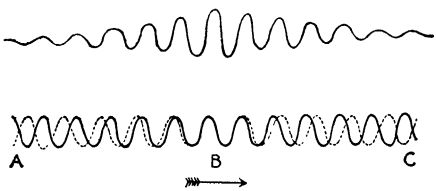


FIG. 13

We will now derive an expression for the group-velocity. Let the longer wave λ' (dotted line in Fig. 14) move with a velocity $V'>V$, the velocity of the shorter wave λ. Let T be the time required for the point marked V' to overtake the point marked V. When this event has occurred, the centre of the group, defined as the point of maximum resultant amplitude, and originally at B, will have moved back a distance of one wave-length. Now the crest V' is approaching the crest V with a velocity $V'-V$, therefore $(V'-V)T=$distance $V'V=\lambda'-\lambda$. If we write $dV=V'-V$ and $d\lambda=\lambda'-\lambda$, we have

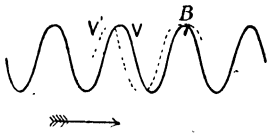


FIG. 14

$$T=\frac{1}{\frac{dV}{d\lambda}}.$$

During the time T the λ-wave train moves forward a distance VT, therefore the centre of the group as defined above has moved a distance

$$X = VT - \lambda = \left(V - \lambda \frac{dV}{d\lambda}\right) T,$$

and the group-velocity is given by dividing this quantity by T. Calling U the group-velocity, we have

$$U = V - \lambda \frac{dV}{d\lambda}.$$

If the medium is free from dispersion $dV/d\lambda = 0$, and $U = V$.

We shall have occasion to use the formula for group-velocity in studying the action of a prism on white light, which will be taken up later.

The difference between U and V only comes into play in determinations of the velocity of light in strongly dispersive media, the correction to be applied amounting to 7.5% in the case of bisulphide of carbon.

Michelson, employing the revolving mirror method, which has been shown by Rayleigh to yield U and not V, found the velocity of light in air 1.758 times greater than in this fluid, while determinations made by measuring the refractive indices gave the value 1.64. If we increase 1.64 by 7.5% we obtain the value 1.76, which is in close agreement with the value observed by Michelson.

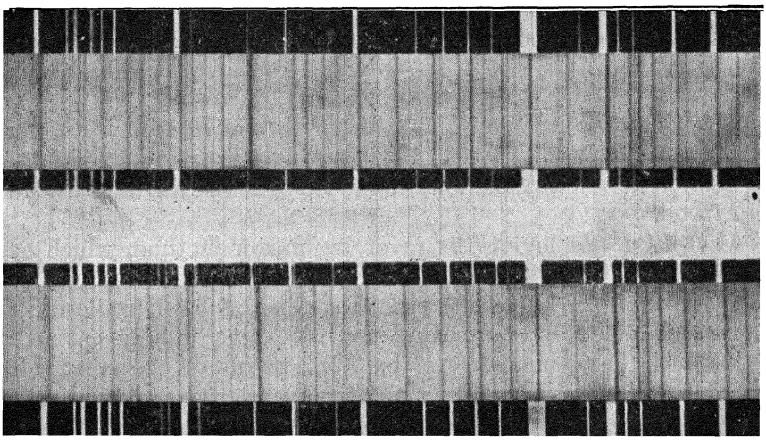
Lord Rayleigh,[1] in one of his earlier papers, stated that the aberration method gave V and not U, and that the close agreement between the velocity of light as determined by Römer's method, which gave U, and the aberration method which gave V proved that the velocity in space was independent of wave-length. Ehrenfest [2] showed, however, that the aberration method gave U and not V. This was admitted by Lord Rayleigh in a subsequent paper.[3]

The real evidence that waves of all lengths travel with the same velocity in free space, is furnished by the variable star Algol, which shows no color sequence when increasing in brightness, as would be the case if waves of different lengths travelled with different velocities.

The Döppler-Fizeau Principle. — Döppler, in 1842, called attention to the change in the pitch of a sound, which resulted when the source was moving towards or away from the observer, and

[1] Rayleigh, *Col. Papers*, vol. i, 322.
[2] Ehrenfest, *Ann. der Phys.*, *33*, 1571, 1910.
[3] Rayleigh, *Col. Papers*, vol. vi, 41, 1911; *Phil. Mag.*, *22*. 130, 1911.

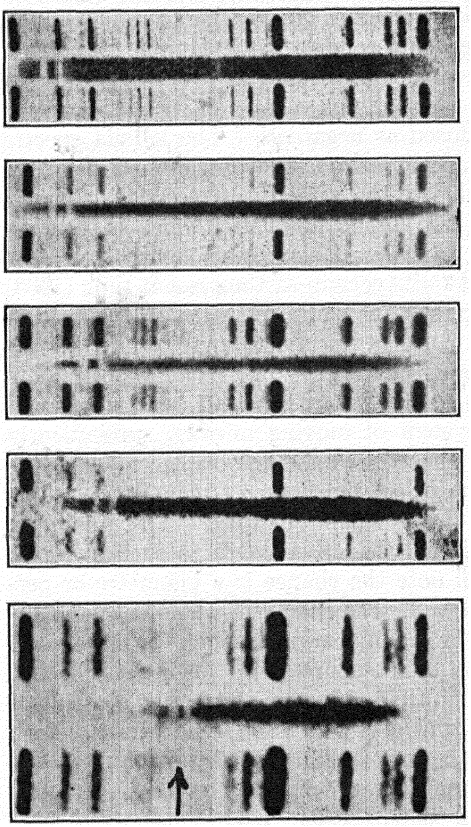
applied the principle to luminous disturbances radiated from bodies in motion, explaining the colors exhibited by certain stars as due to their proper motion. The acoustical phenomenon is most frequently heard when travelling in a railroad train. If a whistling locomotive is passed, the drop in the pitch is very noticeable, especially if the locomotive is moving rapidly in the opposite direction. Döppler's application of the principle to stellar phenomena was unsound, and Fizeau appears to have been the first to show that the effect would manifest itself as a slight shift in the position of the bright or dark lines in the spectrum. If the source of light is moving towards the observer, the frequency of



the disturbance as it passes the observer is increased, and the wavelength diminished: the spectrum lines are therefore shifted towards the violet: the reverse is true when the source is moving away in the line of sight. By photographing the spectrum of a star alongside of a comparison spectrum, it is possible to determine, not only whether the star is moving towards or away from us, but also the velocity with which it approaches or recedes. The principle has had wide applications in astrophysical research, and the rapidly accumulating data regarding stellar velocities will, at some future date, in all probability furnish the key to the solution of that greatest problem of astronomy, the nature of the motion of the multitude of suns which make up the universe.

Double stars have been discovered by the Döppler effect, the components of which no telescope will show separated, and their time of revolution about their common centre of gravity determined. Such stars are called spectroscopic binaries. The first was

discovered at the Harvard Observatory by Pickering. Observations of a number of spectra of this star, taken at different times, showed that the lines became double at stated intervals, an effect which could only be accounted for by assuming the source of light to consist of two bodies which alternately approached and



Sun spectrum centre. H and K absorption lines at left.
Helium spectrum above and below.

M. 32.
Velocity 200 kms. per. sec.

N.G.C. 385.
Distance 23,000,000 light years.
Velocity 3,000 miles per sec.

N.G.C. 4884.
Distance 45,000,000 light years.
Velocity 4,200 miles per sec.

Nebula in Leo cluster.
Distance 105,000,000 light years.
Velocity 12,000 miles per sec.

receded, in other words two bodies revolving around their common centre of gravity.

Keeler applied the principle to the study of the rings of Saturn, and showed that each portion of the ring was rotating at the speed which an isolated satellite would have at the same distance from the planet. Two photographs of the spectrum of a double star showing the Döppler effect, made by Professor Frost, are reproduced in Fig. 15 with comparison spectrum of iron above and below.

The shift between the dark lines of the star and the bright reference lines in the two photographs is plainly evident. In this case we have a star revolving about a dark body, its velocity in the line of sight varying with its position in the orbit.

The most astonishing results of all, however, have been obtained in the case of the extra galactic nebulae which show apparent Döppler shifts representing velocities of recession from 3000 miles per second to 12,000 miles per second. In 1929 Hubble found that practically all of these nebulae are receding with velocities proportional to their distance. Photographs of some of these spectra are shown in Fig. 16, reproduced as negatives. The nebula spectra are at the centre with a comparison spectrum of helium above and below. The shift due to the velocity of recession is shown by the position of the two broad absorption lines of calcium (H and K) indicated by an arrow in the lowest figure.

Astrophysicists are not quite certain at the present time (1933) as to whether these enormous line shifts represent real velocities, or are due to some other cause as yet undetermined.

Döppler Effect in the Laboratory. — The effect was first obtained in the laboratory by Bélopolsky [1] in 1901, who reflected a beam of light from a system of moving mirrors, subsequently analyzing the light with a spectroscope. The displacement of the spectrum lines was of the calculated order of magnitude, which was, however, an exceedingly small quantity. The minimum velocity capable of modifying the wave-length to such a degree that the spectroscope will note the change is a kilometre or perhaps half a kilometre a second. The change of wave-length resulting from reflection from a moving mirror is double the change resulting from the motion of the source with the same velocity. Bélopolsky made use of multiple reflections from two systems of mirrors, mounted on the rims of a pair of opposed wheels, which could be revolved at high speed. In this way he was able to obtain a shift of the spectrum lines which, though small, was easily measurable.

The experiment was repeated in 1907 by Prince Galitzin and J. Wilip with Bélopolsky's apparatus. They employed an echelon spectroscope and the mercury arc, and obtained much larger shifts than those observed previously on account of the much greater power of the spectroscope. The mirror wheels rotated at a speed of 45 revolutions per second, which represented a linear velocity of the mirrors of 30 metres per second. Six reflections were used and the displacement amounted to $\frac{1}{32}$ of the distance between the spectra of different orders. This was a double displacement ob-

[1] *Astrophysical Journal, 13,* pp. 15–24.

tained by two exposures, with the direction of rotation of the mirrors reversed between them. It is a very small amount, as we shall see when we come to the study of the echelon, but the calculated velocity of the mirrors agreed well with the observed. For example, the velocity calculated from the line shift in one case was .405 km. per second while that determined by measuring the speed of the wheels was .379.

Fabry and Buisson [1] in 1920 described a much simpler apparatus with which they measured the effect in the laboratory. A disk of white paper was rotated at a speed of 200 revolutions per second and illuminated by a Cooper-Hewitt mercury arc placed close to it. Seen from the side the disk appeared as a greatly elongated ellipse, one side approaching, the other receding, at a speed of 100 meters per second. A Döppler effect corresponding to a velocity of 200 metres was found for the light coming from opposite ends of the ellipse, amounting to one-sixth of a fringe separation in a Fabry and Perot interferometer, with a plate separation of 65 mms. It is of interest to note that a seaplane travelling 380 miles per hour (1931 Schneider cup record) or 550 feet per second, would, if carrying a mercury lamp, exhibit a Döppler effect of about the same magnitude!

The Döppler effect has been found by Stark in the case of the light emitted by the canal rays in vacuum-tubes. The canal

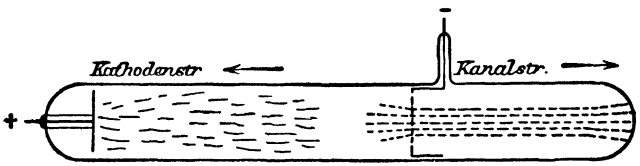


FIG. 17

rays occur where the cathode is perforated with small holes, and they are known to consist of the positively charged residue of the atom after the negative electron has been expelled. They are hurled down the tube with a prodigious velocity, and if the stream is pointed towards the spectroscope a line is observed shifted towards the violet. If the tube is oriented so that the stream is directed away from the instrument, the shift is in the opposite direction. If, however, the canal ray stream stands perpendicular to the collimator of the spectroscope, no shift is observed. The effect is not at all difficult to observe, and the canal ray tube can be made in a few minutes from some small pieces of glass tubing. The electrodes can be sealed in with sealing wax, if the

[1] *Jour. de Phys.*

discharge is prevented from reaching the wax joint. The construction is shown in Fig. 17. The cathode should be made from a piece of aluminum plate a trifle over a millimetre thick, perforated with numerous holes not over a millimetre in diameter. The anode is formed from a short piece of aluminum wire. The tube should be filled with hydrogen and connected to a mercury pump and its rounded end placed against the slit of a two or three prism spectroscope. The effect is best seen at the blue line of hydrogen, on account of the higher dispersion of the instrument in this region.

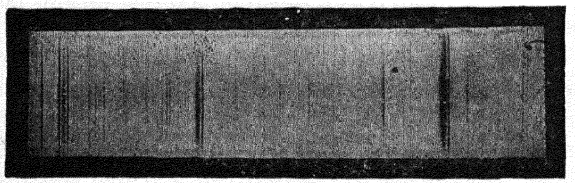


FIG. 18

At a pressure of a millimetre or two the line appears perfectly sharp and in its normal position. As the exhaustion proceeds a wing appears on the violet side, which presently detaches itself from the line and creeps slowly away from it. This gradual movement is due to the fact that as the vacuum becomes higher the canal rays travel at a higher velocity, and the change in the wavelength becomes greater. There is always present the line in its normal position, which shows that there is an emission as well from hydrogen atoms which do not partake of the motion of the canal rays. The velocity of the rays can be determined by measuring the position of the shifted part of the line. A photograph by Stark of the hydrogen lines showing the effect is reproduced in Fig. 18.

Moving Source and Moving Observer. — The change in the period T, of the radiation coming from a source of light moving with a velocity v, is given by the equation

$$T' = T\left(1 \pm \frac{v}{c}\right),$$

in which T is the actual period of the vibration, T' the period of the radiation, and c the velocity of light. It is to be carefully observed that when the source is in motion, the frequency of the vibration in the source differs from the frequency with which the waves pass by the observer, the former being unaffected by the motion.

The above equation for the change in the frequency of the vibration applies to the case of a stationary source of light and a

moving observer, as well as to a moving source. The two conditions are, however, represented by the same formula only when the velocity of translation is small in comparison with the velocity with which the waves travel. That this is true is evident from the following very elementary consideration. Let A be a source which emits ten waves per second, the waves travelling with a velocity of one metre per second. Let an observer B advance against this wave-train with a velocity of ½ m. per second. It is evident that the waves will pass him at the rate of 15 per second. Though the actual wave-length remains unaltered, the frequency of the vibration so far as B is concerned has been increased from 10 to 15. The deviation of waves by a prism depends upon their frequency; consequently in the case of light waves we obtain the spectrum line in a shifted position when the spectroscope is moving with a high velocity towards the source. Suppose now that B remain at rest, and the source A to move towards him with a velocity of ½ m. per second. During the time occupied by the source in moving ½ m., it emits ten waves. These waves will be crowded together into a space of half a metre, that is, between the point occupied by the first wave of the train of ten waves, at the end of one second, and the point occupied by the source at the same time. In other words the wave-length has been halved. This wave-train will sweep by the observer with a velocity of 1 m. per second, or with a frequency 20. The Döppler effect is therefore greater for the case of a moving source than for a moving observer, when we are dealing with velocities comparable with the velocity with which the waves travel.

The number of waves of frequency N, coming from a fixed source, which in one second pass an observer moving towards the source with a velocity v, is $N+v/\lambda$ or $N+vN/V$ in which $V=$velocity of light.

If, however, the observer is fixed, and the source moves with a velocity v, the wave-length is changed from $\lambda=V/N$ to $\lambda=V-v/N$ and the number of waves which pass the observer per second is the velocity V divided by this number, or $VN/V-v$. If a source moves with the velocity equal to or faster than that of the radiation, no periodic waves are given out. There is, however, a single wave like the bow wave of a ship. The sharp click of a high velocity bullet as it flies past us is an illustration of this. Calculate the change of wave-length produced by reflection from a mirror moving towards the source with a velocity equal to one-half that of the radiation.

Rectilinear Propagation of Light: Huygens's Principle. — One of the objections which was first urged against the wave-theory

of light was its failure to account for the rectilinear propagation of luminous disturbances, and the formation of shadows. Waves of sound and water waves were observed to bend around the corners of obstacles, and it was perhaps naturally argued that if light consisted of a wave-motion, it should behave in a similar manner. The objection was partially answered by Huygens, though it remained for Fresnel to give the complete explanation.



FIG. 19

Huygens's conception of the manner in which wave-motion was propagated was as follows: He regarded every vibrating point on the wave-front as the centre of a new disturbance: these secondary disturbances, travelling with equal velocity, are enveloped by a surface identical in its properties with the surface from which the secondary disturbances start, and this surface forms the new wave-front.

For example, in Fig. 19, consider O a luminous point, and AB a portion of the spherical wave-front. Adjoining points a, b, c, d, etc., on this wave-front are vibrating in unison and can be regarded as centres of new disturbances, which spread out around them as indicated by the dotted lines. It is evident that these secondary waves are enveloped by the spherical surface $A'B'$, and this surface is the new wave-front. If the luminous point is at a great distance, and we are dealing with a plane-wave, we have the condition shown in the lower figure.

This view of wave-propagation is known as the Huygens principle. It can be applied to the calculation of the position of a reflected or refracted wave-front, by regarding the points on the reflecting or refracting surface, as they are collectively or successively struck by the incident wave, as individual centres of new disturbances. For example, consider a wave-front AB descending in an oblique direction on a reflecting surface AC. The points a, b, c, d of the surface will be struck in succession by the points a', b', c', d' of the wave-front, consequently they will become successively the centres of secondary disturbances, as indicated in Fig. 20, which are enveloped by the plane surface $A'B'$. This is the reflected wave-front, and we shall see later on that it makes the same angle with the reflecting surface as the incident wave.

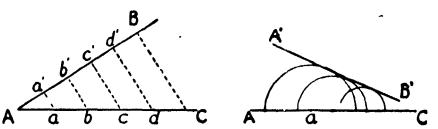


FIG. 20

Rectilinear Propagation. — Assuming Huygens's conception of the mechanism of wave-propagation to be correct, how are we to account for the rectilinear propagation of light? Suppose we have a luminous body at O (Fig. 21) and an opaque screen, a coin, for example, at A. We know that no light pehetrates into the conical region behind the coin (neglecting for the present a phenomenon known as diffraction).

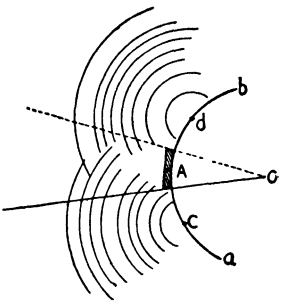


FIG. 21

But if all points on the wave-front are acting as independent sources, why does not the entire wave appear luminous to an eye behind the obstacle?

The answer given by Huygens was that these secondary waves produced no appreciable effect at a point unless they were *at that point* enveloped by a common tangent plane, or that the only effective portion of a secondary wavelet was the small point at its apex which touched the plane tangent to all of them. Huygens regarded the impulses as coming at irregular intervals, and his explanation of the rectilinear propagation of light amounted simply to the assumption that only one point on the secondary wavelet was effective in producing light, which is contradictory to experiment

Fresnel was the first to give a really satisfactory explanation Making use of the principle of interference discovered by Young, by which two luminous vibrations may destroy one another, he arrived at the somewhat startling conclusion that the absence of light in the shadow of a body was due to destructive interference between the secondary wavelets. This explanation not only accounted for the darkness behind the obstacle, but explained perfectly the slight bending of the rays around the edges, a phenomenon known as diffraction, which had been previously explained by assuming the edge to exert a modifying action on the luminous rays which passed close to it.

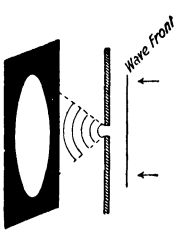


FIG. 22

It was no longer necessary to assume that only a minute portion of the secondary wave was operative in producing light, which as a matter of fact is contrary to experimental evidence, as can be shown by allowing a plane-wave to fall on an opaque screen perforated with a very small aperture. The point on the wave-front not cut off by the screen acts as a centre of a disturbance, which spreads out into the space behind the screen as shown in Fig. 22.

and a card placed in the position shown will be illuminated over an area many times greater than that of the aperture.

Waves of sound behave in a similar manner, and it is actually possible to photograph the secondary wavelet. The accompanying photograph is one of a series made by the author to illustrate certain features of wave-motion. The method by which the pictures were made will be discussed later on.[1]

For the present it will be merely necessary to state that in every case the sound photographed is the crack of an electric spark, which gives, of course, a single pulse, instead of a train of waves. The series shown in Fig. 23 was made to illustrate the principle of Huygens. The spark which started the wave was arranged to snap directly above a screen provided with a narrow slit. A short distance below this first slit a second was mounted, and it will be seen that the two narrow apertures become in succession the centres of secondary waves which diverge precisely as if the source of the sound, that is to say the spark, were situated in the aperture itself. In No. 1 of the series, the wave, which started at the point A, has just encountered the first screen. The aperture, which appears in the photograph as a break in the horizontal white line, becomes the centre of a new hemispherical wave, the gradual development of which is shown in Nos. 2, 3, and 4. In No. 5 the secondary wave has collided with the second screen and been reflected, the aperture in this screen becoming in its turn the origin of a new secondary wavelet. These pictures show that if all but a small part of the original wave is screened off, this small part becomes a complete wave, and again if a small portion of this secondary wave is allowed to pass through a small aperture, it becomes in turn a complete wave.

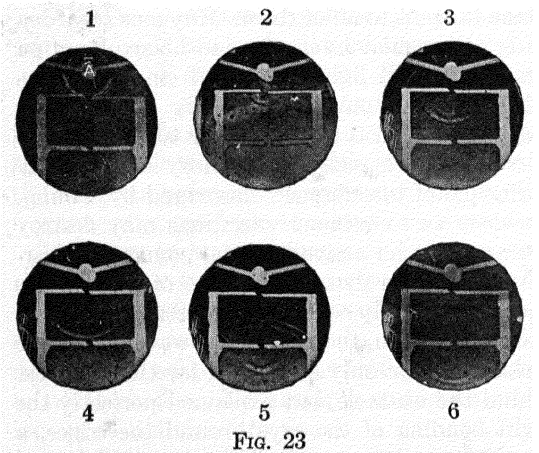


FIG. 23

Before considering in detail Fresnel's explanation we must make an assumption regarding the nature of the secondary wavelet,

[1] Wood, "Photography of Sound Waves," *Philosophical Magazine*, August, 1899.

which is based on the circumstance that no disturbance is radiated *backward*. An opaque screen which absorbs all of the energy falling on it has no effect whatever on the vibration of the medium between it and the luminous source.

From this we infer that the secondary wavelet is propagated only forward, and lies wholly in front of the plane tangent to the wave front at the centre of the wavelet. We are also justified both by theory and experimental evidence in assuming that the effect of the secondary wavelet is greatest on the line which is normal to the tangent plane at the point of tangency. This will be better understood by reference to Fig. 24, where AB is the wave-front, a the centre of any secondary wavelet, and xy the tangent plane behind which we assume that the secondary wavelet never spreads. The effect of the wavelet is greatest along the line, or in the direction ab, less along ac, and falls off continuously, having the value 0 in the direction ax. This may be summed up by saying that the effect of the secondary wavelet decreases with increasing obliquity. The reason for the absence of a back-wave has been given.

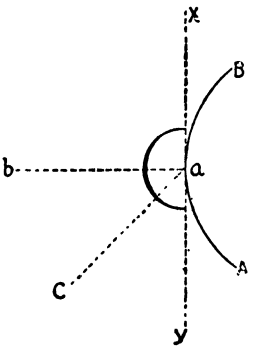


FIG. 24

Effect of a Plane-Wave on an Exterior Point. — Consider a plane-wave (Fig. 25) moving towards P, an exterior point: we require the effect at this point of all the secondary wavelets emanating from the wave-front. Draw a perpendicular from P to the wave-front, intersecting it at C, the pole of the wave with respect to P. Around C describe circles on the wave-front such that the first is half a wave-length further from P than C is, the second 2 half wave-lengths, etc., consequently the secondary disturbances coming from any circle will reach P half a wave-length ahead of those coming from the circle encircling it.

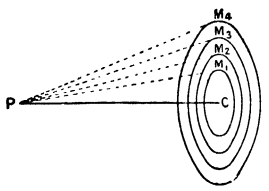


FIG. 25

We regard the effect of the disturbances coming from each ring as proportional to its area and as decreasing with increasing distance and obliquity. Let us now investigate the areas of the rings.

If $PC = b$ the radii of the circles are $\sqrt{b\lambda}$, $\sqrt{2b\lambda}$, $\sqrt{3b\lambda}$, etc., and the areas or πr^2

$$\pi b\lambda,\ 2\pi b\lambda,\ 3\pi b\lambda.$$

Neglecting the square of λ, we find the area of the central circle and each surrounding zone to be equal or $\pi b\lambda$.

For a zone at distance R from P we have its width given by $\lambda/2(R/\sqrt{R^2-b^2})$, as in the case of the element of the linear wave.

Its circumference is $2\pi\sqrt{R^2-b^2}$, and its area, or the product of these two quantities, is $\pi\lambda R$.

The effect due to the disturbances coming from a single one of the zones will be proportional to its area and inversely proportional to its distance. The slight increase in the area of the zones as we recede from the centre of the system is compensated by the increased distance, so that, other things being equal, we could regard the successive zones as producing equal and opposite effects at the point. The zones, however, become less and less effective as we recede from the centre owing to the increased obliquity. We can therefore represent the resultant effect by a series of terms of alternate sign which decrease slowly at first, and then more rapidly, eventually becoming zero, thus:

$$S = m_1 - m_2 + m_3 - m_4, \text{ etc.}$$

The sum of this series is usually stated as being equal to one-half of the first term plus one-half of the last term; the method usually adopted to prove this consists in balancing the second term against half of the first and half of the third, and so on. Schuster has shown that this treatment is too arbitrary, no reason being given why the balancing is not effected in some other way, for example, by considering the second term balanced by three-quarters of the first and one-quarter of the third, which would make the resultant outstanding effect approximately equal to one-quarter of that due to the first member acting alone. Schuster shows in what cases the addition of the series can be effected in the manner indicated.

The treatment is given in his Theory of Optics where he shows that the above expression is correct only for a series of terms decreasing according to a definite law, which, however, holds in the case with which we are dealing.

The problem thus reduces to a determination of the effect due to one-half of the central zone.

The secondary wavelets from this zone unite into a disturbance the phase of which is midway between those of the wavelets from the centre and rim, for we may divide the zone into a series of concentric rings of equal area, the effects of which at the point are equal in amplitude, and of phases ranging over half a complete period. These vibrations may be compounded as vectors by the method given on page 226. The resultant amplitude will be very nearly the diameter of a circle, the semicircumference of which is

made up of the vectors which represent the amplitudes contributed by the elementary zones into which we have divided the central circle. The direction of the diameter makes an angle of 90° with that of the first vector, consequently the phase of the resultant is a quarter of a period behind that due to the element at the centre. We must consequently consider that the secondary waves start with a phase one-quarter of a period ahead of that of the primary wave. The amplitude of the resultant bears the same ratio to the amplitude which would be produced if all the disturbances arrived in the same phase, that the diameter bears to the semicircumference, *i.e.* $2/\pi$. The matter of the acceleration of phase of the secondary wavelet of a quarter of a period, with respect to the phase of the primary wave, has sometimes been regarded as a sort of mathematical fiction. If the advance of phase really exists, a secondary wavelet, if isolated, would reach a distant point with a phase a little in advance of that of the primary wave which originated it. That this is actually so was shown by Gouy,[1] who formed a system of interference fringes with light reflected from a pair of Fresnel mirrors, and then introduced into the path of one of the interfering beams a screen perforated with a minute pin-hole, which became at once the source of secondary spherical wavelets. These interfered with the uninterrupted waves, and formed a system of fringes with a color distribution which showed that the disturbances from the pin-hole were advanced a quarter of a wave-length ahead of the primary wave which was intercepted by the screen.

Assume the amplitude on the wave-front to be unity, and consider that the secondary wave from a small element of its surface produces a resultant effect represented by kds. If r is the radius of the zone, its resultant effect will be $2/\pi \cdot k\pi r^2$. Now $r^2 = b\lambda$, and the amplitude due to the whole zone is therefore $2kb\lambda$. The whole wave will produce an amplitude one-half as great, which we may equate to unity, since we have assumed unit amplitude on the wave-front, and a plane wave is propagated without loss of amplitude. From this we find that k, the factor which represents the effect of the secondary wave, is equal to $1/b\lambda$. That the amplitude due to the secondary wave should vary inversely as the distance b is to be expected, but it may not be at once obvious why it should vary inversely with the wave-length. There is no mystery about the matter, however. If we keep the distance b fixed and increase the wave-length, we are obliged to increase the size of the zone, if the conditions are to remain as before, that is, if the zone is to produce the same effect at the point. The secondary disturbances

[1] Gouy. "Sur la propagation anomale des ondes," *Comptes Rendus*, 1890.

are now coming from a larger area, while only producing the same resultant effect, consequently the effect produced by any small element of surface *ds* will be proportionally smaller.

Let us now put the theory to experimental test.

Suppose we screen off all of the wave-front except the central circle of the zone series. There is now no encircling zone to partly neutralize it, and the illumination is greater than that due to the entire wave. This can be accomplished by placing a screen provided with a small circular aperture at such a distance from the point P that the area of the aperture is equal to the area of the central zone, when the amplitude at once becomes double, and the illumination four times that due to the unobstructed wave. It is of course apparent that the actual size of the zones on the wave-front in a given plane depends on the distance of the point P. As this distance increases the zones widen out. On a wave-front distant about 5 feet from the point, the zones would be of the size shown in Fig. 26, so that if our small circular aperture was of the size of the central circle in the figure, the illumination at a point on the normal 5 feet behind the aperture would be greater than if the screen were not present. And now comes a very curious fact: suppose we increase the size of our aperture until it contains another zone. The disturbances coming from this ring will be out of phase with those coming from the central circle, and will entirely destroy them. Thus by increasing the size of the hole we can reduce the illumination to zero.

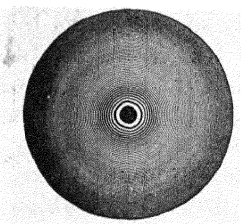


FIG. 26

An experiment proving this is described in the Chapter on Diffraction, page 225.

If we substitute a small circular disk for the aperture we find that the illumination on the axis of the shadow is unaffected by the interposition of the circular disk. By increasing the size of the disk we cut off another zone, still without influencing the illumination, and this may be continued, not indefinitely, but until, owing to the increasing obliquity, the effect of the zones begins to diminish appreciably. We thus see that the centre of the shadow of a circular body may, under certain conditions, be as brightly illuminated as the surrounding field, a proposition due to Poisson.

Fresnel's memoir on diffraction was presented to the French Academy and reported on by Poisson, who raised the objection that if the treatment were applied to the case of a circular disk (a case which had not been treated by Fresnel), it would lead to

the conclusion that the illumination along the axis of the disk would be the same as if the disk were absent, which was supposed to be a *reductio ad absurdum*. In this case it is clear that the illumination will be represented by the above-mentioned series, with as many members removed as there are zones covered by the disk, which will be, as before, one-half of the first exposed zone, and if we assume the zones to produce equal effects, the illumination should be the same as without the disk. As a matter of fact, the experiment had already been recorded by Deslisle, but it had been forgotten, and was rediscovered by Arago and Fresnel, who observed the bright spot in the centre of the shadow of a circular disk. This experiment also is more fully described under Diffraction.

Zone-Plate. — An interesting verification of Fresnel's theory is furnished by the zone-plate, usually attributed to Soret, though Lord Rayleigh constructed one four years previously, as shown by an entry in his notebook, without however publishing the result, thinking no doubt that it contributed nothing new to the theory, already well established by the experiment just described.

If we describe on a large sheet of white paper circles, the radii of which are proportional to the square roots of the natural numbers, we shall have very nearly an exact drawing of the zone system, the neglected terms containing the square of λ introducing a very slight error. If now we blacken the alternate rings with ink, and take a greatly reduced photograph of the whole on glass, we shall obtain a device which will enable us to screen off the alternate zones on the wave-front. Suppose we intercept a plane-wave with such a plate and consider the illumination at a point so situated behind the plate that the central circle of the plate corresponds in size and position to the first zone on the wave-front. The black rings stop all the secondary disturbances from the alternate or odd zones, which previously neutralized those coming from the even ones, consequently all the secondary disturbances coming from that portion of the wave-front covered by the plate reach the point in the same phase, and the illumination will be very intense. The whole surface of the zone-plate will send light to the point, the action being very similar to that of a convex lens. The distance of the illuminated point from the zone-plate we may speak of as its focus, and we readily see that the smaller the zones the shorter the focal length.

Lord Rayleigh's entry, dated April 11, 1871, is as follows: "The experiment of blocking out the odd Huygens zones so as to increase the light at centre succeeded very well and could be shown in quite a short space. The negatives should not be varnished. I have little

doubt that the number of zones blocked out might be advantageously increased much beyond what I used (15). No great accuracy is required in filling in the odd zones with black." A reduced copy of a drawing made by the author in 1898 is reproduced on Plate 2. This can be used for making reduced copies by photography on lantern slide plates or better on a collodion emulsion on thin plate glass. See Lippman photographs for directions for preparing such an emulsion.

It is well to make several plates of different focal lengths. That they have properties similar to lenses may be well shown by holding one of suitable focus, say half a metre, between the eye and a distant lamp. If the central zone is brought over the flame, the whole plate fills up with light like a lens. By combining a zone-plate with a low-power eye-piece we can form a telescope which will give a sharp image of a brilliant object, such as an incandescent lamp.

Lord Rayleigh, in his article on Wave-Theory in the *Encyclopædia Britannica,* called attention to the fact that if it were possible to provide that the light stopped by the alternate zones could be allowed to pass, but with a reversal of phase, a fourfold intensity in the illumination at the focus would result. In this case the secondary disturbances from all the zones, both odd and even, would reach the point in the same phase. This suggestion prompted the author to construct phase-reversal zone-plates [1] by making the zones of a thin film of gelatine on glass, the thickness of the film being such as to retard the waves one-half wave-length. These were made by coating a glass plate with a thin film of gelatine containing a little bichromate of potash. Contact prints are made from the reduced photographs in sunlight, and washed for a few seconds in warm water.

A telescope can be made with an objective formed by a zone-plate of say 50-cm. focus, and the eye-piece of a very short focus plate made by photographing the drawing with a microscope objective of low power. The telescope should be made with two tubes of such length (25 cms. each) as to permit of bringing the smaller plate within 25 cms. of the larger, in which position a magnified *erect* image of an incandescent lamp (with a horse-shoe filament) is observed, the eye-piece functioning as the concave lens of an opera glass. In other positions inverted images are seen, the eye-piece acting as a convex lens for viewing the images at the various foci. An explanation of this property of the zone-plate is given in the Chapter on Diffraction.

Phase-reversal zone-plates have recently been made by the

[1] *Phil. Mag.*, June, 1898.

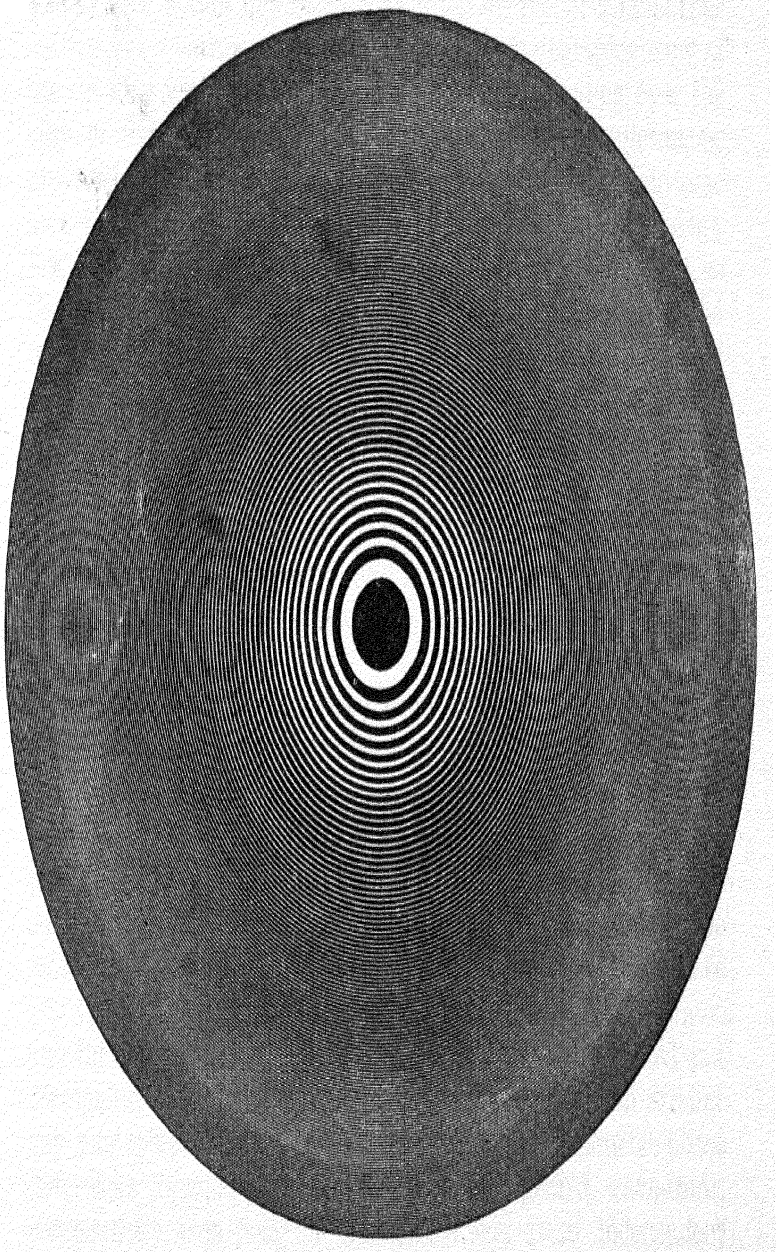


PLATE 2. ZONE-PLATE, FROM A DRAWING.

following process: A rotary ruling machine was constructed by mounting a "turn-table" (as used for sealing microscope cover-glasses to slides) on the carriage of a small comparator. This could be rotated at moderate speed by a small motor, and the ruling point, a gramophone needle carried on a pivot, depressed by pulling a string, until it came in contact with a glass plate coated with a thin film of wax mounted on the turn-table thus cutting out a narrow ring circle. A reduced photographic copy of the drawing was mounted on a strip of glass extending over the edge of the carriage, and this was watched with a microscope provided with a cross hair. The carriage was advanced one zone at a time, and the concentric zones traced with the steel point. For the outer zones a single contact is sufficient for each, but as the zones increase in width, the point is kept in contact with the revolving plate during the advance of the carriage. The waxed plates are prepared by dipping narrow strips of thin plate glass in a solution of 10 grams of black etching wax in 100 c.c. of benzene. Bitumen would probably answer the purpose. The film is very nearly perfectly transparent. If of the right thickness, a candle flame shows brilliant spectra, and a colored central image when viewed through the edge of the plate held close to the pupil of the eye.

Reflection and Refraction by Unpolished Surfaces. — One of the most interesting and instructive applications of the Fresnel construction is to the diffuse reflection and refraction which occur when light strikes unpolished or matt surfaces, such as paper, plaster of Paris, or ground glass. An unpolished surface destroys all phase relation between the elements on the wave-front. The secondary wavelets start from the elevated portions of the surface first, since these portions are struck first by the incident wave, and the reflected wave-front, instead of being plane, is pitted and corrugated in an irregular manner. It is impossible to arrange any zone system on such a surface, for there are all possible phase differences irregularly distributed over the reflected wave-front, consequently each point on the surface acts as an independent luminous source, sending light out in all directions. We can apply the Fresnel theory to reflection of this sort in the following way.

Suppose we have a plane surface XY (Fig. 27) and a luminous point S, and are considering the effect at P, which we will suppose to be the point to which a ray SA would be reflected. We have, however, at A an elevation of height H, and the secondary wavelet will leave the point B sooner than it would have left the point A were the elevation absent. We can see that the effect at P will be the same in either event, provided the difference between the path

SBP and *SAP* is small in comparison to the wave-length. At normal incidence it is obvious that this path difference will be $2H$, therefore a surface having elevations on it of such magnitude that twice their height is not small in comparison to the wave-length will not reflect regularly at normal incidence. With a given roughness long waves may be regularly reflected, and short waves irregularly. It can be seen from the right-hand figure that the difference

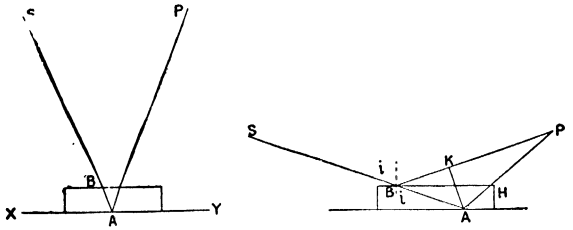


FIG. 27

of path becomes less as the angle of incidence increases, being in the case figured $BA - BK$, which is less than H, and that at grazing incidence it will become 0.

It can be shown geometrically that the path difference is represented for all incidences by $2H \cos i$, the value of which must not exceed a small fraction of a wave-length if regular reflection is to occur.

$$SB = SA - AB = SA - \frac{H}{\cos i}$$

and $$BP = AP + BK = AP + AB \cos(\pi - 2i) = AP - \frac{H}{\cos i} \cos 2i,$$

$$SA + AP - (SB + BP) = \frac{H}{\cos i}(1 + \cos 2i) = 2H \cos i.$$

Since the path difference decreases as the angle of incidence increases, it is obvious that for a given roughness we shall get regular reflection when the incidence angle is so great that $\rho\lambda = 2H$ cosine i, where ρ is a small fraction; therefore if we gradually increase the incidence angle, the long waves will be reflected first, and then the shorter. Smoked glass, which at perpendicular incidence will show no image of a lamp at all, will at nearly grazing incidence give an image of surprising distinctness, which is at first reddish, becoming white as the angle increases.

Let us consider next the effect of a matt surface on refraction. Here the phase differences are due to retardations of the portions of the wave-front encountering the elevations, on those portions encountering the depressions. With a given degree of roughness the retardation will be greater when the substance has a high re-

fractive index, or more accurately when the difference between the refractive indices of the media bounding the rough surface is large. When the retardation between two adjacent paths is larger than a small fraction of a wave-length, we have diffuse transmission. If we take a sheet of ground glass and wet the surface, the glass transmits more direct light than it did before, since we have lessened the difference between the refractive indices of the bounding media. If we substitute benzole for water the glass becomes still more transparent, and by bringing up the refractive index of the benzole by an addition of Canada balsam, we can cause the ground surface to disappear entirely.

It can be shown easily that the regularity of transmission decreases as the angle of incidence increases, the opposite of what we found in the case of reflection.

If the refractive index of the substance is 1.5, then $H(1.5-1)$ or $H/2$ must be small in comparison to the wave-length of light, if the light is to be regularly transmitted at perpendicular incidence. Inequalities can then exist, the heights of which are, say, not greater than $\frac{1}{4}\ \lambda$, which is four times as great a discrepancy as we could have on a reflecting surface.

Summing up, we have (for perpendicular incidence) for regular reflection, $2H=\rho\lambda$ and for transmission $H/2=\rho\lambda$.

If we procure a piece of ground glass, which will barely show the outline of a lamp flame by transmitted light, and thinly silver a portion of the ground surface, we shall have a reflecting and transmitting surface of the same degree of roughness. It will be found that by reflected light the outline of the flame is indistinguishable. Ground glass of this description may be made by grinding two pieces of ordinary ground glass together, with fine emery and water, the process being the first stage of polishing.

We thus see that a rough surface may regularly reflect the long waves while diffusing the shorter ones. Lord Rayleigh has made some interesting experiments upon the reflection of heat-waves from ground-glass surfaces too rough to give any trace of regular reflection with visible light. The ground surface was silvered and the radiations of a Welsbach lamp reflected from it. In some cases two reflecting surfaces were used. It was found that the radiation, freed by the process from the shorter waves, was reflected almost as well by a third ground and silvered surface as by a polished silver mirror. The method is analogous to that originated by Rubens and Nichols for isolating long heat-waves by repeated reflection from quartz or rock-salt surfaces.

A paper by Lord Rayleigh on "Polish" in the *Philosophical Magazine* will be found of interest.

CHAPTER II

THE REFLECTION OF LIGHT FROM PLANE AND CURVED SURFACES

When light strikes the boundary surface separating two media of different optical densities, some of the energy is reflected back into the first medium, and some crosses the boundary and is transmitted through, or absorbed by, the second medium. We have shown in the previous chapter that if the surface is smooth to within one-eighth of a wave-length, we shall have regular reflection, but as a matter of fact, we are practically unable to make a surface so perfect that absolutely no light is diffused. Admit a ray of sunlight into a dark room and concentrate it with a lens on the surface of the most perfect mirror attainable; were diffuse reflection not present the illuminated spot on the mirror would be invisible, which is never the case.

Even with the most highly polished optical glass, or a surface of the cleanest mercury, the small patch on which the concentrated sunlight falls is faintly visible in all directions: The slight irregularities due to the molecular structure of the surface preclude the *complete* destructive interference of the secondary wavelets in the region outside of the reflected beam. The nearest approach to complete reflection and transmission occurs in the case of a very thin, freshly split, flake of mica, in which the perfect regularity in position and spacing of the molecules on the crystal lattice ensures almost complete absence of diffusion. This case will be more fully discussed in the Chapter on Scattering of Light.

The percentage of diffused light decreases as the angle of incidence increases, as has been shown in the previous chapter, regular reflection taking place even on matt surfaces at grazing incidence.

In studying the reflection of light from plane and curved surfaces we shall investigate not only the direction of the reflected rays, but also the form of the reflected wave-fronts.

Reflection of Waves and the Formation of Images. — The study of the phenomena of the reflection of light-waves from plane and curved surfaces in their relation to the formation of images is of importance, since many optical instruments operate with mirrors instead of lenses. Any finite source of light such as a lamp filament or an illuminated surface may be regarded as a collection of lumi-

nous points distributed over a surface. In the formation of a real image by a mirror or lens we can consider that a portion of the spherical wave, expanding from a single luminous point, is transformed by reflection or refraction into a contracting sphere which collapses to a point, thus forming a single element of the image. The wave-fronts emanating from other points on the source are incident upon, and reflected from, the mirror surface at different angles and consequently shrink to points distributed in space in accordance with their points of origin. The *image*, considered from the wave-front point of view, can thus be regarded as a space distribution of points at which spherical wave-fronts are collapsing. We will now consider the nature of the wave-front reflected from surfaces of different type.

Reflection of a Plane-Wave from a Plane-Mirror. — Here the incident rays are parallel to one another and normal to the wave-front. We have seen that the reflected wave-front will be the surface enveloping the secondary wavelets given off by the various points on the mirror's surface as they are struck in succession by the incident wave. The general method of constructing the reflected wave-front is shown in Fig. 28. Let AB' be the surface of the mirror, and AB the incident-wave-front, the rays being indicated by lines. At the moment figured, a secondary disturbance is about to leave the point A. This secondary disturbance will have spread out all around A, to a distance equal to BB' at the moment when the point B on the wave-front encounters the surface. The secondary wavelets from points C, D, E, etc., intermediate between A and B', will have lesser radii. To construct them draw $A'B'$ parallel to AB. This will give us a subsequent position of the wave-front, assuming the *mirror not present.* Join these two wave-fronts by perpendicular lines, which represent rays, which cut the mirror's surface at C, D, E, etc. The wavelet around C must obviously have a radius equal to CC', while that around D has a smaller radius, DD', and so on for all the other points. If we describe these spheres (circles in the diagram) we shall find that they are enveloped by a plane surface, which makes the same angle with the mirror's surface as the incident wave; since we have similar right triangles with a common hypotenuse.

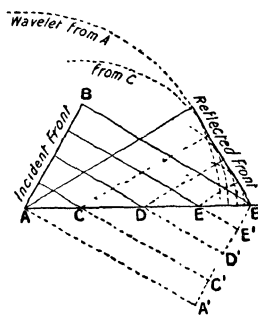


FIG. 28

If we apply the same construction in the case of a spherical wave, we find a reflected wave of spherical form radiating from a point

as far below the mirror as the source is above. An image of this type is called "virtual," having no real existence, as in cases to be considered presently.

Reflection of Sound-Waves. — A striking analogy exists between sound and light; we can show nearly all of the phenomena of reflection, refraction, and diffraction by means of sound-waves. An electric spark is the centre of a spherical sound-wave, which expands at the rate of about 331 metres per second, and at the same time is the centre of spherical light-waves, which expand at the rate of 300,000 kilometres per second. We have no means of directly showing the form of the wave-front of the *luminous* disturbance. We can calculate its form before and after reflection and show the agreement between these forms and the rays as actually observed, but we cannot actually show the wave-front. In the case of the air disturbance, however, the wave-front cannot only be seen, but photographed. We are dealing with a spherical shell of condensed air and, by a suitable optical contrivance which will be described in the next chapter, we can study at our leisure the changes which the wave-front undergoes.

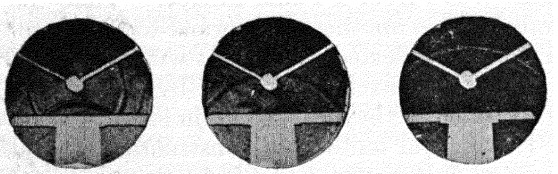


FIG. 29

The author has prepared an extensive series of photographs of sound-waves for the purpose of illustrating optical phenomena.

The case that we have just considered, namely, the reflection of a spherical wave from a plane surface, is shown in Fig. 29. The sound-wave is started by an electric spark which has just passed between two brass balls, seen in line, one behind the other at the centre of each picture. The wave of condensed air is illuminated and photographed by the light of a second spark occurring a moment later. By properly regulating the time interval between the two sparks a progressive series of views is obtained showing the wave-front at different stages of its development.

The form of the reflected wave or echo is seen to be identical with the form of the light-wave as calculated by Huygens's principle.

Mirrors at Right Angles: Return of Ray to Source. — Rays incident on two mirrors meeting at a right angle, in general suffer two reflections, some of them returning to the source. This is true for all rays in the plane of the paper in Fig. 30*a* parallelism of the incident and twice reflected rays being evident from most elementary geometry. Rays are returned to the source regardless of

the direction from which they come, by three mirrors forming the corner of a cube (Fig. 30*b*). Such a mirror can be made by cutting off the corner of a solid cube of glass, and polishing the cut surface, reflection in this case being of the "total internal" type. An easier method is to cement three square mirrors together with sealing wax. Mosaics of such glass reflecting prisms were designed for naval purposes during the war. The flagship of a squadron could ascertain the position of the other vessels at night by sweeping the horizon with a searchlight (of too low intensity to render them visible by illumination), and the rays reflected by the cubic mirrors were visible only from the flagship, since they were returned in this direction regardless of the exact orientation of the mirror mosaic. Similar mirrors, more crudely made, are in common use as "danger" warnings for motor cars, the signs reflecting back the light of the car's lamps.

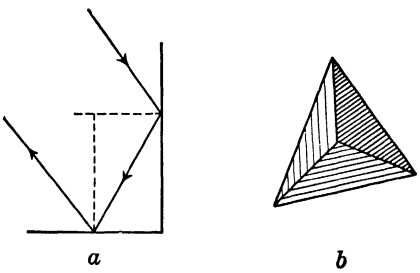


FIG. 30

Reflection by Ellipsoidal Mirror. — If a spherical wave start at one focus of an ellipsoid of revolution, the reflected wave will be spherical in form, and will collapse to a point at the other focus, or rays of light issuing from one focus come accurately together at the other focus. A surface capable of bringing rays of light accurately to a focus, either by reflection or refraction, is said to be aplanatic, consequently an ellipsoidal mirror is aplanatic for rays issuing from a point situated at either focus. This can be shown by the following construction.

Around one focus of an ellipse describe a circle which falls just outside of the furthest extremity of the ellipse. Draw a number of radii to the circle, and around the points where the radii cut the ellipse describe circles with radii equal to the distances from the respective points to the outer circle (measured along the radii). The circles will be enveloped by another circle (the reflected wave-front), the centre of which is at the other focus of the ellipse.

Shadow photographs of the reflection of a sound-wave in an elliptical mirror by Foley [1] are shown in Fig. 31. The expanding sphere is seen to have been changed by reflection into a contracting sphere, which shrinks to a point at the other focus. The same thing can be shown by making a shallow, flat-bottomed, elliptical dish of wood, filling it with mercury, and touching the surface of the fluid at

[1] *Phys. Rev.*, *35*, 373, 1912.

one focus of the ellipse. By placing the dish in the sunlight, and receiving the reflected light on a screen, the experiment can be shown to a class.

The ellipsoidal mirror has one interesting property, to which attention has not apparently been drawn. It collects the energy of a widely divergent system of rays, compressing them into a cone of much smaller angular aperture. This is clearly shown by Fig. 32*a* in which *A* is the source and *B* the image. Here a cone of rays of 180° aperture (hemispherical wave) has, after reflection, been reduced to a cone of about 40°, and after passing through a focus at *B* diverges and falls on a lens *L* so placed as to render

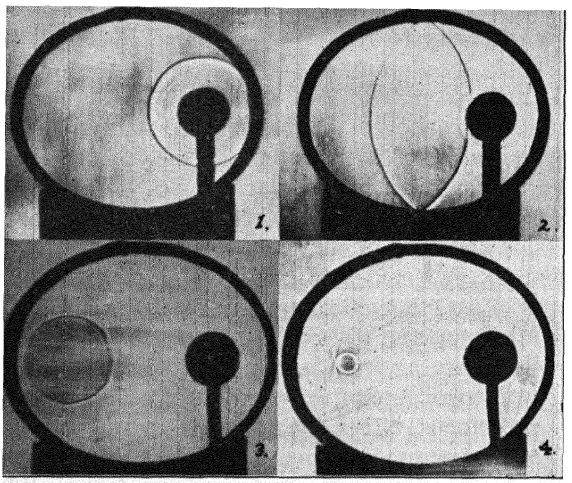


FIG. 31

them parallel. It is clear now that we shall be projecting more light through the lens *L* than if we dispensed with the ellipsoid and placed the source at *B*. In the former case we utilize half of the total light emitted by the source (180° cone), in the latter only that contained in a cone of 40°. One's first impression would be that nothing would be gained by substituting an *image* for the actual source. Looking at the matter from the wave-front point of view, we find that there is an enormous concentration of amplitude on one side of the contracting spherical wave-front which is collapsing to the focus, as shown by Fig. 32*b* in which the amplitude (intensity) is represented by the width of the line, the lowest intensity being to the left, *i.e.* away from the source *A*. The energy of one-half of the spherical wave expanding from *A* is, after reflection, located on a wave-front in the 40° cone at the

right of B. At first sight a difficulty may appear if we consider that the distances between A and the sphere around B, via the reflecting surface are all equal. It is suggested that the student ask himself the question (before reading further), "Why is not the amplitude everywhere the same over the sphere around B, if it is everywhere equidistant from the source?" (measured along the rays of course). The difficulty is only apparent, for we have neglected to take into consideration the fact that the elements of the wave are expanding *before* reflection, and contracting after reflection. In the former case the amplitude decreases as the wave advances, in the latter it increases. The element of wave reaching B via D is decreasing in amplitude over the long path AD, and increasing over the shorter path DB. The reverse is the case for the element of wave going via C.

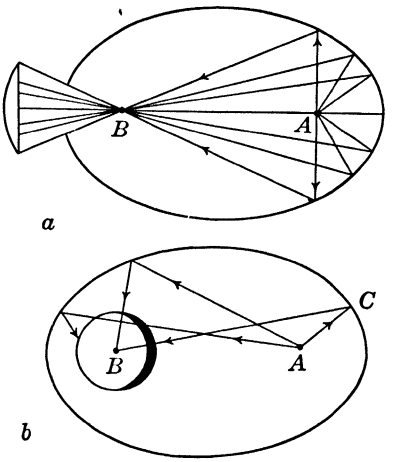


FIG. 32

At first sight it may appear that this arrangement violates the law that an image can never be brighter than the source, but if we take into account the fact that the image is larger than the source, — magnification resulting from the circumstance that the image is further away from the right-hand side of the mirror than the source, the difficulty about intrinsic brilliancy vanishes. In the case of a single radiating atom at the focus the image will be a diffraction disk enormously larger than the atom.

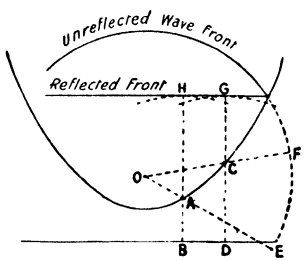


FIG. 33

Reflection from a Parabolic Mirror. — If we construct the reflected wave-front in the case of a wave starting at the focus of a parabolic mirror, we shall find that the reflection transforms the spherical wave into a plane-wave. The reflected rays, being normal to the wave-front, are parallel, and are consequently projected in a narrow beam out of the mirror. This is the principle on which the naval searchlights are constructed.

Let O be the focus of the parabolic section of the mirror (Fig. 33).

Construct the imaginary spherical wave-front (unreflected) EF, and around points A, C, etc., on the parabola construct circles, or secondary wavelets, with radii equal to the distances of the points from the imaginary wave-front measured along the radii of the circle EF.

The wavelets will be enveloped by a straight line, the section of a plane. It is easy to see that this line is straight, or that the reflected wave-front is accurately plane. Every point on the parabola is equidistant from the focus and the directrix; $\therefore OA = AB$ and $OC = CD$. Around A and C we have circles with radii equal to AE and CF respectively. Now, $OE = OF$, being radii of same circle, and $DG = OF$ and $BH = OE$; $\therefore DG = BH$ or the enveloping line is everywhere equidistant from the directrix, and consequently parallel to it. The reflected wave is, therefore, a true plane. The projection of a truly plane-wave from a paraboloid mirror is practically never realized, since the source of light is always of finite size — that is, the waves do not all start from the exact focus.

The amplitude of the plane-wave passing out of the paraboloid is greatest on the axis decreasing as we pass to portions of the wave reflected from parts of the mirror lying at greater distances from the source. This is the same effect as that just considered for the ellipsoid.

In Fig. 34 we have the reflection of a sound-wave from a parabolic mirror. The converse of this case is also true. Plane-waves

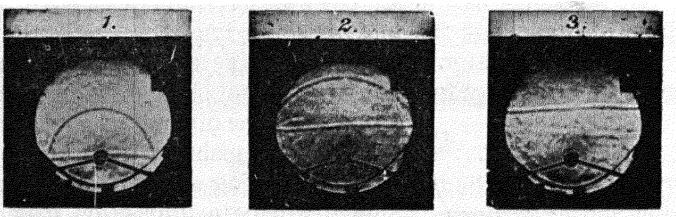


FIG. 34

entering a parabolic mirror are transformed by reflection into converging spherical waves which shrink to a point at the focus of the paraboloid. This means that parallel rays, or rays coming from an object situated at a great distance, are brought accurately to a focus by a mirror of this form, or the parabolic mirror is aplanatic for parallel rays.

The surface of a liquid in uniform rotation assumes the form of a paraboloid under the influence of centrifugal force, and the author once constructed a reflecting telescope of mercury twenty inches in diameter which operated on this principle. The focal

length could be varied from three to twenty or more feet by altering the speed of rotation. The instrument resolved stars three seconds of arc apart, showed the small craterlets on the moon, and yielded wonderfully bright images of nebulae when running with a short focus. It was, however, merely a scientific curiosity. See *Astrophysical Journal*, 1909 and 1910.

Projection of Narrow Beams to Great Distances. — A system of secret signalling by means of very narrow light beams was developed by the author during the war. The optical projectors for flash signals universally employed at the time were similar to automobile headlamps and gave a beam of considerable angular aperture. They could not, therefore, be used for signalling from the rear to advanced positions. The principles upon which optical projection depended did not appear to have been generally understood, for very short focus mirrors of rather large aperture were employed, under the mistaken impression that it was necessary to collect the rays within as large a cone as possible and render them parallel, in order to reach great distances. As a matter of fact the focal length of the mirror (or lens) does not enter the problem at all. The aperture is the only thing that counts. This will be clear from the following consideration. Suppose we wish to employ a mirror or lens three inches in diameter, then, if we use a true point source, placed exactly at the focus, a beam three inches in diameter will be projected to infinity with no lateral spreading (except the small amount due to diffraction). Our source, however, must have some finite size, and the various elements of its surface project a complex of three-inch parallel beams inclined slightly to each other, the *inclination increasing as the focal length of the mirror or lens is decreased.* A greatly enlarged image of the *source* is thus formed on a distant screen, which is made up of the three-inch disks formed by the aggregate beams of all of the true points of the source. This image will increase in size as the focal length of the mirror is reduced. We gather more light, but we distribute it over a larger tract of country and nothing is gained in visibility. The observer's eye sees the entire aperture of the mirror filled with light of the same intrinsic intensity as that of the source, therefore the greater the diameter of the projector, the greater the distance at which it is visible. The instrument as actually constructed for military purposes consisted of an achromatic telescope of three-inch aperture and 18-inch focus with a small concentrated filament incandescent lamp mounted exactly at the focus, just in front of the eye-piece. This projected a beam, the width of which at a distance of two miles was only about ten feet. On looking into the eye-piece, the

distant landscape, highly magnified, and the concentrated tungsten filament of the lamp were both seen in good focus. The telescope was aimed by bringing the point at which the flash signals were to be directed into coincidence with the filament. The principle of the reversibility of rays then ensured that the flashes reached the desired point. This lamp was visible in field glasses, in full sunlight, at a distance of twelve miles when operated by an eight-volt battery. With two observers ten feet apart at a distance of a mile, it was possible to flash signals to one that were quite invisible to the other.

Formation of Images by Parabolic Mirrors. — Parabolic mirrors are now universally used for astronomical photography. The time of exposure depends upon the ratio of the aperture of the mirror to its focal length. It might appear, at first sight, that the time of exposure could be reduced to any degree by employing a sufficiently deep paraboloid. It is true that the amount of light sent to focus can be increased indefinitely by building up the paraboloid, but this light will not all go to the formation of a single image as is clear from Fig. 35, which shows that the image formed by the portion of the paraboloid at B is inclined at 90° to the image formed by the portion at A, for the image is perpendicular to the rays. It is thus clear that as soon as we attempt to increase the aperture of the mirror, distortion of the image occurs, not only as a result of this variable inclination, but also from the circumstance that the size of the image varies for different parts of the mirror since the focal length of an element at B is greater than that of one at A.

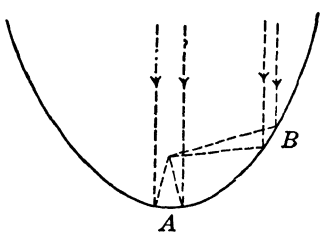


FIG. 35

If we are concerned only with the concentration of energy, as when employing a mirror for producing high temperatures with focussed sunlight, the paraboloid should be made as large as possible for a given focus. Thin platinum foil can be melted by sunlight concentrated with a searchlight mirror 3 feet in diameter and 18 inches focus. Much higher temperatures could be secured by employing larger mirrors of the same focal length, *i.e.* by extending the paraboloid surface. The diameter of the sun's image formed by the mirror is about 8 mms. but will increase rapidly as we utilize portions far removed from the vertex, the diameter of the image for any zone being proportional to the distance of the zone.

Reflection by Hyperboloid. — A spherical wave originating in one focus of an hyperboloid is, by reflection, changed into a sphere

whose centre is at the other focus. The reflected rays appear, therefore, to come from this point, each focus of the hyperboloid being the virtual image of the other. This follows at once from the fact that lines joining the two foci of the hyperbola with any point on the curve make equal angles with the normal at the point. Let F and F' be the two foci (Fig. 36), and consider any ray drawn from F' to the branch of the hyperbola nearest it, say to the point A. Then FA and $F'A$ make equal angles with the normal, and the reflected ray AB produced backwards coincides with AF. A spherical wave started at F' will, after reflection, be a sphere with its centre at F. If a wave is started at F and reflected in the *convex* surface of the same branch of the hyperbola (the branch nearer F being removed), it will be transformed into a sphere expanding from F' as a centre.

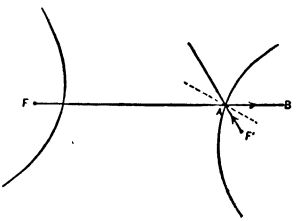


FIG. 36

If now we consider a spherical wave-front contracting towards the focus F' (*i.e.* reversing the last case) it will, after reflection, converge towards F. In other words, the convergence of the rays will be decreased without interfering with their ability to come to a focus at a point. This is important in connection with the use of hyperboloidal mirrors in modern reflecting telescopes.

Reflecting Telescopes. — The earliest proposal for a reflecting telescope was made in 1663 by James Gregory, who, however, did not succeed in making a practical instrument. His scheme is of interest as it involved the use of both the paraboloid and ellipsoid. Its construction is shown in Fig. 37, *1*. A is a parabolic mirror, perforated at its centre, with a focus at B. At C is a concave ellipsoid with its conjugate foci at B and D. It therefore forms an image at D of the image at B formed by the large mirror. Telescopes of this type were subsequently manufactured by James Short of Edinburgh in 1732 which gave good results, being known as Gregorian telescopes. The image at D was viewed by an eyepiece through the perforation in the mirror.

The first successful reflecting telescope was, however, constructed by Newton about 1670. Previous instruments had been made with single lenses of very long focus, to lessen the chromatic aberration, and as Newton's investigations of dispersion had convinced him of the impossibility of constructing an achromatic lens, he turned to the concave mirror, the focal length of which is independent of the color of the light. He placed a small mirror on the axis between the objective and the image (near the latter),

which reflected the rays out through the side of the tube, where the image was viewed with an eye-piece as in Fig. 37, *2*. Telescopes of this form are called Newtonian reflectors. The mirror, however, obstructs some of the incident light, and Herschel accordingly inclined his mirror so that the image fell at the side of the tube. This produces a little distortion, however, though it is small if the inclination is not more than two or three degrees. If we could make a paraboloid in which a portion of the true surface a little to one side of the axis was used, the mirror when properly oriented would give an image free from distortion in an oblique direction. See section on Focal lines. This could be accomplished by local corrections of the surface. In practice it is found that the definition of the Herschel reflector is best with the mirror in a certain position, which can be found by rotating it in its own plane. In this position we have the nearest approximation to the ideal condition mentioned above. The circumstance results from slight irregularities in the curvature.

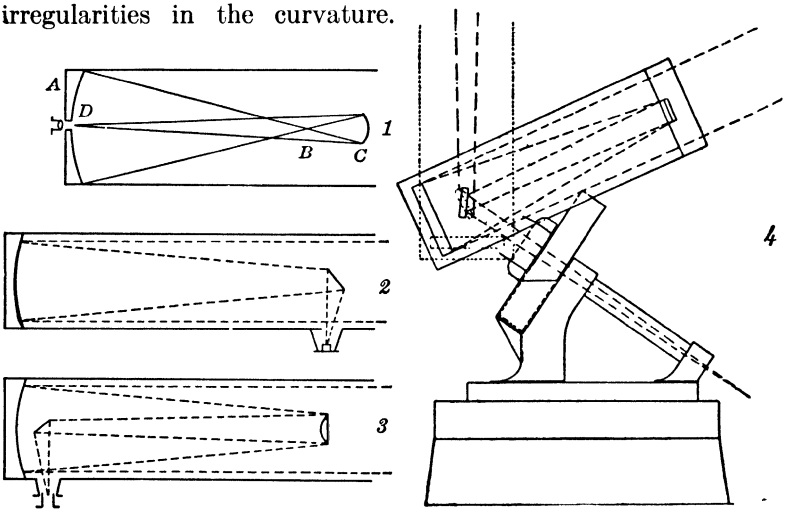


FIG. 37

A second type of instrument was devised by Cassegrain who placed a convex mirror of long focus between the paraboloid and its focus, as in Fig. 37, *3*. This decreased the convergence of the rays which come to a focus in the aperture of the large mirror, as in the Gregorian type. The convex mirror must be an hyperboloid as we have seen. The largest reflecting telescope in the world is the monster instrument of the Mt. Wilson Observatory in California, with a mirror of 100.4 inches diameter and a focus of 43 feet. Its construction is similar to that of its predecessor, with

a 60-inch mirror of 25-foot focus. This instrument can be used first as a Newtonian, as shown in Fig. 37, *2*; second as a Cassegrainian with a focus equivalent to 100 feet, Fig. 37, *3*; and third as a Cassegrainian with a focal length of 150 feet. To transform it into the Cassegrain type a convex hyperboloid is put in place of the plane-mirror used in the Newtonian type. This decreases the convergence of the rays, which are either reflected to one side at the base of the tube, as in Fig. 37, *3*, or thrown down through the hollow polar axis as in Fig. 37, *4* (150-foot focus). In using the hyperboloidal mirror, the actual distance traversed by the rays in coming to a focus is not 150 feet, but the size of the image is the same as that produced by a mirror having an actual focal length of 150 feet.

Reflection from Spherical Surfaces. — If the reflecting surface be a portion of a sphere, the effects are more complicated, except in the special case of waves starting at the centre of curvature. The rays do not all meet at a point, as in the cases which we have considered, but envelop a surface known as the *Caustic*. An example of a caustic is the cusped line of illumination seen on the table-cloth when the light of a lamp strikes the inner surface of a silver napkin ring. We have seen that a concave paraboloid brings parallel rays accurately to a focus. A concave spherical mirror does not do this. Rays near the axis come to a focus approximately at a point, but as we recede from the axis we soon find the reflected rays falling wide of the focus. This effect is known as Spherical Aberration. In constructing telescope mirrors, opticians strive to give the surface as nearly as possible the figure of a paraboloid. The nature of the wave-front in cases where caustics are formed is not at once apparent. The subject is usually treated by ray methods, and we shall accordingly begin by considering one or two examples geometrically, although the evolution of the wave-front, and the relation between the wave-front and the caustic, form a more interesting study.

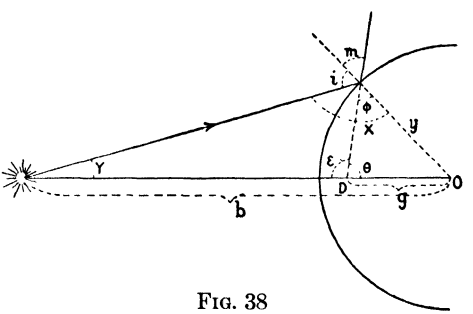


FIG. 38

Reflection from Convex Spherical Surfaces. — In studying reflection by ray methods we can regard a curved surface as made up of an infinite number of plane surfaces, for each one of which

the law of equal angles holds. Let us take as the first case the reflection of light radiating from a luminous point by a convex spherical mirror.

When the light is incident on a small portion of a sphere in a nearly normal direction, we may regard the reflected rays as emanating from a point behind the mirror. If, however, we employ a large arc of the mirror this does not hold even approximately, and we require an expression for the position of this focus in terms of the angle of incidence.

Let the radius of curvature of the mirror be y, the distance of the radiating point from the centre of curvature b, and the angle of incidence i; find g, the distance of the focus D from the centre (Fig. 38). It can be shown by elementary trigonometry (see earlier edition) that

$$g=\frac{by}{2b\cos B-y} \quad \text{for CONVEX MIRROR.}$$

Considering D as the luminous point, and solving for b, we get

$$b=\frac{gy}{2g\cos b-y}, \text{ the equation for a CONCAVE MIRROR.}$$

These equations show us that the rays reflected from different annular elements of the sphere (determined by B) if produced backwards cut the axis of the mirror nearer and nearer to its surface as B is increased.

For rays near the axis ($B=0$ and $\cos B=1$) we have

$$g=\frac{by}{2b-y}.$$

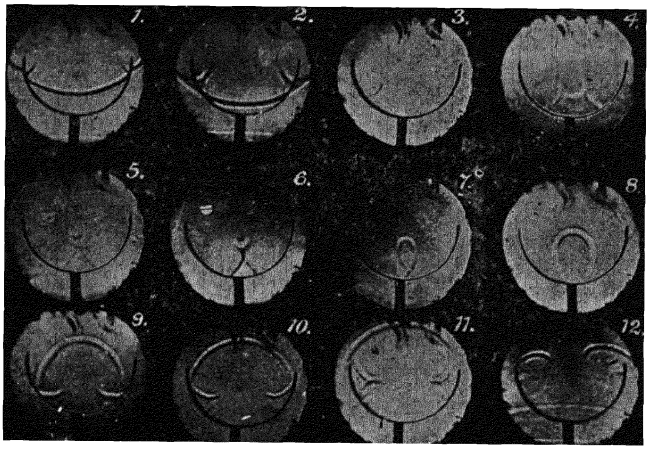
If the source is at a great distance, *i.e.* b very large in comparison to y we have

$$g=\frac{y}{2};$$

that is, the focus is midway between the surface and the centre of curvature. If we construct a number of reflected rays we shall find that they envelop a caustic surface, which is virtual in the case of a convex spherical mirror and real in the case of a concave. We cannot therefore form a clear image with a spherical mirror of large aperture, since the rays do not focus all at the same point; in other words, the reflected wave is not spherical as in the case of the ellipsoid and paraboloid.

Reflection of Plane-Waves from Concave Spherical Mirrors. — The author became interested in the problem in 1898 as a result

of a study of photographs made of sound-waves entering a hemicylindrical mirror. The reflected wave-fronts were of quite intricate form, a series of them being reproduced in Fig. 39. The reflected front is cusped, and in one stage (No. 4) has a form not unlike that of a volcanic cone with a bowl-shaped crater. It appeared to be worth while to construct reflected wave-fronts by geometry, showing the front in successive positions and then photograph the drawing on a kinematograph film which, when projected, would show the movement of the wave and its transformations in a graphic manner. The method already described has certain disadvantages in the present case; it does not readily give us the



complete wave-front, and it gives but a single front. In this particular case the reflected front is rather intricate, and a construction that will enable us to follow it in its evolution is desirable. This can be accomplished by employing a second method.

ABC is the mirror, *AOC* the plane-wave (Fig. 40). Around points on *ABC* as centres describe circles tangent to the wave. These circles will be enveloped by another surface, *ADE*, below the mirror (the orthogonal surface). If we erect normals on this surface, we have the reflected rays, and if we measure off equal distances on the normals, we have the reflected wave-front. By drawing the orthogonal surface we avoid the complication of having to measure off the distances around a corner. The orthogonal surface is an epicycloid formed by the rolling of a circle of a diameter equal to the radius of curvature of the mirror on the mirror's sur-

face, and the normals can be erected by drawing the arc *FG* (the path of the centre of the generating circle), and describing circles of diameter *BE* around various points on it. A line joining the point of intersection of one of these circles with the epicycloid, and the point of tangency with the mirror, will, when produced, give a reflected ray; for example, *JK* produced, for circle described around

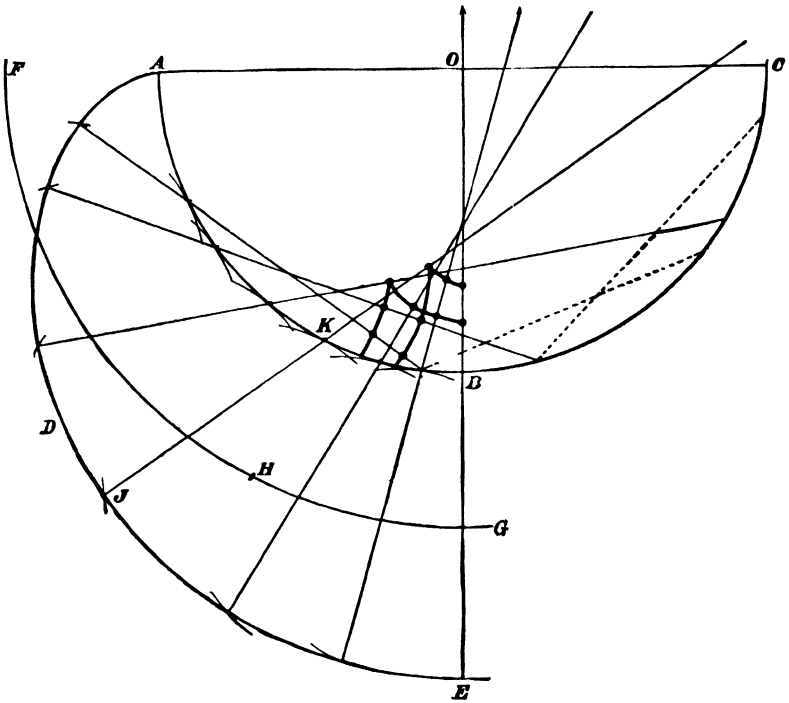


FIG. 40

H. This construction once prepared, the series of wave-front pictures can be very quickly made. Three or four sheets of paper are laid under the construction and holes punched through the pile by means of a pin, at equal distances along each ray (measured from orthogonal surface).

The centre of the mirror and the point where its axis meets the surface are also indicated in the same manner. The sheets are now separated, and corresponding pin-holes are united on each sheet by a broad black line, which represents the wave-front. After a time it becomes necessary to consider double reflections, and to do this we are compelled to construct twice-reflected rays (indicated by dotted lines), and measure around a corner each time. The results

obtained are shown by Fig. 41. These are diagrams taken at intervals on a kinetoscope film prepared by the author for illustrating the wave evolutions.[1]

About one hundred constructions were made, in the manner just described, and photographed in succession on the film, which

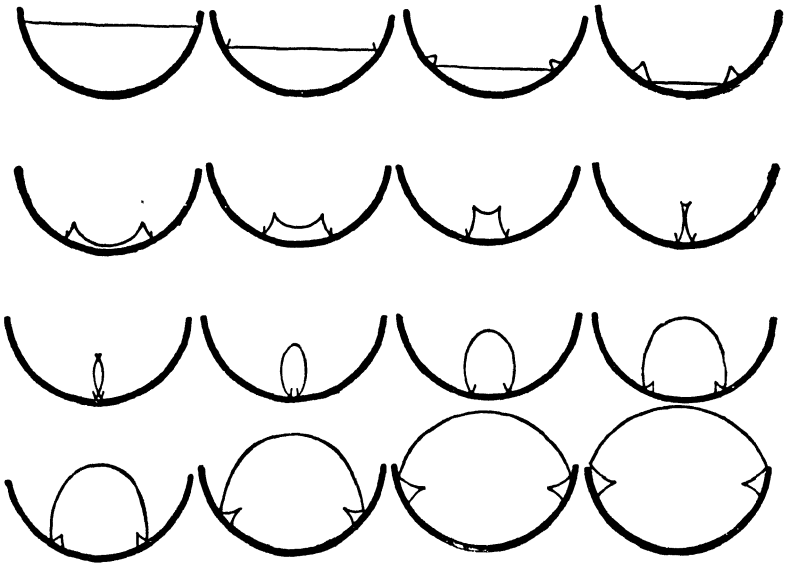


FIG. 41

was shown at the Royal Society in 1899, at which time there were only two moving-picture machines in London, the one used being loaned by Robert Paul. This film may be regarded as the ancestor of the modern cinema cartoons.

In Fig. 42 we have a number of wave-fronts in different stages of reflection, and it is at once apparent that the cusp traces the caustic surface, indicated by a dotted line.

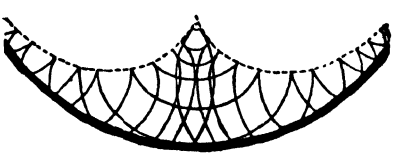


FIG. 42

This gives us at once a *physical*, as distinguished from a *geometrical*, definition of a caustic, which is a surface traced by a moving cusp of the wave-front. The curvature of the crater increases as we go from the bottom to the rim, at which point the radius becomes zero. The inner edge is then continually passing through a focus and ap-

[1] Wood, "Photography of Sound-Waves, and Kinematographic Demonstration of Reflected Wave-Fronts," *Proceedings Royal Society*, vol. 66.

pearing on the outside building up, as it were, the sides of the cone. The cusp can thus be regarded as a moving focus, where there is a concentration of energy, and since it travels along the caustic, the increased illumination along this line is accounted for.

Photographs of the saucer-shaped reflected wave-front in the case of a spherical wave starting at the focus of a spherical mirror

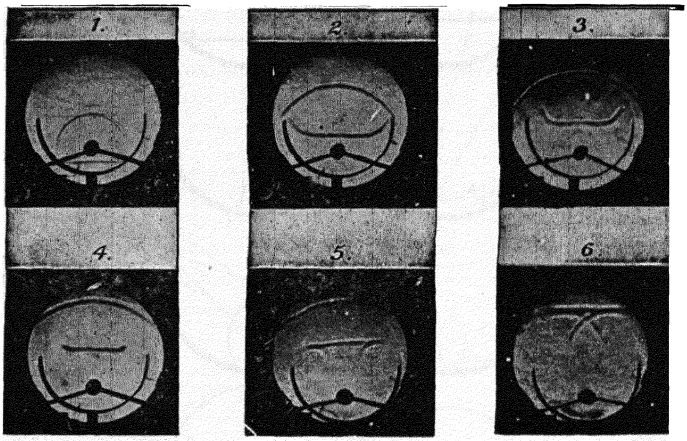
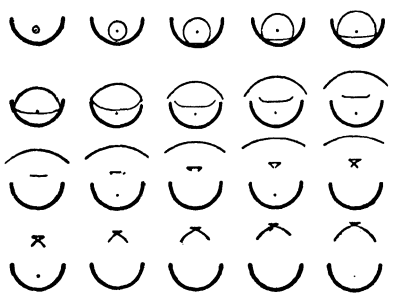


FIG. 43

are reproduced in Fig. 43. A parabolic mirror reflects a plane-wave as has been shown, while in this case no portion of the wave is truly plane, and its area gradually contracts as it advances, its rim being a moving focus which traces a cone-shaped caustic as shown in Fig. 44. This can be demonstrated by



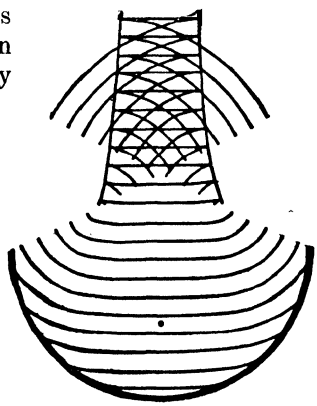


FIG. 44

silvering the outside of a hemispherical glass evaporating-dish or half of a large, round-bottomed flask. The concave mirror thus formed should be mounted on a stand, and a two-candle "pea"

electric lamp arranged so that it can be moved along the axis of the mirror.

If we place the lamp in the focus of the mirror, and hold a sheet of ground glass in front of it at the proper distance, we can show the luminous ring formed by the passage of the sides of the saucer-shaped wave through a focus.

The illumination within the ring is due in part to unreflected light, and in part to the flat portion of the reflected wave.

Formation of Images. — In the formation of a real image by a mirror, the rays radiating from a point are brought together again approximately at a point, or the reflected wave-fronts are converging spheres. The formation of these images, and the study of their position and distance from the mirrors, belong to geometrical optics. The influence of the form of the wave-front on the distinctness of the image may, however, be considered in connection with what has gone before. We have seen that the paraboloid and ellipsoid are the only surfaces that reflect spherical fronts; other curved surfaces give, in general, caustics. If we use only a very small portion of the sphere, and receive the light normally, we have only the cusp of the caustic, which is very nearly a point. There is a certain amount of what is called longitudinal aberration, owing to the fact that the rays reflected from the periphery come to a point situated nearer the mirror than those reflected from the points near the axis. If the image be thrown on a screen, the bright spot formed by the central rays will be surrounded by a circular ring of light formed by the peripheral rays which have already come to a focus and are diverging once more. The position of sharpest definition lies between the focal points of the marginal and central rays. The circle of light is smallest at this point, and is known as the circle of least confusion. The line AB in Fig. 45 shows the position of this point.

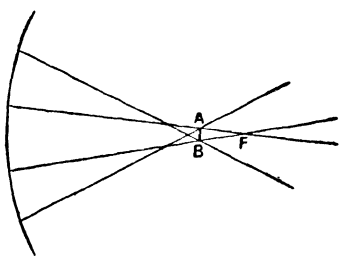


FIG. 45

Focal Lines. — Suppose now that the incident light falls on the mirror in an oblique direction. This is the condition if we consider a small portion of the hemispherical, concave mirror at some distance from the axis. The reflected wave will come to a focus in a line instead of a point; as we increase the distance of the screen from the mirror, the line will decrease in length, increasing in width until it is transformed into a line at right angles to the first.

These lines are known as the primary and secondary focal lines

respectively, and can be shown by holding a concave mirror in an oblique position and reflecting the light coming from a small, brilliant source on a screen placed at various distances from the mirror. We can best form an idea of how these lines are formed by considering the question first by a ray method and then by a wave-front method.

Let AB be the axis of the mirror, near the edge of which a bundle of rays parallel to the axis falls. Construct the reflected rays from a linear strip BC as shown in the sectional view (Fig. 46): we shall have a flat, converging fan coming to a focus at F', then diverging and cutting the axis at F''. Now rotate the whole figure through several degrees around AB as an axis; the parallel sheet of incident rays will trace the rectangular incident bundle, the line BC will trace an approximately rectangular area of the mirror, F' will move through a short circular arc, approximately a straight line (the primary focal line), while the diverging fan will trace out wedge-shaped portions of space on each side of the axis, which have a common linear boundary at F'' (the secondary focal line). The reflected rays between the two focal lines fill a space similar in shape to the sphenoid of crystallography. If any difficulty is found in forming a picture of this rotation figure in the mind, it can be removed by cutting out of cardboard a diagram representing a section of the mirror, incident and reflected rays as figured above, and mounting it on a knitting needle placed in coincidence with the axis AB. By rotating the needle through a small angle, the formation of the focal lines and the sphenoidal bundles of rays can be readily seen.

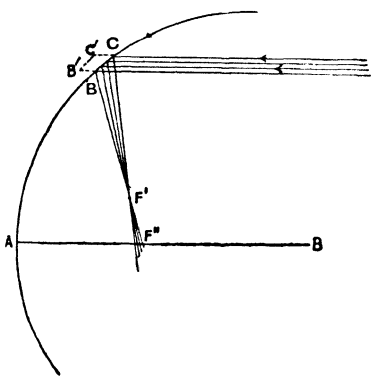


FIG. 46

If we require a mirror which will form a focal point in an oblique direction, the portion BC must be cut from a paraboloid surface the axis of which is AB. This is for parallel rays, and the subject has already been discussed under reflecting telescopes.

If we require a mirror which will form a focal point in an oblique direction for rays coming from a source at a small distance, the mirror must be a portion of an ellipsoid, say the portion at D in Fig. 49, where A and B are the conjugate foci. Such mirrors would be very difficult to make, as the curvatures are different along different meridians.

Let us next endeavor to explain the formation of focal lines by considering the form of the wave-front.

The curvature of the wave-front as it leaves the mirror under these conditions is different along different meridians. If we cut a piece out of the side of a hen's egg we shall have something of analogous form.

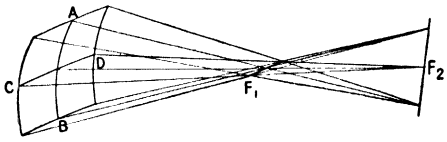


FIG. 47

Let AB be the direction of greatest curvature and CD that of least curvature (Fig. 47). To start with, suppose the curvature be equal along all lines parallel to AB, and suppose all lines parallel to CD to be straight. This will give us a cylindrical wave which will come to a linear focus at F_1, the length of the line being equal to the length of the cylindrical wave.

If we start with a square wave-front we shall find it contracting to a line as we approach F_1, and expanding beyond F_1, first as a horizontal rectangle, then a square, and finally a vertical rectangle.

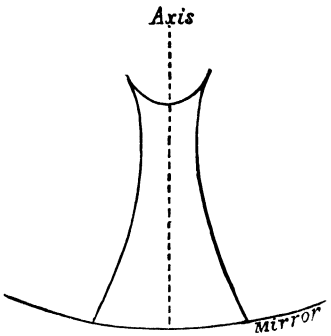


FIG. 48

Now let us impress a slight curvature parallel to CD. The result of this will be that our square will now contract in both directions, only in one less rapidly than in the other, and the line at F_1 into which it shrinks will be shorter than before, and instead of being straight will be slightly concave towards F_2. From here it can be regarded as an expanding wave in a vertical plane, and a contracting wave in a horizontal plane. It is easily seen that the line at F_1 will now open out, first into a horizontal rectangle, as before, then a square (as the two sides *closing in* become equal to the top and bottom *moving out*), then a vertical rectangle, and finally a vertical line at F_2, as the sides come together.

It is interesting to enquire as to the nature of the rectangular wave-surface between F_1 and F_2. From its nature we see that it must be concave towards F_2 in the horizontal plane, and convex in the vertical, the surface resembling a small portion cut out from the inside of a thick cylindrical ring. We can, indeed, find surfaces of this form on our geometrically constructed wave-fronts.

Consider the diagram shown in Fig. 48 (which will be recognized

as the "volcanic cone" form), remembering that the complete wave-front is formed by the rotation of this figure around the axis of the mirror.

The bowl of the crater is concave along every meridian, but it is at once apparent that any portion of the outer slope has the required saddle-shape, being concave in horizontal planes and convex in vertical planes. From this it is evident that the outer wall of the volcanic cone, before it crosses the axis of the mirror, always represents the portions of the wave-front between the primary and secondary focal lines.

That this is true is evident, when we recollect that the first focal line is formed by the intersection of rays on the caustic surface, or, regarded from the wave point of view, by the passage through their foci on the cusp of the wave, of adjacent elements of the wave-front. The second focal line lies on the axis of the mirror; consequently the wave-front between the lines is that portion of the surface which has passed through a focus on the cusp, but which has not crossed the axis.

It will be found that a small glass model of the wave-front, shown in cross section in Fig. 48, is extremely useful in making the whole matter clear. It can be made by drawing down a large thin tube, melting the end down flat, and then sucking it in a little.

Fermat's Principle. — We sometimes find it stated that a ray of light in passing from one point to another by way of either a reflecting or refracting surface, chooses a path such that the time of transit is a minimum. This principle was stated by Fermat more than two centuries ago. It is true, however, only for plane surfaces. In the case of reflection from a plane surface the incident and reflected rays make equal angles with the normal, and we know from elementary geometry that this path is the shortest that can be traced from one point to the other by way of the surface. The same is true for convex surfaces, but for concave surfaces we find that in certain cases the path is a maximum instead of a minimum.

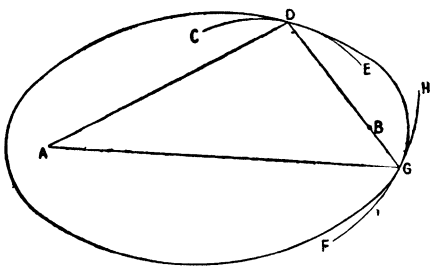


FIG. 49

That the path is sometimes a maximum can be seen by the construction shown in Fig. 49. We will consider the passage of a ray from the point A to the point B by way of the reflecting spherical surfaces CDE and FGH. Around the points A and B as foci we

construct an ellipsoid which we will suppose to be tangent to the two reflecting surfaces at D and G. This ellipse is an aplanatic surface for rays issuing from either focus, consequently the time of transit of a ray from one focus to the other by way of the elliptical surface is the same for every point on the surface. Now the ellipsoid and the two spherical surfaces have common tangent planes at D and G, consequently D and G will be the points on the spherical mirrors so oriented that they can reflect rays from A to B. It is easy to see that the path AGB is shorter than any other path between A and B by way of the sphere which is exterior to the ellipse, while in the case of the other sphere the path actually pursued (ADB) is longer than any other path which we can draw from A to the surface and from thence to B. In this case we see that the path chosen by the ray is such as to make the time of transit a maximum. The conditions for a maximum or minimum may be expressed by saying that the variation of the time of transit with the change of path, ceases at the points for which the path is either a maximum or minimum, or $\delta(AD+DB)=0$.

CHAPTER III

REFRACTION OF LIGHT

In the preceding chapter we have discussed the forms and behavior of the wave-fronts reflected back into the first medium, when light falls upon the boundary between two media of different optical density. A portion of the energy, however, always passes into the second medium, except, perhaps, in the special case of total reflection, and even in this case mathematical analysis shows us that there is a disturbance beyond the boundary, though only penetrating to a distance of a few wave-lengths. The energy crossing the boundary may either be absorbed by the second medium, or propagated according to the laws governing luminous disturbances in it.

In the present chapter we shall consider only the case of wave-propagation in an isotropic medium, or one in which the velocity of propagation is independent of direction. Later on we shall investigate the refraction of light in bodies in which the velocity is different in different directions.

We will begin by considering the refraction of a plane-wave at a plane surface.

Refraction of Plane-Wave at Plane Surface. — Suppose a plane-wave incident at an angle of 30° on a flat surface of glass, and assume the velocity in the glass to be less than the velocity in air, as we shall subsequently show it to be.

The various points on the glass surface become in succession centres of secondary disturbances as they are struck by the incident wave. These secondary wavelets spread out in both media, and it has been shown by Huygens's construction that the reflected wave is the envelope of those spreading out in the first medium. If we apply the same construction to the second medium, supposing for the sake of simplicity that the velocity of the wave-propagation in it is only one-half as great as in the first, the wavelets in the glass will have radii half as large as the corresponding wavelets in air, and the enveloping surface or refracted wave-front is turned through an angle (Fig. 50). The rays, or normals of the wave, are therefore bent an equal amount.

It was determined in 1621 by Snell that in every case of refraction the incident and refracted rays make such angles with the normal to the surface, that the ratio of their sines is constant for

any two given media. Snell's law of refraction we now know holds only for isotropic media. It can easily be deduced from Huygens's construction in the following way.

Construction. — The angle of incidence i is the angle between the incident ray and the normal to the surface. It is also the angle between the wave-front and the surface. The same is true for the angle of refraction. Let the velocity in air be v, and the velocity in glass be v' (equal to say $v/2$), and let t equal the time required for the wave in air to traverse the distance BB' (Fig. 50). Then $BB'=vt$, and the radius of the secondary disturbance around A in the glass will be AA' or $BB'/2$. We have then $BB'/AB'=\sin i$ and $AA'/AB'=\sin r$,

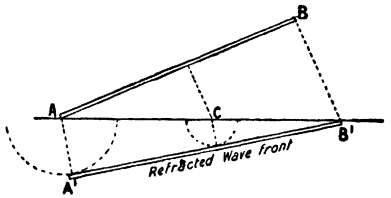


FIG. 50

or
$$\frac{\sin i}{\sin r}=\frac{BB'}{AA'}=\frac{v}{v'}=\text{constant, in this case 2.}$$

This constant is the relative refractive index between the media, and the above relation holds for every value of i, if the second medium is the one in which the disturbance travels at a lesser velocity. The refractive index is usually designated by μ, and in the above case is of course 2, which is higher than is usually the case.

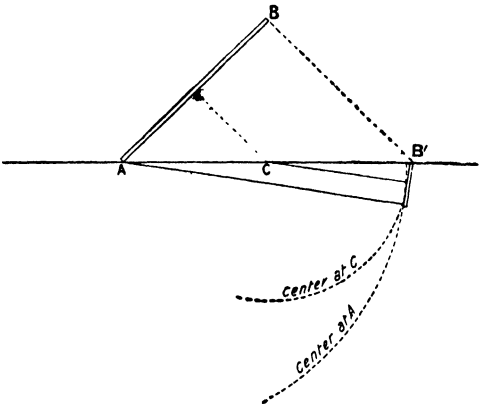


FIG. 51

Total Reflection. — We have seen that in passing from a rare to a denser medium a refracted ray always exists, no matter how great the angle of incidence. This is not true if we reverse the conditions, for now the relative refractive index will be less than one, and we shall find that, if i exceeds a certain value, sin r is greater than one. But no angle has a sine greater than one, therefore there can be no refracted ray. Let us apply Huygens's construction to the case. The secondary waves in the second medium will have

radii greater than the corresponding ones in the first medium, since their velocity of propagation is greater. By dividing the radii of the reflected wavelets by μ, we obtain the dimensions of the refracted wavelets. We shall find that, up to a certain value of i, these secondary disturbances will intersect the surface *within* the projection of the incident wave upon it; in other words, a tangent plane can be drawn from the point B' (Fig. 51). At a certain value of i, however, the secondary wavelet around A will intersect the surface at B', and the same will be true for all of the other wavelets: (since, if the short leg of one of the right triangles divided by μ gives us the hypotenuse, the same will be true of all the other similar right triangles). The tangent plane drawn from B' to these wavelets will be normal to the surface, and will touch the wavelets in a single point only (in the sectional diagram). The refracted ray therefore will travel along the surface, or more correctly there will be no refracted ray.

The value of i for which this condition exists can be found by combining $\sin i = AC/AB$ with $AC/\mu = AB$, which gives us $\sin i = \mu$, μ of course being the refractive index of the rarer medium with respect to the denser. If, as is customary, we consider μ as the refractive index of the denser with respect to the rarer the last equation becomes

$$\sin i = \frac{1}{\mu}.$$

The angle determined by the above expression is known as the *Critical Angle*. If it be exceeded, the secondary wavelets cut the surface beyond the point B' and no tangent plane can be drawn, therefore no refracted ray exists. The energy in this case is totally reflected.

If in the formula $\sin i / \sin r = \mu$ we assign various values to i, and solve for r, we shall find that every possible value of i between 0° and 90° gives a corresponding value of r if μ is greater than one, which is always the case when the ray passes from a rare into a denser medium. Take the case of rays of light entering the level surface of a pond, at all possible incidences from 0° to 90°. The zenith light passes straight down, the horizon light is refracted in a direction given by $\sin 90° / \sin r = 1.33$, or $\sin r = 1/1.33$, which gives for r a value slightly greater than 48°. In other words, no ray in the water makes an angle with the normal greater than about 48°. The light therefore which enters an eye under water consists of rays embraced by a cone of 96° angular aperture (Fig. 52) instead of 180°, as is the case when the eye in air is directed towards the zenith. If therefore, when submerged in water, the eye be directed

towards the surface, the sky appears compressed into a circle of light subtending an angle of 96°, the appearance being precisely as if the water were covered with an opaque roof with a round hole directly overhead. If, however, we are in diving armor, and look upward through the plate glass window of the helmet, the illusion of the hole vanishes, for now the horizon rays are refracted back into their original direction on passing into air once more, as is shown in Fig. 53, the 96° cone widening out to 180°.

In this connection it is of interest to ascertain how the external world appears to a fish below the surface of smooth water. The

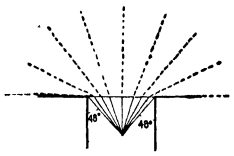


FIG. 52

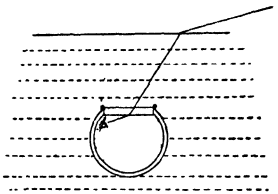


FIG. 53

objects surrounding or overhanging the pond must all appear within the circle of light previously alluded to. There must be a great deal of distortion of objects which are not very nearly overhead, but we can gain absolutely no idea of their appearance by opening the eyes under water, since the lens of the human eye is only adapted to vision in air, and when submerged is quite unable to distinguish the shape of objects. There is, however, no difficulty in photographing the circular window of light and the external world as seen through it. It was found after a little experimenting that better results were obtained with a pin-hole than with a lens, and a small camera was constructed which could be filled with water and pointed in any direction. If pointed vertically it recorded the view seen by a fish in a pond; if horizontally, the view as seen by a fish looking out through the side of an aquarium. It is obvious that the plate must be immersed in water, as otherwise refraction occurs as in the helmet of diving armor.

The fish-eye camera can be made of a wooden or metal box measuring about 12×12×5 cms. (inside measure).[1] A hole 3 cms. in diameter is bored through the centre of one of the sides, over which is cemented a piece of mirror glass with the silvered and varnished side facing the interior. The glass must be quite opaque, *i.e.* free from pin-holes in the silvered film. A very small hole should be made through the film by scratching it carefully with a needle, before the plate is cemented to the box. This small aperture passes the rays of light which form the image to the photographic

[1] R. W. Wood, "Fish-eye Views," *Phil. Mag.*, Aug. 1906.

plate which lies against the opposite side of the box. The box must be light-tight, and filled with clean water. A little consideration will show that the part played by the water in the pond is, in this case, played by the glass plate. A number of views secured

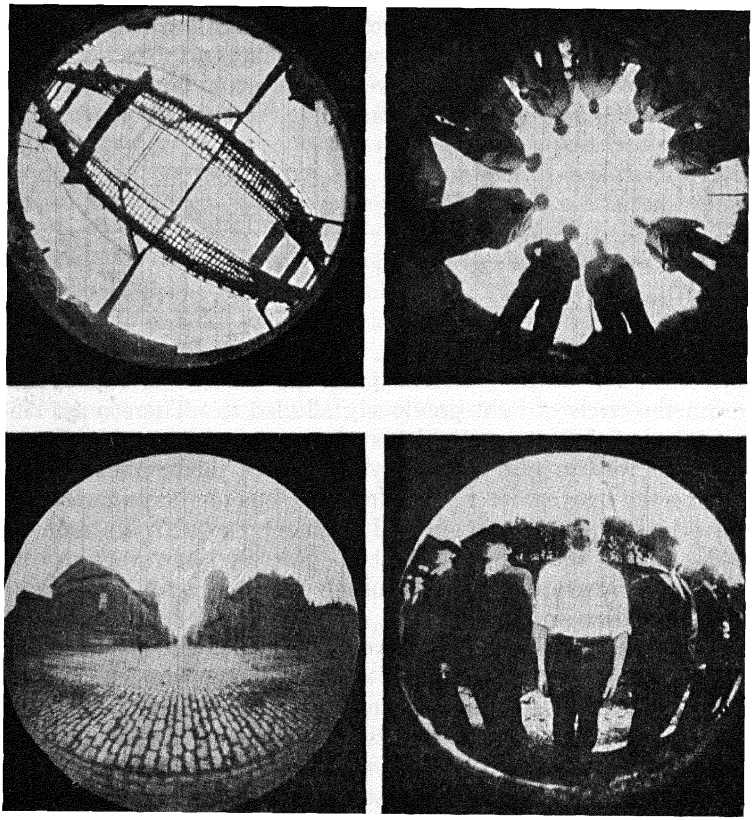


FIG. 54

with the apparatus are reproduced below, Fig. 54. The camera obviously has an aperture of 180°.

One of the views is of a railroad bridge passing overhead, the other represents the appearance of a crowd of men standing around a pond, to a fish below the surface. The two lower views were taken with the camera pointing in the horizontal direction, *i.e.* the views correspond to what a fish sees when looking out through the side of an aquarium. One of them shows a view looking both

up and down a street, the other a row of men standing in a straight line taken from a point only 50 cms. in front of the central figure. These last two show in a very effective manner that the angle of view embraces 180°.

Total Reflecting Prism. — A right-angle prism, as shown in Fig. 55 constructed of glass or quartz, gives practically total reflection for rays of all wave-lengths for which the material is transparent. Quartz prisms of this type are used for reflecting ultra-violet light without loss of energy, as in the construction of some spectrographs. If a ray enters such a prism in the direction indicated by the dotted line, it suffers two internal reflections and emerges in a direction perpendicular to the incident ray, regardless of the angle of incidence i, provided it be large enough to cause the two internal reflections. Since this holds accurately only if the two acute angles are exactly equal, and the right angle exactly 90°, such prisms are tested by viewing an object by the emergent dotted ray, and then turning the prism to and fro on a vertical axis (perpendicular to the page in Fig. 55). If no movement of the object is seen the prism is accurately figured.

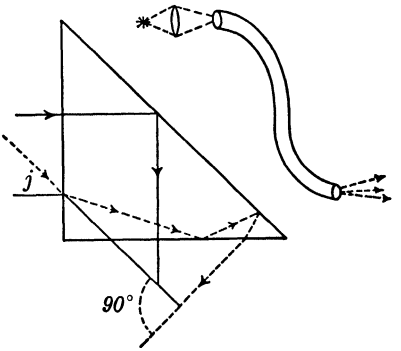


FIG. 55

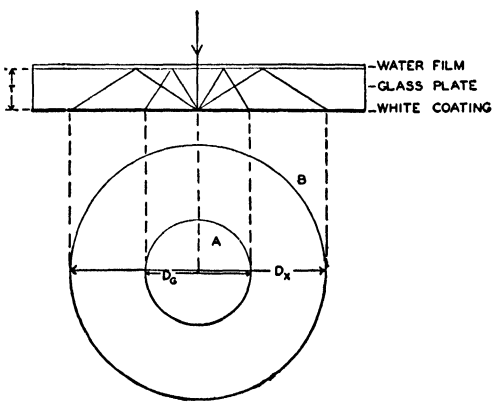


FIG. 56

Total Reflection within Rods. — Light can be conducted from one point to another without much loss of energy by total internal reflection from the walls of a rod of glass, or better of fused quartz, as shown at the right of Fig. 55 the device being sometimes of use in experimental work. It can be conducted in this way to a considerable distance along a flexible thread of fused quartz, which is more transparent than glass.

Total-Reflection Refractometer. — A very simple and inexpensive refractometer by which refractive indices of liquids can be determined to the third place of the decimal, has been described by A. H. Pfund. "It consists of a plane-parallel glass plate having the approximate dimensions $10\times10\times0.5$ cms. The lower surface of this plate is coated with white paint which is allowed to dry. From above (as shown in Fig. 56) a brilliant 'point-image' of an auto headlight lamp is projected on the glass-paint interface. This image acts as a new source of light which sends out rays in all directions. At the angles of emergence equal to and greater than the critical angle all rays are returned to the lower white surface, thus producing an illuminated area which has a black circular disc A at its center. If a film of water be spread over the upper surface, a second and larger circular ring B makes its appearance. As is evident this ring is due to rays which are totally reflected at the glass-water interface. From a knowledge of the plate thickness and ring-diameters, refractive indices may be calculated.

"Let t = plate thickness in cm.
n_g = refractive index of glass plate
n_x = refractive index of liquid covering glass plate
D_g = diameter of central disc (no liquid on glass plate)
D_x = diameter of outer ring (liquid *on* glass plate)
then it follows, simply, that:

$$n_g = \frac{\sqrt{D_g^2+16t^2}}{D_g}$$

$$n_x = n_g \cdot \frac{D_x}{\sqrt{D_x^2+16t^2}}.$$

"For ordinary purposes a good grade of plate glass of about 7 mm. thickness will do. It is advantageous to remove the white paint over half the area of the plate and to attach a millimeter scale of white celluloid with Canada balsam (see lower part of Fig. 57). Since ring-diameters may be read from the scale with a reading glass to 0.1 mm., the refractive index n_x is determined to the third decimal place. In view of the relatively low refractive index of plate glass, the liquids studied should not exceed 1.48 in refractive index.

"A wider range is secured by employing a plate of dense flint glass of refractive index 1.7.

"The apparatus is constructed as follows (Fig. 57).

"A is an auto head-light bulb well blackened, but for a small window, with lamp-black, water glass and a little glycerine. The

containing tube *B* is lined with black velvet. The 'point-image' *D* of the lamp filament is formed by means of the lens *C* which effects a reduction in the ratio of about 1:10. To reduce spherical aberration a diaphragm is introduced at *E*. The tube is held in a slotted vertical bar *F* which, in turn is attached to the base-plate *G*. Sharp focusing of the filament image on the lower, white surface of the refractometer plate *H* is brought about by changing the tilt of the tube *B*.

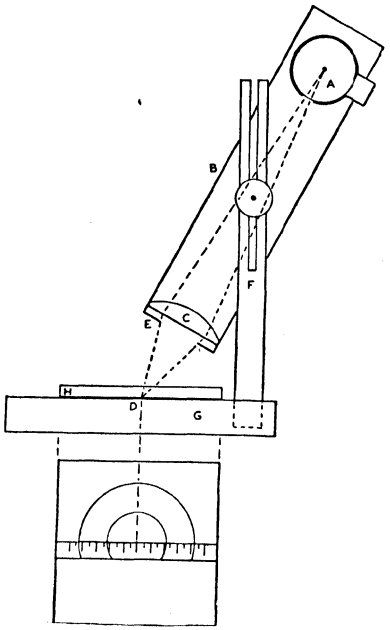


FIG. 57

"The appearance of the rings is shown in Fig. 58 where the outer ring is due to a film of water ($n = 1.333$).

"This photograph was made by substituting a photographic plate (film side up) for the white paint surface, covering it with the glass plate (with a film of glycerine between), and spreading a film of water over the surface of the plate. After an exposure of a few seconds the plate was washed and developed.

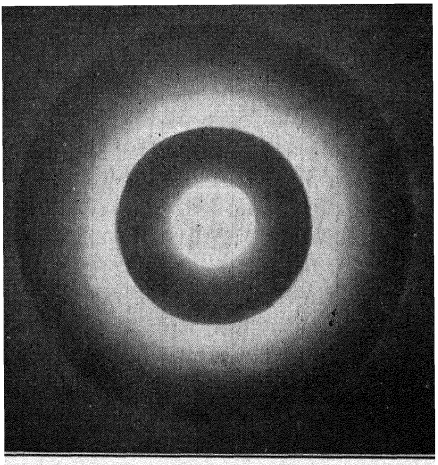


FIG. 58

"In spite of the fact that the filament-image *D* (Fig. 57) is far from being a point image, the sharpness of the rings is remarkable. The absence of chromatic effects at the ring is due to a process of achromatization brought about by a relatively constant ratio of the refractive index of glass to that of water throughout the spectrum. Repeated

trials have shown that ring-diameters may be measured to an accuracy of ± 0.03 mm. on a dividing engine supplied with long-focus microscope of unit magnification. Under these conditions refractive indices may be determined to the fourth decimal place."

Refraction of a Wave by a Plane-Parallel Plate. — The application of Huygens's construction to the passage of a plane-wave through a glass plate bounded by parallel planes, shows at once that the emergent wave-front is parallel to the incident, no matter how great the angle of incidence. The direction of the ray is therefore unchanged, though each individual ray is shifted to one side by its passage through the plate. Inasmuch as the position of an object at a great distance depends solely on the direction of the parallel rays reaching the eye, it will not be changed by the interposition of a thick plate, at any angle. We can test this by viewing a very distant object through a thick piece of plate glass and turning the plate rapidly around a vertical axis to the right and left. Objects near the plate, however, will be found to shift their apparent position considerably as the plate is turned. If the two objects and the eye be in the same straight line it may seem at first sight as if the intervention of the oblique plate would in no way affect their apparent positions, for parallel rays from the distant object are unchanged in direction by passage through the plate, and the same is true of the rays from the near object. If, however, we remember that the ray is shifted laterally, the difficulty disappears, for the lateral shift, while it does not alter the apparent position of an object at infinity, displaces an object situated at a finite distance. This will be made clear by reference to Fig. 59.

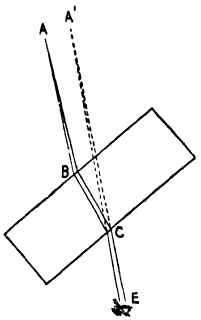


FIG. 59

Let A be a point not far from the plate. It is seen by an eye at E by means of the rays pursuing the path $ABCE$, and its apparent position is A^1. If the oblique plate be removed, the point A will be seen by the direct pencil of rays AE, and will appear in its true position A.

If we make the same construction for parallel rays coming from a distant point we shall find that the apparent position, or the *direction* from which the rays by which it is seen come, is unchanged.

The Ophthalmometer. — This principle is made use of in the ophthalmometer, an instrument devised by Helmholtz for determining the curvature of the lens of the eye, by measuring the

diameter of the image of a source of light seen reflected from the curved surface.

The instrument enables us to measure the distance between two points, or the diameter of an object, without taking into account its distance, by an optical method.

It consists of a small telescope with two plane-parallel thick glass plates in front of the objective, arranged so as to rotate about a common axis, the angle of inclination between the plates being measured by a graduated circle. If the object is at a great distance rotation of the plates produces no effect of course. If at a small distance we see it doubled as soon as the plates are inclined, and by setting the plates in such a position that the two images touch each other end to end, and reading off the inclination we can determine the length l of the object by the formula

$$l = 2a \sin \Phi \frac{\sqrt{n^2 - \sin^2 \Phi} - \cos \Phi}{\sqrt{n^2 - \sin^2 \Phi}},$$

in which a is the thickness of the plates, n their refractive index, and Φ the angle through which each plate is turned from the zero position (plates parallel).

As will be seen, the distance of the object does not come in at all. As we bring the object nearer, it appears larger, but the angular shift necessary to produce a given displacement increases at the same rate. In other words there is compensation. The instrument can be used only for measuring objects at a moderate distance.

Refraction of a Spherical Wave at a Plane Surface. — Suppose a spherical wave originating at O (Fig. 60) to be refracted at the plane surface AB. If we construct the refracted wave-front by the method of Huygens, making the points on the refracting surface the centres of secondary wavelets whose radii are found by dividing their distances (measured along rays) from the wave in its unrefracted position, by the refractive index of the medium, we shall find that the incident wave is flattened down into what at first sight appears to be a sphere of less curvature.

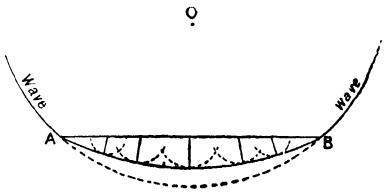


FIG. 60

It can be proved (see earlier edition) that the refracted rays are normal to an hyperboloid, and the wave-fronts are parallel curves, located by measuring off equal distances from the hyperboloid. They will *not* be themselves hyperbolae, for the parallels

to a conic are in general curves of the eighth degree. The evolute of the hyperbola is the caustic of the refracted wave, in this case virtual of course. After refraction, then, the different elements of the wave-front appear to come from points distributed along the caustic. If then we transfer our eyes from one position of the wave-front to another, the position of the radiant point in space will apparently alter. The same thing is true when the waves are refracted from a dense to a rare medium, the caustic in this case being the evolute of an ellipse, and the refracted waves parallels of an ellipse. The formation of a caustic under these conditions is shown in Fig. 61.

A small portion of the wave around the ray leaving the surface normally comes from the cusp of the caustic, which we may regard as a point, consequently this portion of the wave is approximately spherical.

The cusp of the caustic from which this portion comes is elevated above the true radiant point, consequently the refraction appears to bring the point nearer the eye. The bottom of a vessel of water consequently appears to be nearer than it really is.

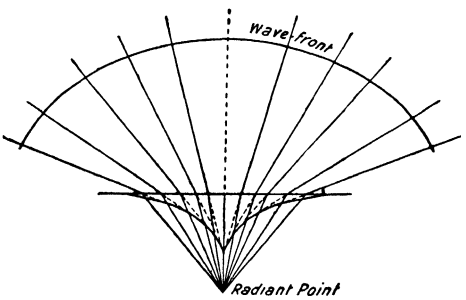


FIG. 61

The apparent elevation of the bottom of a body of still water is a matter of common observation. It is most marked when the eye is only a little above the plane of the surface and the bottom at a considerable distance is under observation. The rays which leave the surface at nearly grazing emergence come from that portion of the caustic which is very near the surface, as is apparent from Fig. 61. The bottom at a distance may thus appear elevated almost to the surface.

In the case just considered the caustic is virtual and not real, and since it is in reality non-existent, there are no moving cusps on the wave-front as in the cases considered under reflection. If we could reverse the emergent wave, and at the same time remove the water, the caustic would become real, and cusps would develop upon the wave-front. This cannot of course be done experimentally, but attention is drawn to it in order to clearly define the difference between a real and virtual caustic.

Refractive Index of a Plate Measured by the Microscope. — This apparent elevation of a point due to the decrease in the radius

of curvature of the wave-front when it emerges into the air can be used for measuring the refractive index of a glass plate, of which we know the thickness d. An object seen through the plate appears nearer by the amount $a=d(n-1)/n$.

If we focus a microscope upon an object, and then place the glass plate over it, we shall have to raise the microscope through a distance a to bring the object into focus. The refractive index is then given by $n=d/d-a$. We should use an objective of as short focus as possible, consistent with its use with the plate. It is best to provide the eye-piece with cross hairs and focus by absence of parallax, *i.e.* so that there is no relative motion between the object and the cross hair as the eye is moved from side to side.

Two modifications of the method may be cited. Make a mark on the upper and lower surface. To change from one to the other the microscope must be raised a distance a: then $n=d/a$. Make a small dot with white paint on the upper surface and illuminate it from above, on a dark background. Focus first on the object and then on its image seen reflected from the lower surface, moving the miscroscope a distance a: then $n=2d/a$. The values are correct to the third place of the decimal if the observations are made with great care.

Fermat's Law. — In the case of reflection we have seen that the path of a ray from one point to another by way of a reflecting surface is either a maximum or a minimum. The same is true in the case of refraction. If the refracting surface is plane, the time of transit is a minimum, and we have what is known as the principle of least time. If the refracting surface is curved, the time may be either a maximum or a minimum, according to whether the refracting surface lies within or without the aplanatic surface, the same as in the case of reflection. Fermat's law may be deduced from Snell's law by the maximum and minimum method of the calculus, the full treatment being given in earlier editions.

Refraction by a Prism. — In the case of refraction by a prism we have to determine the deviation of a ray or wave-front, by passage through a medium bounded by two planes which make an angle with each other; this angle is called the angle of the prism. If the refractive index of the prism be greater than that of the medium in which it is immersed, as is usually the case, the deviation of the ray is always away from the vertex, that is, towards the base of the prism. This is obvious at first sight, except perhaps in the case in which the incident ray falls on the prism in the direction shown in Fig. 62, for here the deviation at the first surface is *towards* the vertex, while that at the second is towards the base, the final direction depending on the relative magnitude of

these two deviations. The angle of refraction at the second surface is greater than that at the first, and since the deviation increases as this angle increases, the deviation towards the base at the second surface is greater than the deviation towards the apex at the first.

We will now derive an expression for the deviation. Let the angle of the prism be α (Fig. 63) and let i and r be the angles of incidence and refraction at the first surface, r' and i' at the second.

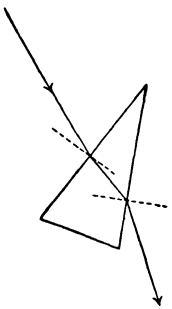


FIG. 62

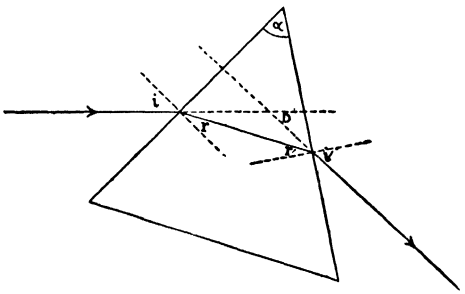


FIG. 63

The deviation is obviously D, the angle between the emergent ray and the incident ray. The deviation at the first surface is $i-r$, at the second $i'-r'$, while

$$D=(i-r)+(i'-r')=i+i'-(r+r').$$

But $r+r'=\alpha$, since α+the two base angles of the prism=2 right angles and $(r+r')$+the base angles=2 right angles.

$$\therefore\ D=i+i'-\alpha.$$

Refractive Index of a Prism. — Let us now suppose the angle of incidence to be such that the ray passes through the prism parallel to the base. In this case $i=i'$ and $r=r'$, and if we can measure D and know the angle α we can easily determine the refractive index of the prism.

We have $D=2i-\alpha$, or $i=\dfrac{\alpha+D}{2}$, also $r=\dfrac{\alpha}{2}$.

Substituting these values in

$$\mu=\frac{\sin i}{\sin r}\ \text{we have}\ \mu=\frac{\sin\tfrac{1}{2}(\alpha+D)}{\sin\tfrac{1}{2}\alpha}.$$

We must now find some method of arranging the angle of incidence so that the path of the ray through the prism will be parallel to the base, *since it is only for this condition that the above formula holds.*

This adjustment is very easily made, for the deviation of the ray can be shown to be a minimum when the passage through the prism is symmetrical.

Effect of Refraction on the Width of the Beam — It is at once apparent, by reference to the diagrams for the construction of the refracted wave-fronts, that the width of the beam is increased in a direction parallel to the plane of incidence, when the rays pass from a rare to a dense medium, or a given portion of the wave-front is spread out over a larger surface. In passing from a dense to a rare medium the reverse is the case, the wave-front being compressed into a smaller area. This change in the width of the beam diminishes or enlarges (in one direction) the apparent size of objects seen under these conditions, a matter which will now be investigated.

Magnifying Power of Prisms. — When the prism is set at minimum deviation the widths of the incident and emergent beams are the same, otherwise not. For example: in Fig. 64 when the incident beam falls normally on the first surface, and leaves the second surface at a large angle with the normal, the width of the beam has been contracted. If we view an object under these conditions, the eye being placed in the contracted beam, we shall find that it is magnified in the direction in which the beam has suffered contraction. A circular opening in a card backed by a sodium flame is a suitable object, and will be found to appear as an ellipse. If an achromatic prism is available a circular white object can be used, when the effect is very striking. If on the other hand the incident light makes a large angle with the normal, the emergent wave-front is expanded in width, and, if the eye be placed in it, the object will appear decreased in size in this dimension, a circular card appearing as if turned edgewise. Brewster suggested that by using two achromatic prisms at right angles to each other, magnification might be shown in both directions, and the action of a telescope imitated.

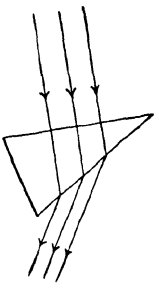


FIG. 64

Lord Rayleigh has given a very neat demonstration of magnifying power, based on Fermat's law, which is applicable to telescopes as well as to the case just cited. It proves by a wave-front method that the contraction of a beam of light, or the compression of a wave-front, causes magnification. Consider a wave-front of width AB (Fig. 65) refracted at the surface CP, and compressed thereby to width $A'B'$. By Fermat's law the time of transit over the path ACA' is equal to the time of transit over the path BPB', being a

minimum in each case. This we may express by saying that $\int \mu ds$ (the reduced path) is the same along each ray. If from any cause B is retarded relatively to A, say an amount BE, B' will be retarded an equal amount relatively to A', namely $B'E'=BE$.

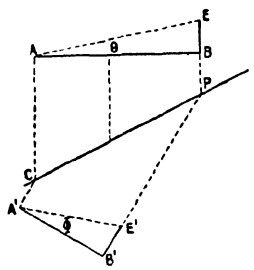


FIG. 65

(This is of course true only if AB and $A'B'$ are in the same medium; the first surface of the prism, parallel to AB, is not represented.) If this retardation be considered as represented by a rotation of the wave-front AB through angle θ it will be measured by $(AB)\theta$. The wave-fronts AB in the two positions can be thought of as two separate fronts coming from two distant stars subtending an angle θ at the point of observation. The retardation of B' must be of the same amount, consequently the rotation of the wave-front $A'B'$ will be much greater than θ, being measured by $\Phi(A'B')$. Since the retardations are equal we can write

$$\theta(AB)=\Phi(A'B') \text{ or } \frac{\theta}{\Phi}=\frac{A'B'}{AB}.$$

Now Φ is the angle formed by the rotation of $A'B'$ the compressed wave-front, consequently we may regard it in its two positions as two fronts coming from stars which subtend an angle Φ, as much greater than θ as AB is greater than $A'B'$.

The same reasoning can be applied to telescopes, the compression here being symmetrical, a plane-wave of large area emerging from the eye-piece as a plane-wave of small area, the magnifying power being equal to the *ratio of the widths of the stream of light before and after entering the telescope.*

The reduction in size, or minification, resulting from expansion of the wave-front, can be shown in a striking manner with a prism spectroscope and a sodium flame, for the D lines can be resolved with a slit width of a millimetre or two if the prism is rotated to a position such as to make the rays incident at an angle of nearly 90°. It has been suggested that a great gain in light could be obtained in this way, but it is obvious that the loss due to the utilization of but a small part of the light from the collimator more than offsets the gain.

In recent years concave glass diffraction gratings have come into use for the study of the remote ultra-violet spectra and the best results have been obtained with nearly grazing incidence, as the reflecting power of the glass is greater under these conditions. Extremely narrow lines are obtained with a fairly wide slit.

The Fraunhofer lines can be seen in the sun's spectrum by mounting a plane grating (speculum) in sunlight at grazing incidence, minification compressing the solar disk to a narrow line.

Refraction by a Lens. — In the Chapter on Reflection it has been shown that a parabolic mirror transforms a plane-wave into a contracting spherical wave, while an ellipsoidal mirror exerts the same action on spherical waves originating at one of the foci. It is possible to construct refracting surfaces having the same property. We will begin by computing the refracting surface, which shall be aplanatic for spherical waves.

Let O (Fig. 66) be the luminous point, and O' the conjugate focus where the converging waves are to shrink to a point. By Fermat's principle the reduced paths along the different rays will be equal, and the disturbances will all reach O' in the same phase, resulting in intense illumination. Suppose O to lie in a medium of ref. index 1, practically in air, while O' is in a medium of ref. index 2 between which we require an aplanatic surface of separation. Let r and r' be the distances of any point on the surface from O and O', then $r+\mu r'=$Constant, the equation of a Cartesian oval. Choose a point on the line joining O and O' such that $r=5$ and $r'=7$. The constant for this particular case will be 19. Now describe around O a circle of radius 6 and around O' a circle of radius of $(19-6)/2$. The intersection of these circles will give two more points on the aplanatic surface, which can be gradually built up by giving to r constantly increasing values.

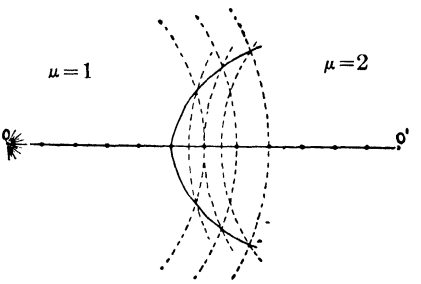


FIG. 66

The general form of the equation of a Cartesian oval is

$$\mu r+\mu' r'=\text{Constant, from which we get } \mu\frac{dr}{ds}+\mu'\frac{dr'}{ds}=0.$$

In the case just considered the conjugate foci lie in different media. If they are to be in the same medium we require an intervening medium capable of effecting the required change in the form of the wave-front. We thus come to the Aplanatic Lens.

Spherical light-waves, originating at a point in air, are to be transformed by a lens into converging spheres which come to a focus at another point also situated in air. Suppose the lens to be midway between the two points and the curvature of its two surfaces the

same. The spherical wave will be changed into a plane-wave by the first surface, consequently the simplest way to construct the form of the lens will be to trace the surface aplanatic for a plane-wave by the method given above, modifying it, however, in such a way as to make the sum of any path measured along a ray from the plane-wave to the aplanatic surface, and the reduced path from this point to the focus, a constant. These aplanatic surfaces are, however, of very little practical importance, for they can be reproduced only when the departure from a spherical surface is very slight. In the process of lens making the surfaces which are being ground

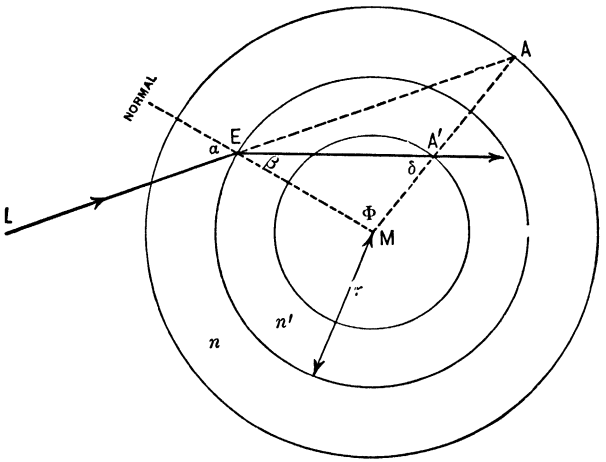


FIG. 67

together assume of their own accord a spherical form, since two surfaces, to fit together in all positions, must be of constant curvature. Lenses with spherical surfaces do not bring rays accurately to a point, or in other words do not give converging waves which are truly spherical. This results in what is known as spherical aberration, which has been treated sufficiently for the purposes of this book under *reflection*. A spherical surface may be made approximately aplanatic by local grinding, if the amount of material to be removed be not too great. This process is known as correcting the lens for spherical aberration, or figuring, and is largely a cut and try operation.

Refraction by Sphere. — There is one special case in which the spherical lens is aplanatic, which is made use of in the construction of the microscope. Weierstrasse gives this simple method of constructing the refracted rays when incident upon a sphere.

Suppose a sphere of refractive index n' and radius r immersed in a medium of refractive index n. If the sphere is in air, $n=1$.

Describe a circle of radius r representing the sphere, and around its centre two other circles of radii $(n'/n)r$, and $(n/n')r$, respectively, as in Fig. 67. Draw the ray LE incident upon the sphere at E, at incidence angle α, and continue it until it cuts the outer circle at A: then join this point with the centre, the line cutting the inner circle at A'. The line EA' is then the refracted ray. If now we have a convergent system of rays falling upon the sphere, which would unite at A if the sphere were absent, it is clear from the construction that the sphere will bring them all together in a point focus at A'. Conversely if rays emanate from A' within the sphere, they will, after refraction out through the lower half of the sphere, traverse paths which, if produced backwards, meet at A. A and A' are called the aplanatic points of the sphere, and the sphere is an aplanatic surface for these two points, one of which is real, the other virtual.

The proof of the above construction is as follows.

Triangles EMA and EMA' are similar having $\angle\Phi$ in common, and its enclosing sides in equal ratio

$$\frac{EM}{MA'}=\frac{AM}{EM}=\frac{n'}{n}.$$

$$\therefore \angle EA'M=\angle AEM,$$

and since $AEM=\alpha,\ \delta=\alpha.$

In triangle EMA'

$$\frac{\sin\delta}{\sin\beta}=\frac{EN}{A'M}=\frac{n'}{n}.$$

$$\therefore \frac{\sin\alpha}{\sin\beta}=\frac{n'}{n}.$$

Microscope Objective. — The existence of the aplanatic points just proved was utilized by Amici in the construction of microscope objectives of wide aperture. A section of such an objective is shown in Fig. 68.

A hemispherical lens I receives the wide cone of rays from a point at L. After refraction by the plane surface, they pursue directions as if coming originally from L^1. If L^1 is the aplanatic point, refraction by the spherical surface will render them still less divergent without introducing any spherical aberration. Since refraction at a plane surface introduces aberration, better conditions obtain if a drop of some oil having the same refractive index as the glass is introduced between the plane surface and the object, which must now be at L^1, the aplanatic point. This virtually imbeds the object within the medium of the sphere, and the wave leaving it has its centre at L^2. The next lens II is a meniscus, its first or concave surface having the same radius of curvature as

the spherical wave leaving lens *I*, *i.e.* its centre of curvature is at L^2. The wave therefore enters lens *II* without change of form. The second surface of the meniscus is spherical and of such a curvature that L^2 is an aplanatic point for it. The wave on leaving the meniscus is still spherical, the centre being now pushed back to L^3. The rays are rendered still less divergent and finally convergent by the two achromatic lenses *III* and *IV*.

The Amici principle has a disadvantage which makes it impossible to use it more than twice. It introduces chromatic aberra-

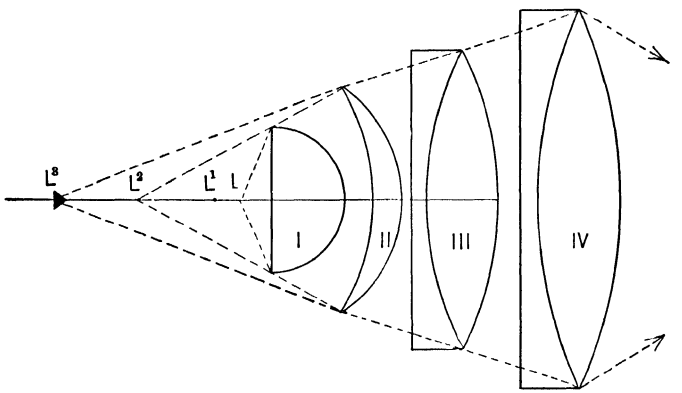


FIG. 68

tion, which can be compensated by the overcorrected (for color) achromatic lenses *III* and *IV* provided the divergence is not too small. If it is too small, as it would be if the principle was made use of again, it would be impossible to compensate it, and at the same time render the rays convergent.

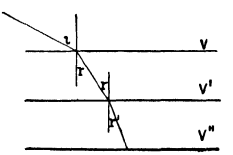


FIG. 69

Refraction of Light in Non-Homogeneous Media. — The consideration of the laws of refraction in media in which the refractive index varies continuously from point to point leads us to a most interesting class of phenomena, the most common examples of which are the illusions known as Mirages.

As an introduction to the subject it will be well to investigate the refraction of a ray of light by a number of media of different refractive indices arranged in horizontal strata of equal thickness. Let the velocities of light in the different strata be represented by v, v', and v'', and let the angle of incidence on the first layer be i (Fig. 69) — we require the deviation of the ray by the two boundaries.

We have $\frac{\sin i}{\sin r}=\frac{v}{v'}$ for first boundary. $\therefore \sin r = \sin i\frac{v'}{v}$.

The incidence angle at the second boundary is obviously r, therefore we have

$$\frac{\sin r}{\sin r'}=\frac{v'}{v''} \text{ or } \frac{\sin i\frac{v'}{v}}{\sin r'}=\frac{v'}{v''}, \text{ which gives } \frac{\sin i}{\sin r'}=\frac{v}{v''},$$

showing that the direction of the ray in the third medium is the same as if the intervening medium were not present. Now suppose the number of layers to be increased indefinitely, and the thickness of each to be reduced indefinitely. This gives us a medium of continuously varying refractive index, and we see that the direction of the ray at any point is the same as if the upper layers were removed, and the ray entered the flat surface of a medium of refractive index equal to that which the non-homogeneous medium has at the point in question. Suppose a ray to be travelling in a horizontal direction in a medium of this nature. As the ray is moving in a direction in which the refractive index does not change, it may seem at first sight as if there would be no change of direction. The discussion of the case by ray methods would lead to this conclusion, a result which plainly shows the danger of handling optical problems in this way. No matter how limited the width of the ray, the wave-front, — the motion of which constitutes the rays, — must have a finite size, and the upper and lower edges of the front are moving in regions of different optical density. The upper edge will consequently move faster than the lower, and the front will gradually wheel around, which means that the direction of propagation is constantly changing.

Astronomical Refraction. — The optical density or refractive index of the earth's atmosphere decreases as we ascend from the surface, consequently the rays of light, which reach our eyes from the stars, move in curved paths, except when the star is in the zenith. Since the direction in which the star appears to be is the direction from which the ray comes when it enters the eye, the true position of the star can be determined only by taking into account the refraction of the atmosphere. The effect of refraction is to make the star appear higher up above the horizon, or nearer the zenith than it really is. For stars at the horizon the elevation amounts to 36 minutes of arc.

Now we have seen that the final direction of the ray is independent of the layers intervening between the medium in which the observation is made and the region from which the light comes; it

is therefore apparent that the change in direction can be determined by determining the refractive index of the air at the point where the instrument is situated, which can be done by observing its temperature, pressure, etc.

The curvature of light rays in the atmosphere also influences the apparent positions of objects on the earth's surface, the usual effect being an elevation of an object above its true position, a circumstance which must be taken into account in all geodetic observations. As a result of this refraction it is possible to see the sun and the eclipsed moon above the horizon at the same time.

The displacement due to atmospheric refraction increases very rapidly as we near the horizon as can be seen from the flattened solar disk when its lower edge is on the horizon, and careful observation shows that the lower portion is more flattened than the upper. The upper edge of the disk remains visible two and a half minutes after it has actually passed below the true horizon, and rises ahead of time by the same amount. The day is thus five minutes longer than it would be in the absence of an atmosphere. Horizontal strata of air at different temperatures sometimes distort the setting sun's circular outline into curious shapes.

The " Green Ray." — The green ray, or more properly speaking, the green flash is an interesting accompaniment of the phenomenon just treated. Owing to dispersion the blue and green rays are more powerfully refracted than the red and yellow. The upper edge of the disk, just as it disappears, shows a marked color change, the orange being replaced by a greenish-yellow or even a blue-green color. In the North Atlantic the phenomenon is rather rare, though it can occasionally be seen. It has frequently been contended that the phenomenon is an illusion due to contrast, the green spot being an "after-image." This however is sheer nonsense, as any trained observer can testify. The fact that it is sometimes seen and sometimes not is very probably due to a temperature difference between the air and water as was pointed out by the author some years ago in a letter to *Nature*. With the water colder than the air, the normal gradient of refractive index of the air would be increased by the chilling of the layer near the water. With the water warmer, the reverse would be the case, the condition being as in the case of mirage. This would be unfavorable for the occurrence of the green flash.

While the radius of curvature of a ray of light travelling parallel to the earth's surface is much greater (about 7 times) than the radius of the earth, it is possible to conceive of an atmosphere with a density gradient sufficient to lessen the radius of the ray to that of the planet. If such a condition prevailed a ray would travel

completely around the planet, if the atmosphere were perfectly transparent. In some cases we may even have an atmosphere with a density gradient sufficient to give us an even smaller radius of curvature.

Schmidt's Theory of the Sun. — Schmidt has made the suggestion that the sun may be such a body, and that the disk as we see it may be an optical illusion. He considers the sun to be a mass of gas the density of which increases from the surface towards the centre. At a sufficient depth the radiation of the gas will be

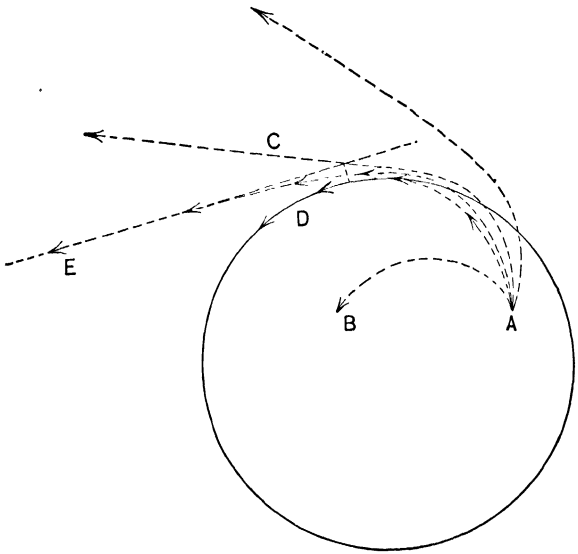


FIG. 70

"white" light, *i.e.* it will give a continuous spectrum. What we shall see, however, will be a white-hot ball with a sharply defined rim. The radius of the ball will be the radius of the sphere taken within the gas mass, upon the surface of which light rays will have a similar radius of curvature. Schmidt calls this the critical sphere.

In Fig. 70 consider the circle to be the critical sphere, and assume light radiated in all directions from a point A deep down within the gas mass. The ray B, with a small radius, will turn back into the mass; the ray D, leaving A at a smaller angle with the normal will travel around the critical sphere; while the ray E, leaving A at a slightly less angle, will pass off into space. Other rays, such as C, will also pass off into space, but will not reach us. If the gas without the critical sphere does not emit light the ray E

will come apparently from the edge of the critical sphere, notwithstanding the fact that it originated much deeper down in the mass, where the gas is radiating light as a result of its high temperature and density. On this theory what we call the diameter of the sun is merely the diameter of the critical sphere, plus a slight increase due to the refraction of the gas outside of it. An atmosphere can in the same way cause an increase in the apparent diameter of the body which it surrounds. We can show this very nicely in the following way. Make a small rectangular glass tank by cementing five squares of glass together with sealing wax. Fill it with melted gelatine and support an empty test-tube in the fluid with a clamp-stand. The bottom of the test-tube should be within half a centimetre of the bottom. After the jelly has solidified, pour hot water into the test-tube, and immediately withdraw it. It will leave a cylindrical hole in the jelly, with a hemispherical bottom. Now pour a mixture of glycerine and powdered chalk into the cavity until it is half full. Fill the remainder with water to which a few drops of milk have been added. The glycerine will gradually diffuse into the gelatine, increasing its refractive index. The condition at the end of a few minutes will be not unlike that of a white body surrounded by a dense atmosphere, for the refractive

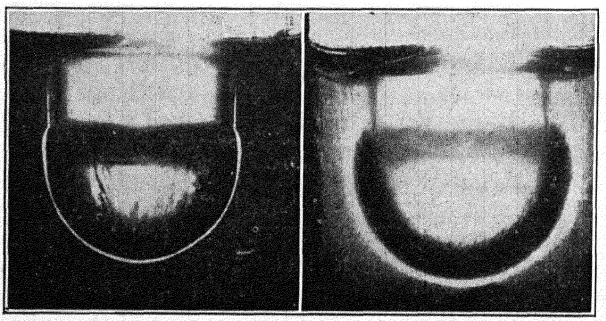


FIG. 71

index will be high at the boundary between the jelly and glycerine, gradually decreasing as we pass out into the jelly.

The magnification resulting can be seen by looking through the side of the trough, the lower portion of the cavity appearing swollen out like a mushroom. If we perform the experiment with pure glycerine and clean water the same thing happens. By placing an arc light behind the tank and throwing an image of the cavity upon a piece of ground glass with a camera objective, placed at the centre of the shadow of the tank, we can see the bright ring of light which appears to surround the bottom of the cavity. This is

analogous to the ring of light which would be seen surrounding the earth by an observer on the moon during a lunar eclipse, or rather a solar eclipse. As the glycerine penetrates into the jelly this ring of light eventually separates from the line of the cavity. Photographs of the gelatine cavity at two different stages of the diffusion are reproduced in Fig. 71.

Refraction by Planetary Atmospheres. — From what has been said it is clear that the diameter of a planet will appear to be greater than it really is if it is surrounded by an atmosphere, and the diameter will appear greatest for ultra-violet light, which is the most strongly refracted. As was pointed out by the author in 1911,[1] in describing work on lunar photography by ultra-violet light it seemed highly probable that photography of the brighter planets by infra-red light might be expected to yield valuable results if the planets are surrounded by light scattering atmospheres, since the surface of the earth as seen from a neighboring planet, would be seen through a haze equal in brilliancy to that of the blue sky, that is it would present much the same appearance as that of the moon when seen at noonday.

The author's experiments on the photography of landscapes and of the planets[2] through infra-red and ultra-violet filters showed that the obscuration of distant objects by haze was completely eliminated by an infra-red filter, and that the darkening of the planet's disk as we approach the limb is less than in photographs made with violet light, which is what would be expected on the theory that the darkening results from atmospheric absorption or scattering. It was urged that Mars be photographed by infra-red, visible, and ultra-violet light at its next opposition, but this was not done until thirteen years later, when Wright secured some very fine pictures of Mars, showing that the diameter of the planet in the ultra-violet pictures was greater than in those made by infra-red light. This was the first real proof that Mars was surrounded by an atmosphere.

Mirage. — The normal variation of the refractive index in the atmosphere is often disturbed by temperature variations, as when the air near the surface is warmed by the heated ground, or when a layer of cold air flows over a layer of warm air, as may occasionally happen. These abnormal conditions in the atmosphere give rise to the phenomenon of mirage, the commonest type being that seen on the desert where the air is heated by the hot sand. In this case the refractive index is abnormally low along the ground, rises to a

[1] Royal Institution Lecture, May 11, 1911.

[2] *Astrophysical Journal*, July, 1912, and xliii (1916). Contrib. Mt. Wilson Obs. 113, 1916.

maximum as we ascend, and then decreases more slowly according to the usual law. Rays of light near the surface are therefore concave *upwards*, while those travelling at greater elevations are concave *downwards*. The result of this is, that rays which would ordinarily strike the ground are turned upwards and reach the eye, appearing as if reflected by the ground, while other rays starting originally from the same point may reach the eye by the usual path. The point thus appears double. The sky at the horizon may thus appear as if mirrored in the sand, and since the only reflecting body in nature capable of acting in this manner with which we are familiar is a smooth sheet of water, the natural inference is that a lake exists between us and the horizon. Where the sky is broken by mountains, we see their inverted images mirrored. The paths of the rays in mirage of this type are shown in Fig. 72.

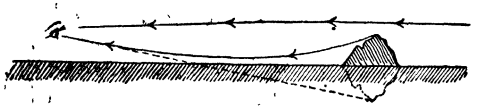


FIG. 72

It can be very beautifully reproduced by a method which was described by the author in the *Philosophical Magazine* in 1899, and which has since been somewhat improved. Three slabs of flat sheet steel each a metre long, 20 cms. wide, and 3 or 4 mms. thick, are mounted on iron tripods and carefully brought into the same plane, so that the upper surface is continuous and flat, which can be ascertained by "sighting" it from one end (upper figure on Plate 3). The surface is sprinkled with sand, to prevent reflection, which may occur at grazing incidence. A sheet of ground glass with an arc light behind it represents the sky, or a mirror mounted so as to reflect the sky when viewed from the opposite end of the desert. The artificial sky must come down to the level of the sanded surface, and in front of it a chain of mountains cut out of pasteboard is mounted, with peaks varying from 1 to 2 cms. in height, and valleys which come quite down to the sand. The desert is heated by a long gas burner made by drilling numerous small holes in a long piece of gas pipe. The gas should be introduced at each end of the long tube, and the flames should be about 5 cms. in height. If we look along the sand, holding the eye an inch or two above the plane of the surface, we shall see, as the desert warms up, what appears to be a brilliant pool of water on the sand, in which the inverted images of the mountains and sky appear reflected. Photographs of this artificial mirage are shown on the plate, Figs. 80.

Non-Homogeneous Cylinders as Pseudo-Lenses. — If there were a similar variation in the refractive index in horizontal di-

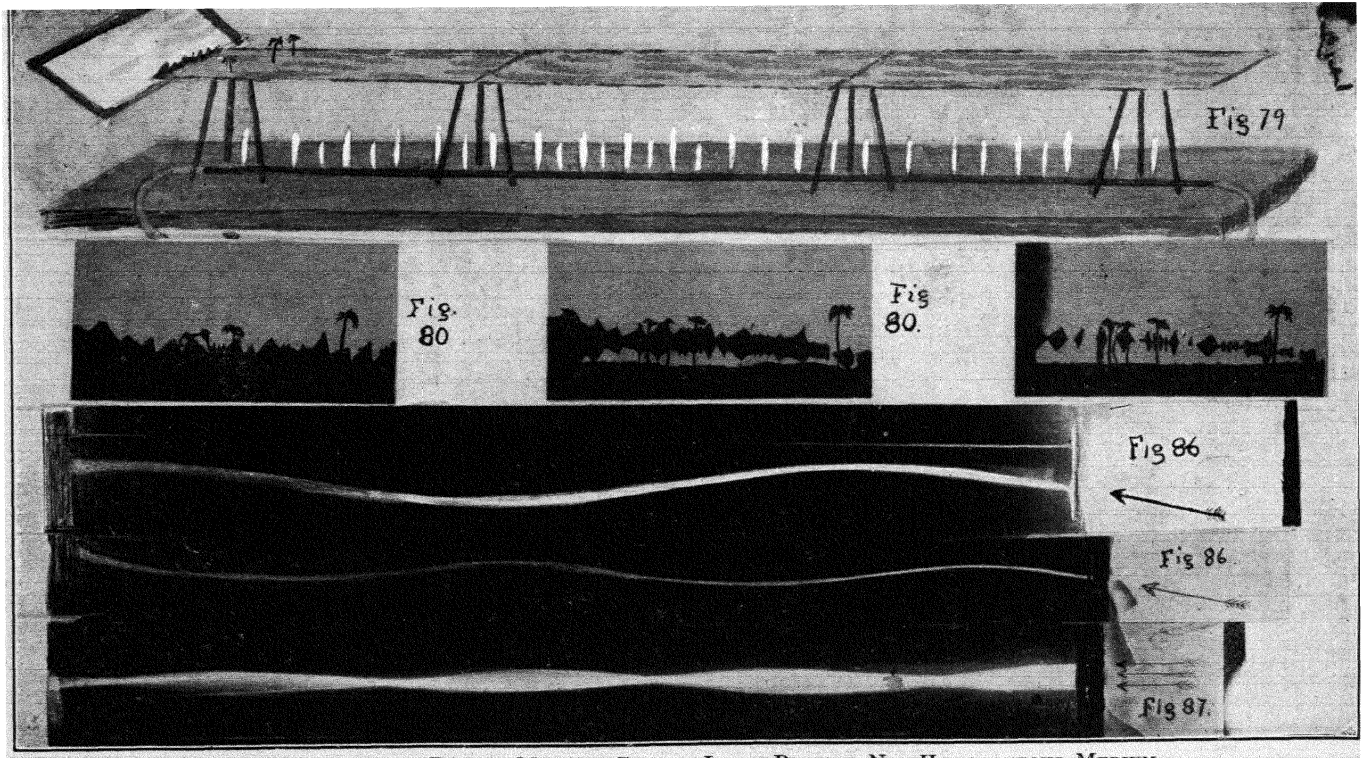


PLATE 3. ARTIFICIAL DESERT MIRAGE: CURVED LIGHT RAYS IN NON-HOMOGENEOUS MEDIUM.

rections, magnification in all directions would occur. In a medium capable of acting in this way the equi-indical surfaces, or layers, of equal refractive index will be coaxial cylinders, the highest refractive index being along the axis. Exner has shown that the eyes of some insects are arranged in this way, the convergence of the rays to a focus resulting from the action of a non-homogeneous medium.

Cylinders of gelatine soaked in water were found by Exner and Matthiessen to behave in the same way.

It is possible to prepare cylinders which have the maximum refractive index on the surface or along the axis, and act accordingly as concave or convex lenses.

The original method has been improved by the author by the use of glycerine. These pseudo-lenses are not at all difficult to prepare and are extremely interesting. A handful of photographic gelatine is soaked in clean water until thoroughly softened. The excess of water is poured off and the mass is then heated until quite fluid, and filtered through a funnel with a small piece of absorbent cotton placed at the bottom of the cone. If the gelatine refuses to run through, add a little more boiling water. Pour a small quantity into a test-tube, and let it stand until solid. Evaporate the remainder over a small flame, stirring constantly until it is of the consistency of syrup. This means boiling it down to one-third or less of its original volume. Now add an equal volume of glycerine, and pour the mixture into a second test-tube. After the jellies have set, crack the bottom of the tubes by a sharp blow, warm them by the momentary application of a Bunsen flame and push out the cylinders.

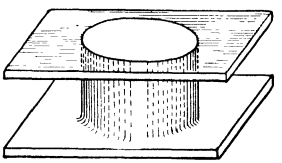


FIG. 73

Cut the cylinders into disks of different thicknesses, with a warm pen-knife. The best thickness is about two-thirds of the diameter. Mount the disks between small squares of thin plate-glass (window glass will do), warming the plates slightly, to ensure getting the jelly into optical contact (Fig. 73). It may be found necessary to prop the upper plate in position until the surface in contact with the glass has "set." The cylinders which are made of gelatine and water are now to be immersed in glycerine, the glycerine jelly cylinders in cold water. The glycerine should be stirred occasionally, as the layers in contact with the jelly take up the displaced water. The action will be found to be well under way in a quarter of an hour, the glycerine gradually diffusing into the jelly, driving out the water, and the water gradually replacing the glycerine. A jelly containing glycerine has a higher refractive index than one

containing water, consequently the cylinders soaked in glycerine act as concave, while those soaked in water act as convex lenses.

The focal length will be found to be only 8 or 10 cms., and very sharp images of the filament of an incandescent lamp or a gas flame can be obtained with them.

Schott has prepared similar cylinders of glass, by pouring the molten glass into iron tubes. The sudden chilling of the outer layer produced tension in the glass cylinder, and a corresponding variation in the refractive index, plane-parallel plates cut from the cylinder acting as concave lenses.

Curved Light Rays. — A ray of light entering a medium of this description will be bent towards, and cross the line of maximum optical density, where it changes its curvature and is again bent towards the line, which it may thus cross again and again, traversing a path which is approximately a sine curve. If we are dealing with a diverging pencil of rays, the rays will alternately converge to and diverge from a focus, passing in this way through a number of successive foci. These effects can be well shown by the following device, which was described by the author in the *Philosophical Magazine* for April, 1899.

A glass trough 50 cms. long by 10 cms. high and 2 cms. wide, with plate glass ends, is filled to the depth of 3 cms. with a strong solution of alum. On this is floated a layer of water containing 10% of alcohol, which is very much lighter than the alum solution, though having about the same refractive index. A mixture of glycerine and 85% alcohol has a much higher refractive index, but a specific gravity intermediate between these two liquids, consequently it is possible by means of a glass siphon, drawn down to a small aperture which is bent in a horizontal direction, to introduce a layer of it between the alum solution and the supernatant water. The necessary precautions and fuller directions will be found in the original paper. The three solutions were previously acidified with sulphuric acid, and rendered fluorescent with sulphate of quinine in order that the paths of the rays could be followed. By cautious stirring the diffusion of the layers into each other can be assisted, and we shall have as a result a medium in which the refractive index increases from the surface towards the median plane, and then decreases from this plane towards the bottom, the condition being similar to the atmospheric condition producing the Fata Morgana, the apparent elevation of objects on the horizon into pinnacles and spires.

If a very narrow beam of light from an arc-lamp, made parallel by means of a condensing lens, be thrown obliquely into one end of the trough, it will be seen to traverse the liquid in the form of a

most beautiful blue wave, the curvature of which varies with the angle at which the ray enters. A ray of light travelling in a sine curve is shown in Fig. 86 of Plate 3, which was photographed directly from the trough.

The alternate convergence and divergence of rays, and the successive foci can be shown by allowing a parallel bundle of rays to enter one end of the trough in a horizontal direction. A photograph of this phenomenon is also shown in Fig. 87 of Plate 3.

Scintillation. — In addition to the more or less regular gradations in the refractive index of the atmosphere there exist striae, or small regions of sudden change due to the mixing of hot and cold currents, somewhat similar to the conditions existing in a mixture of glycerine and water. When a wave-front of light passes through a region where the refractive index is low it gains, and while travelling in a region of high refractive index it loses. The result of this is, that the striae deform the plane-waves of light coming from the stars into corrugated waves, portions of which are convex in the direction of propagation, while other portions are concave. The concave portions naturally converge, while the convex portions diverge; the result being that the energy concentrates itself in certain areas at the expense of the adjacent areas, as shown in Fig. 74.

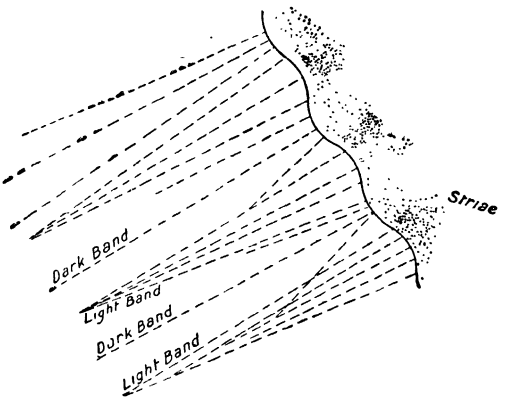


FIG. 74

This uneven distribution of light produces the familiar phenomenon of scintillation or twinkling of the stars. If the intensity of the light from a star were sufficient, we should find that instead of illuminating a white surface uniformly, as does the sun or moon, it would illuminate it unevenly, dark and light patches alternating over the surface. This uneven illumination is actually observed during the few moments immediately preceding the total stage of a solar eclipse, the patches of light and shadow being arranged in more or less parallel bands. They move along over the ground with a velocity depending on the velocity of the upper currents of the

atmosphere, usually from ten to twenty feet per second. The same phenomenon occurs also in the case of star light, except that the light is too feeble to produce shadow-bands which are visible. Their presence, however, can be inferred from the well-known fact that the brilliancy of the star observed by the eye appears to suffer rapid periodic changes, the star appearing bright or feeble according to whether the eye is in a light or dark area of the moving system of shadows. The width of the bands is frequently not over 3 or 4 cms. This means that it may easily happen that one eye is in a dark, while the other is in a bright area at the same moment. If we look at a star with the eyes slightly converged, which we can easily do by focussing them on some object at a distance of five or six feet, and in a line with the star, the star will appear doubled and the two images will fluctuate in intensity, but the fluctuations will not be "in step," one eye seeing the star dark at the moment when the other eye sees it bright.

If a star is viewed through a telescope of large aperture, the resultant illumination at the focus is the integral of the bright and dark bands covering the object-glass at the moment, and this average illumination is practically constant, therefore scintillation is no longer observed. If the aperture of the instrument be contracted by a diaphragm of such size that only the light of a single bright or dark band can enter the instrument, the twinkling reappears. It is possible in this way to actually measure the radius of curvature of the corrugations of the wave-front in the case of star light. Suppose that at a given instant the wave entering the small aperture of the telescope is concave, it will come to a focus at a point slightly nearer the object-glass than the focus of the telescope for objects at infinity. At another instant when the aperture is in a dark band where the wave is convex, the focal point for this wave will be behind the principal focus. As the dark and light bands sweep across the aperture the image of the star will alternately appear sharp and blurred. If the eye-piece is at the focus for the concave wave it will be inside the focus for the convex wave. By pushing the eye-piece in up to a point where it is possible to occasionally catch a glimpse of a sharp image of the star, and then drawing it out to a point outside the focus, for which the same conditions prevail, it will be possible to determine the minimum radius of curvature of the convex and concave portions of the wave-front. Measurements made in this way show that the average radius of curvature is about 6000 metres, although it may sometimes fall as low as 1800 metres, or rise as high as 20,000 metres. Obviously the conditions most favorable for work with astronomical instruments are to be found when the radius of curvature of

the corrugations is very large. This means that the waves are approximately plane.

One other point is worthy of mention in this connection, namely, the difference between planets and stars in the matter of twinkling. In the case of planets the light comes from a luminous disk of an appreciable size, every point of which produces a system of shadow-bands of its own. It is true that the inclination to each other of the rays coming from the different portions of the planet to the eye is very slight, but when we consider that they have traversed a distance of, say 6000 metres, in coming to a focus, that is, in forming a bright band, it is easy to see that the light from one side of the planet may easily produce a system of shadow-bands exactly out of step with those produced by the light of the other side of the disk. The superposition of a large number of shadow systems results in practically uniform illumination and absence of scintillation.

The Method of Striae. — A very ingenious and beautiful method was originated by Töpler [1] (which he named the "Schlierenmethode") for making visible in a transparent medium those regions in which the refractive index differed but slightly from that

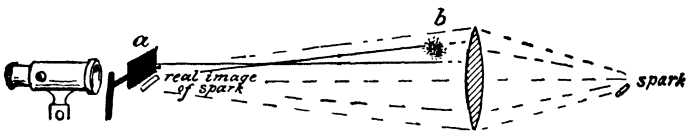


FIG. 75

of the surrounding regions. By employing as a source of light the instantaneous flash of an electric spark he was able actually to see the spherical sound-waves sent off from another spark which had occurred a moment before. Mach has used the method extensively for studying by photography the air waves given off by sparks, and accompanying rifle bullets in their flight, and an extensive series of photographs were made by the author [2] of sound-waves undergoing reflection and refraction, to illustrate some of the fundamental principles of optics. The apparatus for showing these waves can be set up in a few minutes, with very little trouble, and as the experiment is a very beautiful and instructive one it will be described in detail.

The general arrangement of the "Schlieren" apparatus is shown in Fig. 75. A good-sized achromatic lens of the finest quality obtainable, and of rather long focus, is the most important part of

[1] *Wied. Ann.*, cxxxi, 33.
[2] *Phil. Mag.*, Aug., 1899, July, 1900, May, 1901.

the device. The object-glass of a small telescope three or four inches in diameter is about right.

This lens is mounted in front of a suitable source of light (in the present case, an electric spark), which should be at such a distance that its image on the other side of the lens is at a distance of about fifteen feet.

The image of the spark, which we will suppose to be straight, horizontal, and very narrow, is about two-thirds covered with a horizontal diaphragm (*a*), and immediately behind this is placed the viewing-telescope. On looking into the telescope we see the field of the lens uniformly illuminated by the light that passes under the diaphragm, since every part of the image of the spark receives light from the whole lens. If the diaphragm be lowered the field will darken, if it be raised the illumination will be increased. In general it is best to have the diaphragm so adjusted that the lens is quite feebly illuminated, though this is not true for photographic work. Let us now suppose that there is a globular mass of air (*b*) in front of the lens of slightly greater optical density than that of the surrounding air. The rays of light going through the upper portion of this denser mass will be bent down, and will form an image of the spark below the diaphragm, allowing more light to enter the telescope from this particular part of the field; consequently, on looking into the instrument, we shall see the upper portion of the globular mass of air brighter than the rest of the field. The rays which traverse the under part of "*b*," however, will be bent up, forming an image of the spark higher up, and wholly covered by the diaphragm, consequently this part of the field will appear black. It will be readily understood, that with the long path between the lens and the image a very slight change in the optical density of any portion of the medium in front of the lens will be sufficient to raise or depress the image above or below the edge of the diaphragm, and will consequently make itself manifest in the telescope.

The importance of using a lens of first-class quality is quite apparent, since variations in the density of the glass of the lens will act in the same way as variations in the density of the medium before it, and produce unequal illumination of the field. It is impossible to find a lens which will give an absolutely even, feeble illumination, but a good achromatic telescope-objective is perfect enough for every purpose. A more complete discussion of the operation of the apparatus will be found in Töpler's original paper in the *Annalen*. The sound-waves, which are regions of condensation, and consequent greater optical density, make themselves apparent in the same way as the globular mass of air already referred to.

They must be illuminated by a flash of exceedingly short duration, which must occur while the wave is in the field of view.

Töpler showed that this could be done by starting the sound-wave with an electric spark, and illuminating it with the flash of a second spark occurring a moment later, while the wave was still in the field. A diagram of the apparatus used is shown in Fig. 76. In front of the lens are two brass balls (a, a), between which the spark of an induction-coil passes, immediately charging the leyden-jar c, which discharges across the gap at e an instant later. The capacity of the jar is so regulated that the interval between the two sparks is about one ten-thousandth of a second. The field of the lens is thus

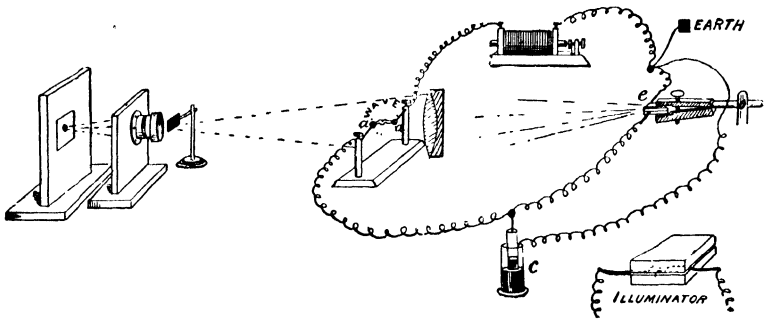


FIG. 76

illuminated by the flash of the second spark before the sound-wave started by the first spark has gone beyond the edge of the lens.

To secure the proper time-interval between the two sparks it is necessary that the capacity of the jar be quite small. This limits the length and brilliancy of the illuminating spark, and with the device employed by Töpler it was impossible to get enough light to secure photographs of the waves. After some experimenting it was found that if the spark of the jar was passed between two thin pieces of magnesium ribbon pressed between two pieces of thin plate glass, a very marked improvement resulted. Pyrex glass or fused quartz will be found more durable. With this form of illuminator five or six times as much light could be obtained as by the old method of passing the spark between two brass balls. The spark is flattened out into a band, and is kept always in the same plane, the light issuing in a thin sheet from between the plates. By this arrangement we secure a light source of considerable length, great intensity, and bounded by straight edges, the three essentials for securing good results. The glass plates, with the ribbon terminals between them, must be clamped in some sort of a holder and directed so that the thin sheet of light strikes the lens:

this can be accomplished by darkening the room, fastening a sheet of paper in front of the lens, and then adjusting the plates so that the paper receives maximum illumination. The image formed by the lens will be found to have very sharp straight edges, on one of which the edge of the diaphragm can be set in such a manner as to allow but very little light to pass when the intervening medium is homogeneous; a very slight change, however, in any portion may be sufficient to cause the entire amount of light passing through that portion to pass below the diaphragm and enter the telescope.

For photographing the waves the telescope is removed and a photographic objective put in its place. A vertical board is firmly clamped behind this in such a position that the image of the balls, between which the sound-spark passed, would be in focus on a plate held against it. This arrangement is used instead of a camera, because it is necessary to move the plate rapidly during the exposure, to prevent the image of more than one wave being formed on the same place. It was found that simply holding the plate in the hand against the vertical board and advancing it slowly from left to right, at the same time giving it a rapid up-and-down motion, answered every purpose.

The images obtained in this manner show the waves in different stages of development, for the time-interval between the two sparks varies between rather wide limits. This is really an advantage, for on a single plate it is possible to pick out a series showing the successive changes in the form of the wave-front produced by reflection, refraction, etc. Each picture shows the circular field of the telescope-lens with the two brass rods crossing it and supporting in the centre the two balls between which the sound-spark passes. The hot air rising from the spark appears in most of the pictures like a puff of steam above the ball.

A few words regarding the apparatus may be helpful to those wishing to repeat the very beautiful experiments of Töpler. An induction-coil capable of giving a three- or four-inch spark is about right, while a good-sized test-tube partly filled with mercury, and standing in a cylinder of mercury, will be found most convenient for a leyden-jar. The balls between which the sound-spark passes should be adjusted so as to obtain almost the maximum spark possible, which will in general be rather less than half as long as the coil will give between its terminals. The best results are obtained when the sparks give off the same crackle found desirable in experiments with Hertz waves. Fresh plates of glass should be put in the illuminator every little while.

It is not at all difficult to get the apparatus to work properly,

and doubtless it could be made to work on quite a small scale with a good photographic objective of rather long focus. The objective of a good-sized spyglass would also give results. Töpler was, I believe, of the opinion that he got more uniform results with an influence-machine than with the coil. He certainly found the time-interval between the two sparks to be more constant. This, however, is no object in photographic work, for the wide variation is the very thing that makes the pictures interesting.

In earlier editions, will be found other experiments with these waves, such as their refraction in tanks filled with hydrogen and carbonic acid, covered with collodion films of soap-bubble thickness. Very much better photographs have been made by Foley by casting the simple shadow of the wave on the plate.[1] Fig. 31.

In examining liquids or solids for striae, or regions of variable refractive index, we can employ a flat gas flame as our source of light, covering the lower part of it with an opaque screen having a straight edge. The lens will form an inverted image of this in front of the objective of the viewing telescope, and all but a strip half a millimetre or so in width is to be cut off from above by a second screen. The object to be examined is placed immediately in front of the lens. A piece of ordinary window glass makes a good object. The heated air rising from the hand can also be seen, and if a tank made of optical glass, filled with warm water, is placed before the lens, a drop lifted out and allowed to fall back can be seen descending through the liquid: the change in the refractive index is obviously due to the cooling by evaporation. Opaque objects placed before the lens appear with brilliantly illuminated margins, the light in this case being diffracted: with the arrangement of screens described only the upper and lower edges appear illuminated, since lateral deflection of the rays is without effect. The method is an extremely useful one, and can be applied to many lines of investigation, and the student should be thoroughly familiar with its possibilities.

Invisibility of Objects. — Opaque substances are seen by the light reflected from their surfaces; transparent substances in part by reflected light and in part by transmitted light. If we analyze carefully the appearance of a cut-glass decanter stopper we shall find it to be extremely complicated. Each facet reflects the image of some object in the room from its surface, and in addition to this shows some other object by refracted rays which have entered some other facet, these latter being in general more or less spread out into a spectrum by dispersion. If the stopper is wholly or in part made of colored glass, the refracted rays passing through the

[1] *Physical Rev.*, *35*, 373, 1912.

colored portions are modified by absorption, and affect the appearance. This remarkable complex, we say, looks like a stopper, and unless we try to paint a picture of it, or have our attention drawn to the details, we are apt to regard its appearance as quite simple.

We thus see that reflection, refraction, and absorption all play a part in making objects visible. It is interesting to examine into the conditions under which objects are invisible. If they are immersed in a medium of the same refractive index and dispersion, reflection and refraction disappear; and if they possess in addition the quality of perfect transparency, they will be absolutely invisible, the light rays passing through them without any modification either in intensity or direction. Could a transparent solid be found whose refractive index was the same as that of air, objects made of it would be invisible. The effect of immersing a transparent solid in a medium of similar optical properties is usually illustrated by dipping a glass rod into Canada balsam or oil of cedar, the immersed portion being practically invisible. A still better medium can be made by dissolving chloral hydrate in glycerine by the aid of heat. Only a little glycerine should be taken, as it is necessary to dissolve some eight or ten times its volume of the chloral before the solution acquires the right optical density. A glass rod, if free from bubbles or striae, becomes absolutely invisible when dipped in the liquid, and if withdrawn presents a curious appearance, the end appearing to melt and run freely in drops.

As a matter of fact, transparent objects are visible only by virtue of non-uniform illumination, as is pointed out by Lord Rayleigh in his article on optics in the *Encyclopædia Britannica*. If the illumination were the same on all sides they would be invisible, even if immersed in a medium of very different optical index. A condition approaching uniform illumination might, he says, be attained on the top of a monument in a dense fog. The author has devised a very simple method of showing this curious phenomenon, which, in brief, is to place a cut-glass decanter stopper within a hollow globe, the interior surface of which is painted with Balmain's luminous paint, and view the interior through a small hole.

CHAPTER IV

THE ABSORPTION AND DISPERSION OF LIGHT

The Absorption of Light. — The transmission of light through a material medium is always accompanied by a certain amount of absorption, regardless of the color or wave-length of the light. Media which we commonly speak of as transparent, if not employed in too great thickness, transmit without appreciable absorption the range of wave-lengths comprised within the region of the visible spectrum.

In general, however, they exercise powerful absorption in the infra-red and ultra-violet regions, and if a sufficiently great thickness is employed, absorption will be found present even in the range of visible radiations. Pure water, which is one of the most transparent substances which we have, in long columns appears distinctly blue, showing that it absorbs more or less completely the red end of the spectrum. The same is true of most varieties of glass. The definition "transparent" is thus seen to be purely arbitrary, there being no such thing in nature as a perfectly transparent substance.

The character of the absorption exerted by any substance can be best observed by receiving the transmitted light on the slit of a spectroscope, when dark regions will be seen in the spectrum, corresponding in position to the wave-lengths absorbed. If the absorbing medium is moulded into the form of a wedge, which is placed in contact with the slit of the instrument, we can observe at once the effect of increased thickness, the form of the absorption curve being pictured in the spectrum. In general, it will be found that as the thickness increases, the absorption band widens out. One edge of the spectrum shows us the absorption of a thin layer, the other edge that of a thick layer, intervening portions corresponding to intermediate thicknesses. The resultant curve is sometimes symmetrical, but more often not so, and we shall see, when we come to consider the theory of absorption, that the form of this curve depends upon a number of different factors. The absorption spectra of about 150 aniline dyes have been photographed and published in the form of an atlas by Uhler and Wood. A wedge-shaped layer of the liquid was used, contained in a quartz cell which was placed in contact with the slit of a large grating spectroscope.

Photographs were in this way obtained showing the position and forms of all absorption bands, both in the visible and ultra-violet regions. Three of these pictures are reproduced in Fig. 77, and show the absorption of nitroso-dimethyl-aniline, auramine, and potassium permanganate. We shall first examine the phenomenon of absorption in a general way, and then in its relation to other closely related phenomena, such as dispersion, emission, and the transformation of the absorbed radiations into other types of energy.

At the beginning of the subject we shall find it convenient to distinguish between two types of absorption: general, in which the absorbing power is very nearly the same for all wave-lengths, at

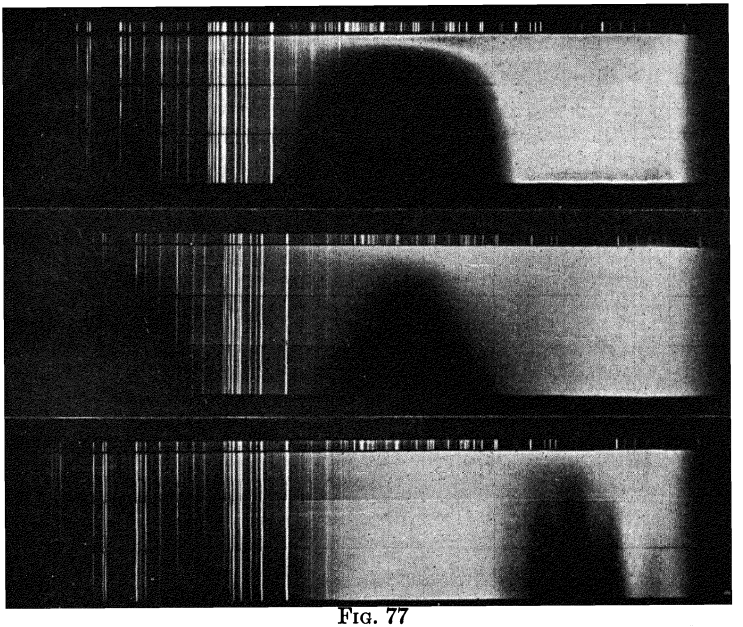


FIG. 77

least over a fairly wide range; and selective, when the absorbed region is more or less limited in extent. The absorption of metal films and lampblack represents the first type fairly well. The light transmitted through thin layers differs but slightly in its composition from the original light, and exhibits therefore but little color. Of course there are exceptions, for, as is well known, thin films of gold transmit an excess of green light, while silver is fairly transparent to the ultra-violet. Aniline dyes, and, in fact, all

colored media, represent the second type, certain colors being freely transmitted, while others are strongly absorbed. When we come to consider the theory of the phenomena, we shall see that the causes of the absorption are radically different in the two cases, though in many cases both conditions may occur simultaneously in one and the same medium.

Laws of Absorption: Lambert's and Beer's. — Lambert's law states that each layer of equal thickness absorbs an equal fraction of the light which traverses it. If we consider layers of the thickness of a single molecule, we can say that each molecule absorbs an equal fraction of the light which passes by it.

$$J = J_0 e^{-Kx} \text{ (Lambert's law)}$$

expresses this relation, J being the intensity of the light after passage through a layer of thickness x, and J_0 the intensity of the light *entering* the layer, *i.e.* the original intensity reduced by reflection at the first surface. The light emerging at the second surface also suffers a reduction by reflection. Formulae taking care of these losses will be given presently.

In a solution, then, the absorption depends upon the concentration and thickness of the layer traversed. Unit layer and unit concentration absorb in the same degree as a layer of thickness 2 with half the concentration.

Calling the absorption coefficient of unit concentration a, the thickness x, and the concentration c, we have, if J_0 is the entering intensity,

$$J = J_0 e^{-acx} \text{ (Beer's law).}$$

No exceptions have ever been found to Lambert's law which could not be attributed to experimental error. Beer's law holds, however, only when the absorbing power of a molecule is not influenced by the proximity of its neighbors, which is not always the case. Cobalt sulpho-cyanate is one of the most striking exceptions, concentrated solutions being sky-blue while dilute solutions are reddish-purple. In this case it is not the proximity of the cobalt sulpho-cyanate molecules to one another, for the purple solution can be changed to blue by adding crystals of calcium chloride to it, or by the addition of potassium sulpho-cyanate, the action being one of dehydration. A drop or two of a solution of cobalt chloride or nitrate mixed with a concentrated solution of potassium or ammonium sulpho-cyanate in a test-tube gives a blue solution, while the same amount added to an equal volume of a dilute solution gives the purple color.

A convenient method of looking for exceptions to Beer's law is to put equal amounts of a concentrated solution in two flat-

bottomed test-tubes, fill one up with water and hold both above a sheet of white paper, looking down at the light through the tubes.

Determination of Extinction Coefficient. — Suppose that we have a partially transparent plate, say of colored glass, and wish to determine the fractional part of the light of any wave-length lost by absorption when the light traverses the plate. In addition to this loss there will be the loss by reflection at each surface which must be taken into account. If I is the initial intensity of the light, and R the reflection coefficient and A the absorbed fraction of the light entering the plate, the intensity emerging from the plate is

$$\begin{matrix}\text{Entering} & \text{Absorbed} & \text{Reflected at Second Surface}\end{matrix}$$
$$I_1 = I - IR - (AI - AIR) - (IR - IR^2 - AIR + AIR^2)$$

or

$$I_1 = I(1-R)^2(1-A)$$

The loss by reflection can be eliminated by immersing the plate in a glass cell filled with a liquid of the same refractive index, or by measuring the absorption of two different thicknesses of the substance, in which case if I_1/I and I_2/I are intensities transmitted by the two layers ($I_2 < I_1$). Their quotient I_2/I_1 is the fraction absorbed by a layer equal to the difference in thickness. The absorption can also be expressed in terms of another constant κ, such that the amplitude of the light-wave of length λ in the medium decreases in the ratio $1 : e^{-2\pi\kappa}$ in traversing a layer of thickness λ. If the thickness of the layer is d, the ratio expressing the decrease of amplitude is $1 : e^{-2\pi\kappa(d/\lambda)}$. If λ is for vacuum, $1 : e^{-2\pi n\kappa(d/\lambda)}$.

Now the intensity of the light is measured by the square of the amplitude, and the intensity therefore decreases in the ratio $1 : e^{-4\pi\kappa(d/\lambda)}$. To avoid the error due to reflection from the surfaces of the layer, or the glass plates between which it is held, it is customary to employ layers of different thicknesses. Let these thicknesses be d_1 and d_2, and the intensity of the incident light be J_0. The intensities after traversing the two layers will be $J_1 = J_0 e^{-4\pi\kappa(d_1/\lambda)}$ and $J_2 = J_0 e^{-4\pi\kappa(d_2/\lambda)}$

$$\log J_1 = \log J_0 - 4\pi\kappa \frac{d_1}{\lambda} \log e,$$

$$\log J_2 = \log J_0 - 4\pi\kappa \frac{d_2}{\lambda} \log e,$$

$$\log J_1 - \log J_2 = \frac{4\pi\kappa}{\lambda} \log e(d_2 - d_1).$$

$$\log \frac{J_1}{J_2} = \frac{4\pi\kappa}{\lambda}(d_2 - d_1) \log e.$$

From this equation we can calculate the extinction coefficient κ, by measuring the intensities of the transmitted beams with a spectro-photometer. The layers of different thickness are best obtained by pouring the liquid into a glass cell containing a glass plate which reduces the thickness of the layer along the bottom of the cell. The spectro-photometer best adapted to the purpose is the instrument designed by Vierordt. It is provided with a double slit, before which the cell is placed in such a position that the dividing line between the two layers coincides with the junction of the two slits. The intensities of the two spectra, which lie one above the other, can be made equal for any value of λ by altering the widths of the slits. Equal illumination is obtained when the slit widths b_1 and e_2 are inversely proportional to the intensities of the illuminating beams, that is, when

$$\frac{J_1}{J_2}=\frac{b_2}{b_1}.$$

Our equation now takes the form

$$\log\frac{b_2}{b_1}=\frac{4\pi\kappa}{\lambda}(d_2-d_1)\log e,$$

$$\kappa=\frac{1}{4\pi(d_2-d_1)\log e}\,\lambda\log\frac{b_2}{b_1}.$$

The absorption coefficient which we have called k is given by

$$k=\frac{1}{2(d_2-d_1)\log e}\log\frac{b_2}{b_1},\text{ since }\kappa=\frac{k\lambda}{2\pi}.$$

Body Color and Surface Color. — The colors of most natural objects result from absorption. The light penetrates their surfaces and then suffers internal reflections or refractions and emerges robbed of the rays which are most strongly absorbed. If this is to happen, it is clear that the substance must not be homogeneous, otherwise the reflections and refractions, which return the unabsorbed light, will not occur. It is thus not strictly correct to say that colored pigments reflect certain colors more strongly than others. If the pigment particles formed a continuous and homogeneous medium, no color whatever would appear in the reflected light, which would be white. If any color appeared, as it might if the pigment were a very powerful absorbent, it would be the tint complementary to the one exhibited by the powder. Cases of this nature we shall consider presently.

Since pigments produce color by absorption, it is at once apparent why a mixture of two pigments does not exhibit the color which we should obtain if we actually mixed the colored lights

which they appear to reflect. The light reflected from the mixture is the residual color which remains after the dual absorption has taken place. For example, if we mix yellow light and blue light, we get white, while a mixture of a blue and yellow pigment appears green. The reason of this is, that the yellow pigment absorbs the blue and violet, the blue pigment the red and yellow, the mixture absorbing everything except the green.

The nature of pigments can be well studied by preparing a number of beads of fused borax, colored with varying amounts of cobalt. If we powder a bead which appeared bright blue by transmitted light, we shall find that the powder is nearly white, the reason being that the light in this case does not penetrate a sufficient thickness of the absorbing medium. A bead colored so dense as to appear black will, however, furnish us with a blue pigment when it is reduced to powder. Pigments, then, are very powerful absorbing media, and, if they could be obtained in homogeneous masses, would be intensely opaque, even in fairly thin sheets.

If we go on increasing the aborbing power, we shall finally observe a phenomenon of a different nature. The color, instead of being absorbed, is selectively reflected. Substances which possess this property are said to exhibit surface color. The aniline dyes are excellent examples, one which in solution absorbs green light, appearing purple by transmitted light, in the solid state reflects green light selectively.

Coefficient of Transmission: Dichromatism. — If the absorbing medium is homogeneous, the quantity of light of a given wave-length which is absorbed will be proportional to the thickness of the medium traversed. If we represent the intensity of the light that enters the front surface of the medium by I_0, the intensity after transmission through unit thickness can be represented by $I_0 a$, in which a is a fraction depending on the nature of the medium and the wave-length of the light. If the same fraction is absorbed by each successive layer, it is clear that the intensity, after traversing a thickness x of the medium, will be Ia^x, the quantity a being called the coefficient of transmission.

The coefficient of transmission varies with the color, and the emergent light is therefore colored. In the case of most absorbing media the color of the transmitted light does not depend to any great degree on the thickness, the depth or saturation merely increasing. In some cases, however, the color depends on the thickness, thin layers, for example, appearing green, and thick layers red. Such substances are said to exhibit dichromatism. Some of the aniline dyes, or mixtures of them, show the phenomenon. Thin layers of a solution of cyanine appear blue, thick layers red. The

addition of a little nitroso-dimethyl-aniline to the solution gives us a green-red dichromatic liquid, as has been shown by Pflüger. The explanation of the change of tint is very simple. Suppose we have a substance which absorbs the yellow and blue. The transmitted light then consists of a mixture of red and green. Let us assume, as is usually the case, that the visual sensation produced by the green is greater than that due to the red. Writing for these intensities $I_g > I_r$, and assuming that the coefficient of transmission of the green is less than that of the red, $a_g < a_r$, it follows that for small thicknesses $I_g a_g^x$ will be greater than $I_r a_r^x$, while the reverse will be true for thick layers. They will be equal for a thickness x given by the equation

$$I_g a_g^x = I_r a_r^x,$$

or taking the logarithms of both sides,

$$x = \frac{\log I_g - \log I_r}{\log a_r - \log a_g}.$$

For this thickness the intensities of the red and green will be equal, and the color of the transmitted light will appear to be yellow, for a mixture of red and green light produces the sensation of yellow when mixed in proper proportions. An excellent mixture for illustrating this can be formed by dissolving "brilliant green" and "naphthaline yellow" in hot Canada balsam and pressing the mixture between two glass plates in the form of an acute prism. The balsam should be previously boiled down until a drop solidifies on cooling, and the dyes should not be added until the fluid has cooled somewhat, otherwise they are apt to decompose. The thin edge of the wedge will appear green, the thick edge red, and the intermediate portions, where we have equality of transmission, yellow.

If some of the same mixture is moulded into a prism of 20 or 30 degrees angle, the mechanism of dichromatism can be beautifully shown by observing a lamp flame through it. The prism will show the red image well separated from the green, and the latter will be found to be extinguished more rapidly than the former as the prism is moved laterally before the eye.

Our equations for color show us as well that the color of the transmitted light, for a given thickness, will vary with the composition of the original light. If the plate of stained balsam is examined by gas-light and then by daylight, it will be found that parts of it will appear red in the former and green in the latter case. A solution of cyanine and nitroso-dimethyl-aniline in alcohol appears red by lamplight and bottle-green by daylight. The same

phenomenon is exhibited by the gem Alexandrite, found in the Urals.

Absorption by Porous Surfaces. — The absorption of light at surfaces formed of lampblack or finely divided metals such as platinum black is accompanied by very little reflection. The question naturally arises as to why a metal with a high reflecting power can, under certain conditions, appear nearly dead black. The roughness of the surface will not account for the fact, for matt surfaces of electrolytically deposited silver appear as white as plaster. Chemically precipitated silver, on the other hand, appears black.

The phenomenon is to be referred to the condition of the surface. Consider a bunch of polished steel needles, turned with their points towards the light. Rays falling upon the surface formed by the points will be reflected down into the interstices between the needles, and practically none of it will escape or be reflected back. A portion is absorbed at each reflection, and the large number of reflections reduce the intensity rapidly to zero. Surfaces of lampblack (soot) and platinum black can be considered as porous, the pores acting as light "traps." The energy penetrates into the spongy mass by multiple reflection, and is speedily transformed into heat by absorption. If the pores are closed up by compressing the mass, its reflecting power is increased or wholly restored. If, too, the angle of incidence is too large to admit of downward reflection into the mass, the light is more or less completely reflected. A surface of smoked glass reflects very perfectly at large angles of incidence, and at the same time yields a sharply defined image of the source of light, as we have seen in the Chapter on Huygens's Principle. The extreme blackness of black velvet is produced in this way, and Pfund has recently obtained an even blacker surface by condensing the vapor of tellurium on the walls of an exhausted tube, the material crystallizing in the form of a thick forest of vertical needles. The tube must be opened of course to observe the effect.

Dispersion of Light. — In our treatment of refraction we have assumed a constant retardation of the waves for a medium of given refractive index. We have seen that the velocity of light in free space is independent of the color or wave-length. Such, however, is not the case in refracting media, for here the waves not only travel slower than in free space, but waves of different length travel with very different velocities. In all such media as air, water, and glass, the long waves travel faster than the short ones; consequently the deviation of the ray, or the angle through which the wave-front turns when encountering the boundary of the

medium, depends on the color of the light as well as on the optical density of the medium.

When white light enters a transparent medium, the long red waves forge ahead of the green ones, which in their turn get ahead of the blue. If we imagine an instantaneous flash of white light traversing a refracting medium, we must conceive it as drawn out into a sort of linear spectrum in the medium, that is, the red waves lead the train, the orange, yellow, green, blue and violet following in succession. The length of this train will increase with the length of the medium traversed. On emerging again into free space the train will move on without any further alteration in its length.

We can form some idea of the actual magnitudes involved in the following way. Suppose we have a block of perfectly transparent glass (of ref. index 1.52) twelve miles in thickness. Red light will traverse it in $\frac{1}{10,000}$ of a second, and on emerging will be about 1.8 miles in advance of the blue light which entered at the same time. If white light were to traverse this mass of glass, the time elapsing between the arrival of the first red and the first blue light at the eye would be less than $\frac{1}{6000}$ of a second. Michelson's determination of the velocity of light in carbon bisulphide showed that the red waves gained on the blue waves during their transit through the tube of liquid and the colors were spread out into a spectrum by the rotation of the mirror, as they are by a glass prism which deviates the wave-fronts of different frequencies through different angles. The absence of any change of color in the variable star Algol furnishes direct evidence that the blue and red rays traverse space with the same velocity. In this case the distance is so vast, and the time of transit so long, that the white light coming from the star during one of its periodic increases in brilliancy, would arrive at the earth with its red component so far in advance of the blue that the fact could be easily established by the spectro-photometer or even by the eye.

Inasmuch as the deviation of a ray of light depends on the change of velocity of a wave on going, say, from a rare into a denser medium, we infer that those rays which are deviated the most, namely the violet, suffer the greatest change of velocity or move the slowest. Later on, when we come to the study of interference, we shall find other evidence that such is the fact.

Newton was the first to systematically study the phenomenon of dispersion. He discovered that ordinary white light was made up of different colors which could be separated from each other by passing the light through a prism. His most complete and convincing experiment may be briefly summed up as follows. The light of the sun was admitted to a darkened room through a small

hole in the shutter, and the narrow beam passed through a prism. Instead of a round white image of the sun, there now appeared on the screen a colored band of light or spectrum, made up in reality of an infinite number of differently colored images of the sun superposed, but slightly displaced with reference to one another.

A small perforation was made in the screen which allowed light of approximately a single color to pass. This ray was transmitted through a second prism, and was found to form a fairly sharp image of the sun on a second screen, proving that monochromatic light suffers no decomposition or dispersion in the prism. By slightly turning the first prism, the spectrum could be moved so as to allow light of any color to pass through the perforation and be refracted by the second prism. Newton found that the colored image of the sun on the screen changed its position with every

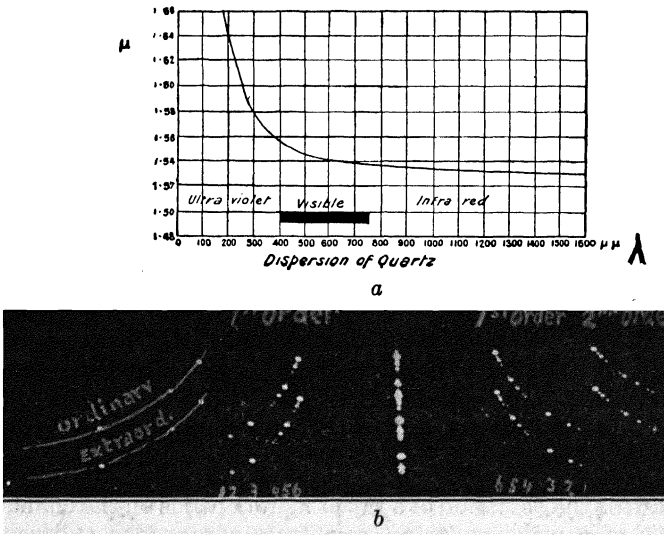


FIG. 78

change of color, the deviation being greatest when the color was violet, and least when it was red.

The refractive index of a substance varies then with the wave-length of the light employed. To determine the relation between the two we can measure the index of a prism for light of known wave-lengths, *e.g.* the bright lines in the spark spectrum, or the dark lines in the solar spectrum, and plot the results on coördinate paper, taking the refractive indices as ordinates, and the wave-lengths as abscissae.

We shall find that the refractive index increases more rapidly than the wave-length decreases as we approach the violet, the curve having the general form shown in Fig. 78*a*, which is the dispersion curve for quartz. Having plotted such a curve for a given prism, we can determine the wave-length of any other line in the spectrum by determining the refractive index for the line, and finding the corresponding ordinate on the curve. If a prism spectroscope is to be used for wave-length determinations, it must be calibrated in this manner, for different samples of glass have very different dispersive powers. Were the deviations proportional to the wave-lengths, the curve would be a straight line, and we should have what is known as a normal spectrum. Such a spectrum can be formed by a diffraction grating, but never by a prism. The dispersion curve can be shown experimentally in the following way. Let a vertical normal spectrum, formed by a diffraction grating, be viewed or projected through a prism standing with its refraction edge vertical. The entire spectrum will be deviated by the prism, but the deviation will increase very rapidly as we near the blue, the spectrum being bent into a curve. Figure 78*b* shows dispersion curves made in this way. A quartz prism in which the rays travelled perpendicular to the optic axis was mounted in front of the lens of a camera, and a small glass diffraction grating placed behind it. The source of light was an "end-on" helium tube placed at a distance of three metres. Instead of spectrum lines we have small circles of light, corresponding to the different wave-lengths emitted by the glowing helium gas.

The quartz prism is doubly refracting and consequently gives two spectra, and the photograph shows the dispersion curve for each. The straight spectrum is made up of the superposed spectra which the quartz prism gives of the central image formed by the diffraction grating. This will be better understood after we come to the subject of diffraction. This method of "crossed prisms," due to Newton, is of use in studying the remarkable phenomenon of anomalous dispersion, which we shall come to presently.

Newton came to the erroneous conclusion that the dispersion was proportional to the refraction, that is to say that substances of high refractive index had great dispersive powers, or gave wide spectra, while the reverse was true for substances of low refractive index. While this is apt to be the case, it is not always true, for we find that there are substances the mean refractive indices of which are small, while their dispersive powers are large, and *vice versa*.

Achromatism. — The fact that dispersion is more or less independent of refractive index makes it possible to arrange two prisms

of different kinds of glass, with their refracting angles turned in opposite directions, which shall have the power of deviating a ray without spreading it out into a spectrum. One of the prisms almost entirely annuls the dispersion of the other, without entirely annulling the deviation, a thing which Newton considered impossible. Such a combination is known as an achromatic prism. Let us see just how such a system operates.

Flint glass has a much higher dispersive power in proportion to its mean refractive index than crown glass. The refractive indices of the two glasses for red, yellow and bluish-green light of wave-lengths corresponding to the *C*, *D* and *F* lines in the solar spectrum are as follows:

	C	*D*	*F*
Flint glass	1.630	1.635	1.648
Crown glass	1.527	1.530	1.536

If prisms of small angle are employed we can write the deviations for these colors as proportional to the refractive indices less 1, that is for flint glass the distance from a point on a screen where the direct ray falls, to the points where the red-green and blue rays fall when the prism is put in the path of the light, will be 630, 635 and 648.

The length of the spectrum, or rather the distance between the *C* and *F* lines, is obviously 648 − 630 or 18.

For crown glass the distances will be 527, 530 and 536, and the distance between the *C* and *F* lines will be 536 − 527 or 9. The dispersion of the flint glass is therefore double that of the crown glass. If now we make a crown-glass prism of twice the angle of the flint-glass prism, the distance between the *C* and *F* lines will be the same as with the flint prism, while the distances of the lines from the spot where the direct ray falls will be twice as great as before, or 1054, 1060 and 1072.

Suppose now we place the two prisms together with their refracting angles turned in opposite directions. The crown prism alone would shift the *F* line to a distance of 1072, but the flint prism shifts it back a distance of 648, and its resulting position is 1072 − 648 or 424 from the spot where the direct ray would fall. The *C* line would be deviated by the crown prism to a distance 1054, but the flint one moves it back 630, and its position is 1054 − 630 or 424. The *C* and *F* lines are thus deviated the same amount, and the dispersion is annulled so far as these two colors are concerned. The combination is achromatic for red and greenish-blue light, deviating both to a distance of 424. Let us now see if the yellow light falls in the same place. The position of the *D* line will be

given by 1060 − 635 or 425, that is, it will be deviated a very little more than the C and F lines, consequently the combination is not perfectly achromatic. By means of two prisms it is possible to bring any two parts of the spectrum together, the other colors lying a little to the right or left of the superposed portions, forming what is known as the secondary spectrum.

The general rule to follow in the construction of an achromatic prism is as follows. To bring any two lines of the spectrum together, the angles of the two prisms must be so proportioned that the distance between the lines in question is the same for each prism. Were the distances between the other lines the same for both prisms, the combination would be truly achromatic, but such is not the case, owing to the irrationality of dispersion.

Achromatic prisms are of very little practical use, but the principle is of great importance in connection with achromatic lenses.

Partial Achromatization by a Grating. — A diffraction grating gives a normal spectrum, that is, one in which the distances between the lines are proportional to their difference of wave-length. Some very beautiful effects can be seen by viewing a distant gas flame through a flint-glass prism of 60°, and a glass diffraction grating of about 2000 lines to the inch. The grating alone gives two sets of spectra turned in opposite directions. Those on one side of the central image are still further expanded by the prism, while those on the other are more or less closed up. By varying the angle at which the prism is placed, and tilting the grating so as to vary its dispersion, we can achromatize for the middle of the spectrum and obtain a green image of the flame, with a red and blue spectral band extending out from it as a purple flare of light. If we turn the grating so that its direction of dispersion makes a small angle with that of the prism we shall see how this colored image and the purple flare has been produced: the spectrum has been bent back upon itself, and resembles a portion of an ellipse. As we rotate the grating still further, this curve gradually opens out, acquiring the form of the dispersion curve of the prism when

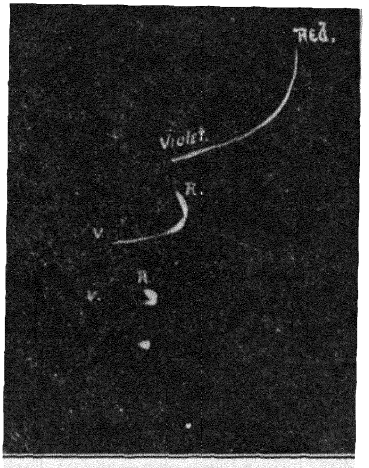


FIG. 79

the rotation has reached 90°, that is when the prism and grating are crossed. Photographs of the spectrum obtained in this way, with the grating in four different positions, are reproduced in Fig. 79. It is clear that when the curved spectrum is flattened together (which occurs when the lines of the grating are vertical) the red and blue will be superposed, while the middle or green portion will be pretty well concentrated at a single spot, giving the green image of the flame.

Direct-Vision Prisms. — By referring to the table of refractive indices for crown and flint glass it is easy to see how a combination of two prisms can give dispersion without deviation, that is, yield an undeviated spectrum. If, instead of giving the crown prism an angle double that of the flint, we make it 1.2 times as great, and make the same calculation as before, we shall find that we have a spectrum the length of which is 7, and the centre of which falls on the spot where the undeviated ray would fall. Such a combination is known as a direct-vision prism, and is employed in cases where any considerable deviation would be detrimental, as when compactness of the instrument is desirable.

Achromatic Lenses. — We are now in a position to consider the principle on which the achromatic lens is made. Any lens can be considered as a prism of varying angle, or rather as a solid formed by the rotation of a thin section of a curved prism around its base. Since the distance of the focus of a lens from its centre depends on the deviation of the rays, it follows that the focus will be different for the different colors, the blue rays which are bent the most meeting nearest the lens, and the red, which are bent to a less degree, coming together farther away, an effect known as chromatic aberration. What we require is a combination which will produce an equal deviation, and consequently a common meeting point for rays of all colors. If we can arrange two prisms of crown and flint glass which will give deviation without dispersion, we can in the same way, by employing a double convex lens of crown and a plane concave of flint glass, give exactly the same deviation to two colors widely separated in the spectrum, and very nearly the same deviation to the other colors, with the result that rays of different refrangibility come together at very nearly the same point.

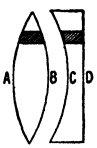


FIG. 80

Reference to Fig. 80 will make the analogy between the achromatic lens and prism clear. The blackened parts indicate how each portion of the lens combination can be considered as two opposed prisms. We found that in the case of the prism the ratio between the angles was 1 : 2, and applying this to the lens it is

easy to see that if the surfaces A, B and C have the same curvature, the surface D of the flint lens must be plane, since the angle of the elementary prismatic portion of the flint lens must be everywhere $\frac{1}{2}$ that of the opposed elementary crown prism. Just as by employing two prisms we could unite two lines of the spectrum, so by the use of two lenses we can bring rays of any two different colors to the same focus.

Resolving Power of Prisms. — By the resolving power we mean the ability of the prism to show a line as double when two different wave-lengths are present. This will obviously depend upon two things: the narrowness of the images of the lines and their distance apart. The former depends upon the width of the beam of light, that is the horizontal aperture of the instrument, for as we shall see when we come to the Chapter on Diffraction the image of a very narrow slit broadens by diffraction as the aperture of the lens is reduced, and it always has a finite width. It is moreover accompanied by fainter images on each side, with dark minima between them. Now Lord Rayleigh has shown that if we have two wave-lengths λ_1 and λ_2 which are very nearly equal, to see the slit image double, the image due to λ_2 must be separated from that due to λ_1 by at least the distance of the first minimum bordering the image formed by λ_1. This lies in such a direction that the path difference between the disturbances coming from opposite sides of the aperture, instead of being zero, the condition at the centre of the image, differs by exactly one wave-length.

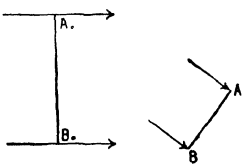


FIG. 81

Though the following treatment by Lord Rayleigh will hardly be understood until after the Chapter on Diffraction has been read, it is introduced at this point for future reference.

"Let A_0B_0 (Fig. 81) be a plane wave-surface of the light before it falls upon the prisms, AB the corresponding wave-surface for a particular part of the spectrum after the light has passed the prism, or after it has passed the eye-piece of the observing-telescope. The path of the ray from the wave-surface A_0B_0 to A or B is determined by the condition that the optical distance, represented by $\int \mu\, ds$, is a minimum; and as AB is by supposition a wave-surface, this optical distance is the same for both points. Thus

$$\int \mu\, ds \text{ (for } A) = \int \mu\, ds \text{ (for } B). \qquad (2)$$

"We have now to consider the behavior of light belonging to a neighboring part of the spectrum. The path of a ray from the wave-surface A_0B_0 to A is changed; but in virtue of the minimum

property the change may be neglected in calculating the optical distance, as it influences the result by quantities of the second order only in the change of refrangibility. Accordingly the optical distance from A_0B_0 to A is represented by $\int (\mu+\delta\mu)ds$, the integration being along the path $A_0 \ldots A$; and, similarly, the optical distance between A_0B_0 and B is represented by $\int (\mu+\delta\mu)ds$, where the integration is along the path $B_0 \ldots B$. In virtue of (2) the difference of the optical distance is

$$\int \delta\mu\, ds \text{ (along } B_0 \ldots B) - \int \delta\mu\, ds \text{ (along } A_0 \ldots A). \qquad (3)$$

"The new wave-surface is formed in such a position that the optical distance is constant; and therefore the *dispersion*, or the angle through which the wave-surface is turned by the change in refrangibility, is found simply by dividing (3) by the distance AB. If, as in common flint-glass spectroscopes, there is only one dispersing substance, $\int \delta\mu\, ds = \delta\mu s$, where s is simply the thickness traversed by the ray. If we call the width of the emergent beam a, the dispersion is represented by $\delta\mu(s_2-s_1)/a$, s_1 and s_2 being the thicknesses traversed by the extreme rays. In a properly constructed instrument s_1 is negligible, and s_2 is the aggregate thickness of the prisms at their thick ends, which we will call t; so that the dispersion (θ) is given by

$$\theta = \frac{t\delta\mu}{a}. \qquad (4)$$

"The condition of resolution of a double line whose components subtend an angle θ is that θ must exceed λ/a. Hence from (4), in order that a double line may be resolved whose components have indices μ and $\mu+\delta\mu$, it is necessary that t should exceed the value given by the following equation:

$$t = \frac{\lambda}{\delta\mu}, \qquad (5)$$

which expresses that the relative retardation of the extreme rays due to the change of refrangibility is the same, namely one wavelength, as that incurred without a change of refrangibility when we pass from the principal direction to that corresponding to the first minimum of illumination.

"That the resolving-power of a prismatic spectroscope of given dispersive material is proportional to the total thickness used, without regard to the number of angles, or setting of the prisms, is a most important, perhaps the most important, proposition in connection with this subject. Hitherto in descriptions of spectroscopes far too much stress has been laid upon the amount of

dispersion produced by the prisms; but this element by itself tells nothing as to the power of an instrument. It is well known that by a sufficiently close approach to a grazing emergence, the dispersion of a prism of given thickness may be increased without limit; but there is no corresponding gain in resolving-power. So far as resolving-power is concerned, it is a matter of indifference whether dispersion be effected by the prisms or by the telescope."

The expression for the resolving-power of a prism is usually written in the form $\lambda/\delta\lambda = t(\delta\mu/\delta\lambda)$, which follows at once from (5).

This equation states that two lines of wave-lengths λ and $\lambda+\delta\lambda$ will be just barely separated when the thickness of the prism's base t, multiplied by $\delta\mu/\delta\lambda$, is equal to $\lambda/\delta\lambda$.

As an example we may calculate the thickness of a prism which will just separate the sodium lines. We must first get a value for $\delta\mu/\delta\lambda$.

This we can do by differentiating the dispersion formula

$$\mu = A + \frac{B}{\lambda^2},$$

$$\frac{\delta\mu}{\delta\lambda} = -\frac{2B}{\lambda^3}.$$

The value of B varies with the material of the prism. Let

$B = .984 \times 10^{-10}$, which is for extra dense flint,

$\lambda = 5.890 \times 10^{-5}$ cms.,

$\delta\lambda = .006 \times 10^{-5}$ cms. (difference between D_1 and D_2).

Therefore $$t = \frac{\lambda^4}{2B\delta\lambda} = \frac{10^{10}\lambda^4}{1.968\delta\lambda} = 1.02 \text{ cms.}$$

The base of the prism must thus be at least a centimetre thick if the sodium lines are to appear separated.

Lord Rayleigh found as a result of a number of experiments that from 1.2 to 1.4 cms. were actually required, depending on the observer.

He also found both from theory and by experiment that a decided improvement in resolving-power resulted from stopping the rays which passed through the centre of the aperture.

Christiansen's Experiment.[1] — While engaged upon some determinations of the refractive indices of white powders by the method of immersing them in liquid mixtures of the same refractive index, Christiansen observed some very remarkable and interesting effects. Owing to the different dispersive powers of the liquid and powder, complete transparency could only be obtained for monochromatic light. If white light was employed the transmitted light

[1] *Wied. Ann.*, Nov., 1884.

was highly colored, the transmitted color corresponding to the particular wave-length for which the two substances happened to have the same refractive index. Finely powdered glass immersed in a mixture of benzol and bisulphide of carbon was found to exhibit the colors well. The powder must be quite free from dirt, the elimination of which is sometimes very difficult. The author has obtained the best results with the powdered quartz, which can be procured from the large chemical houses. The powder is boiled in nitro-muriatic acid to free it from impurities and thoroughly washed in clean water. It is then dried and placed in a test-tube with enough bisulphide of carbon to wet it thoroughly. Benzol is then added a little at a time until the mixture begins to get transparent. It will be found that red light is transmitted first, then yellow, green and blue in succession as more benzol is added. It is best to stop when the transmitted light is yellow. In the general illumination of a brightly lighted room the colors are not very pronounced, and it is best to employ a distant lamp in a fairly dark room as the source of light. If a permanent preparation is desired, the following method gives good results. A quantity of the quartz powder is introduced into a 100-c.c. flask (not more than $\frac{1}{10}$ of the volume of the flask), the neck of which is then drawn down until it has a diameter of only a few mms. The liquid previously adjusted in the manner described is then introduced in sufficient quantity to form a rather thick pasty mass, which will stick in a thick layer to the walls of the flask if it is shaken. The flask is then packed in powdered ice and salt and the neck closed by fusion in the flame. The freezing mixture is necessary on account of the inflammability of the vapor and its tension, and it is a good plan to wrap a towel around the beaker containing the flask and cooling mixture in case of explosion. On removing the flask from the ice it will be found to be quite opaque, owing to the change in the refractive index of the liquid. As the temperature rises red light is transmitted first, and by slightly warming the flask in spots by momentary contact with a flame or even with the fingers all colors of the rainbow may be made to appear simultaneously, the whole appearing like a great opal. The reason of these beautiful temperature changes will be readily understood by reference to Fig. 82.

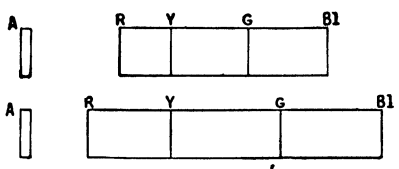


FIG. 82

Suppose A to be a linear source of light which is deviated to the

right and spread out into two spectra by prisms of the same angle, the one composed of quartz, the other of the liquid mixture. The refractive indices having the same value for yellow light, the yellow of one spectrum will fall immediately above the yellow of the other. Since, however, the liquid has a much higher dispersion it will yield a longer spectrum and the other colors will not be in coincidence, or in other words the refractive indices are different for all the other colors. The color transmitted will obviously be the one for which we have coincidence in the above diagram. The other colors will be more or less scattered by irregular refractions and reflections. Suppose now we heat the mixture, the increase of temperature will cause the liquid to expand and its refractive index to decrease, while the effect upon the quartz is comparatively slight. This will mean a shift of the lower spectrum in the diagram towards the left, the green regions of the two spectra coming into coincidence, while a further increase of temperature will bring the blue regions together. The effect of the warming is thus to shift the region of transmission down the spectrum towards the blue.

In general, unless the thickness of the heterogeneous medium is considerable, the light which is not directly transmitted emerges to some extent as diffused light. The color of this diffused light is complementary to the transmitted, and the green image of a lamp flame seen through a thin layer of the paste is surrounded by a purple halo. The colors seen when equal volumes of glycerine and turpentine are shaken together into an emulsion are of similar nature, though erroneously attributed to interference in some textbooks. The opalescent precipitate obtained by the addition of hydrofluorsilicic acid to a solution of potassium chloride has been found by the author to be another case, the color of the transmitted light changing in a most beautiful manner upon the addition of water, which diminishes the refractive index of the liquid, precisely as the rise of temperature did in the case of the benzol mixture. Similar colors were observed in 1896 in a jar of ethyl cinnamate in a Berlin laboratory due to partial saponification. Fuller particulars regarding these curious mixtures will be found in the original papers of Christiansen, and in an interesting paper by Lord Rayleigh.[1]

Determination of the Dispersion of a Substance in the Form of a Powder. — If a transparent substance in the form of a powder be mixed with a liquid of the same refractive index, the whole becomes optically homogeneous, and the opacity resulting from the irregular reflection and refraction of the particles disappears. Owing to

[1] *Phil. Mag.*, xx, 358, 1885.

the irrationality of dispersion it is not possible to obtain a liquid of exactly the same refractive index and dispersion, the mixture being optically homogeneous for a single color only: this color is transmitted as we have seen, while the other colors are scattered, and to a greater or less extent refused transmission. Suppose we wish to determine the dispersion of precipitated potassium fluosilicate, which in the solution of KCl, in which it is formed, shows brilliant opalescent colors by transmitted light. Introduce the mixture into a hollow prism and allow it to stand until the precipitate has settled. Place the prism on the table of a spectrometer, cover the upper part of the prism with a card to cut off the light which passes through the clear liquid and examine the transmitted light with a telescope. It will be found to consist of some definite portion of the spectrum, which can be considerably narrowed by shielding all of the prism except the base. Set the cross hair of the eye-piece on the centre of this band, uncover the upper portion of the prism and note the wave-length of the Fraunhofer line which comes nearest to the cross hair. Determine the refractive index of the liquid for this line in the usual manner, which will be also the refractive index of the powder for the same color. By adding KCl or water we can vary the refractive index of the liquid, making it coincide with that of the powder for the other colors of the spectrum, and in this way the dispersion of the powder can be determined. Compare this with the dispersion of the liquid at such a density, say, that it is optically the equivalent of the powder for green light.

Unless the prism has a thickness of about 10 cms. the transmitted color is very impure in the case of the fluosilicate. In cases such as this more accurate results can be secured by putting the mixture into a long tube closed by glass plates, and making an independent observation of the wave-length of the transmitted light.

The potassium fluosilicate has been found by the author to have the lowest mean refractive index of any known transparent solid, and a dispersion much less than that of water.

Anomalous Dispersion. — In the case of transparent substances the dispersion is said to be normal, that is, the refractive index increases as the wave-length decreases, though the rate of change varies according to the nature of the substance.

In the case of substances which show selective absorption this is not generally the case, the refractive index for the short waves on the blue side of the absorption band being less than the index for the red light on the other side of the band.

This phenomenon has been named anomalous dispersion, but, as

we shall see presently, there is nothing anomalous about it, the so-called normal dispersion being nothing more than a special case of the anomalous. Fox Talbot appears to have been the first to notice the peculiar effect, but his discovery was not followed up. In 1860 Le Roux [1] discovered that a prism containing iodine vapor deviated the red rays more than the blue, the indices at a temperature of 700° C. for the red and violet being 1.0205 and 1.019. Christiansen [2] in 1870 detected anomalous dispersion in the case of an alcoholic solution of fuchsine, which is one of the aniline dyes having a strong absorption band in the green. Of the remaining colors, the red, orange and yellow occur in the same order as in the case of a glass prism. The violet, however, is less refracted than the red, and separated from it by a dark interval. Christiansen's prism was made of two glass plates inclined at an angle of 1 degree, the solution being held between them by capillarity. The subject was next investigated by Kundt, whose papers will be found in *Pogg. Ann.*, 1871, 1872. His observations showed that the phenomenon is to be observed in the case of all bodies which possess what is known as surface color, that is, bodies which selectively reflect certain wavelengths. Kundt applied the method of crossed prisms, due originally to Newton, to the investigation of anomalous dispersion. If a spectrum is formed by a glass prism with its refracting edge vertical, and this spectrum is further deviated by a prism formed of an alcoholic solution of some aniline dye with its refracting edge horizontal, the appearance seen will be similar to that shown in Fig. 83. Kundt established the law that on approaching an absorption band from the red side the refractive index is abnormally increased by the presence of the band, while if the approach is from the blue side the index is abnormally decreased. So great is the difficulty of seeing the effect with the small dispersion obtainable by alcoholic solutions, that the earlier results of Kundt were not at first accepted by some physicists of repute, the effect being attributed to a want

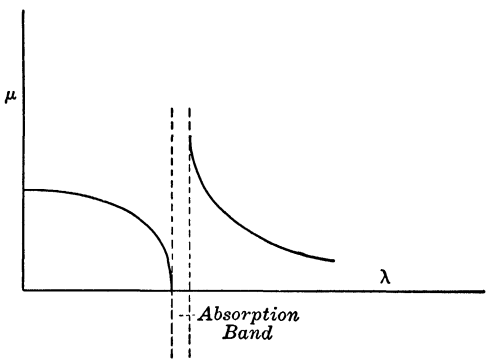


FIG. 83

[1] *Ann. de Chimie et de Physique*, 3d series, vol. xli, p. 285, 1861.
[2] *Pogg. Ann.*, 1870.

of achromatism of the eye. The demonstration by means of crossed prisms, however, removed all doubts regarding the reality of the phenomenon.

Considerable trouble is usually found in repeating Kundt's experiment with fluid prisms.

The phenomenon can be studied to much better advantage by means of prisms formed by squeezing fused cyanine between plates of glass.[1] A certain amount of dexterity is required to make good prisms, which can be acquired only by practice. Small rectangular pieces of thin German plate glass are prepared (measuring about 2×3 cms.), and a thin strip cut from a visiting-card glued along the short side of one. A piece of cyanine [2] about the size of a coarse shot is placed near the opposite side, and the edge of the plate heated over a small flame until the dye fuses, holding another coverstrip in the flame at the same time, in order to have both at about the same temperature. The hot edge of the cover is now to be brought down into the cyanine, and the plate gently lowered until the edge rests on the strip of card. The plates must be at once placed under pressure in a small clamp, where they are to remain until cold. The pressure is to be applied close to the refracting edge of the prism only, as shown (Fig. 84). This is very important. Experience is the only guide to the degree of pressure required.

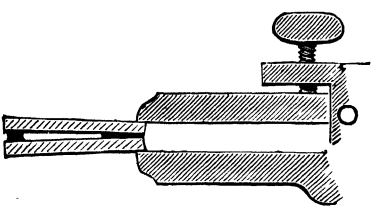


FIG. 84

It will be found that there is a very narrow strip of clear glass at the refracting edge, where the glass plates have come into optical contact. This produces a diffraction-band superposed on the anomalous spectrum, but it is so faint that it is not troublesome. One has only to view a narrow and brilliant light source through the prism, the anomalous spectrum showing colors in the order orange-red, blue, green, the latter being the least deviated.

It is usually necessary to turn the prism slightly to get the green part of the spectrum; that is, the incidence should not be normal.

If a prism of this nature is covered with a small transmission diffraction grating, the lines of which are perpendicular to the edge of the prism, the oppositely curved branches of the diffraction spectra appear most beautifully when an arc light is viewed through the combination. If a grating is not available, the cyanine prism

[1] Wood, *Phil. Mag.*, June, 1901.

[2] Some preparations of cyanine do not fuse. That used for the preparation of prisms was in the form of needle-like crystals, and was prepared by Grüber.

can be mounted over a small aperture in a card and combined with a glass prism of low dispersion, or better a water prism, both being mounted on the table of a spectrometer illuminated with sun or arc light.

Other remarkable cases will be described in the Chapter on the Theory of Dispersion.

Anomalous Dispersion in Its Bearing on Solar Phenomena. — W. H. Julius [1] made the very brilliant suggestion that the "flash spectrum" seen immediately preceding the total phase of a solar eclipse might be due to photosphere light abnormally refracted in the atmosphere of metallic vapors surrounding the sun: in other words, come not from the reversing layer at all, but from the photosphere.

The theory of Julius supposes the sun to be surrounded by an atmosphere of metallic vapors, the density and refractive index of which decrease with increasing distance from the surface. In this atmosphere the rays of light coming from the photosphere will move in curved paths similar to those of rays in our own atmosphere. The reader should refer back to Schmidt's theory of the solar disk.

The refractive index is, however, very small except for wave-lengths very near those which are absorbed by the vapor, consequently the light most strongly refracted, if it could be sorted out and examined with the spectroscope, would resemble very closely the light emitted by the vapors. Julius shows that this sorting out of the more refrangible rays may account for the bright line spectrum usually attributed to the reversing layer, these rays moving in curved paths in the solar atmosphere, thus reaching us after the photosphere has been hidden by the moon.

This phenomenon, namely the production of a bright line spectrum by the anomalous refraction of light from the equivalent of a white-hot source, was reproduced in the laboratory by the author, and independently by Ebert at about the same time. The conditions supposed by Julius to exist at the surface of the sun were imitated as closely as possible, and a spectrum of bright lines was obtained with light from a source showing a continuous spectrum, by means of anomalous dispersion in an incandescent metallic vapor.

For the reproduction of the phenomenon in the laboratory it is necessary to form an atmosphere of metallic vapor in which the refractive index changes rapidly from layer to layer. This was accomplished by allowing the flame of a Bunsen burner fed with metallic sodium to play against the under side of a white plaster

[1] *Astrophysical Journal*, xii, 195.

plate. On looking along the surface of the plate it was seen that a dark space existed between the flame and the cold surface, resembling somewhat the dark space surrounding the cathode of a Crookes's tube. It seemed highly probable that, inasmuch as the temperature of the flame was lowered by contact with the plate,

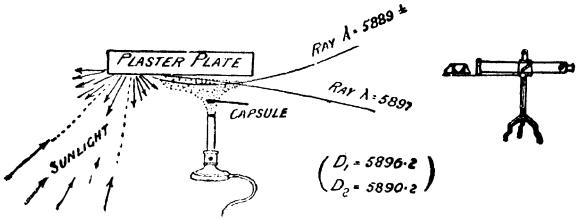


FIG. 85

the density of the sodium vapor would increase very rapidly from the surface of the plate downward. The under surface of the plaster plate having been thus covered with a non-homogeneous layer of sodium vapor, a spot at the edge of the flame was illuminated with sunlight concentrated by a large mirror. This spot radiated white light in every direction and corresponded to the incandescent photosphere of the sun (Fig. 85). A telescope provided with an objective direct-vision prism was directed toward the white spot and moved into such a position that, owing to the reduction in the width of the source of light by foreshortening, the Fraunhofer lines appeared in the spectrum. This represented the stage of an eclipse when only the thin crescent of the sun is visible. The sodium flame appeared superposed on the spectrum, of course. On moving the spectroscope until it was well inside of the plane of the illuminated surface and feeding the flame with fresh sodium, the solar spectrum vanished and there suddenly blazed out two narrow bright yellow lines, al-

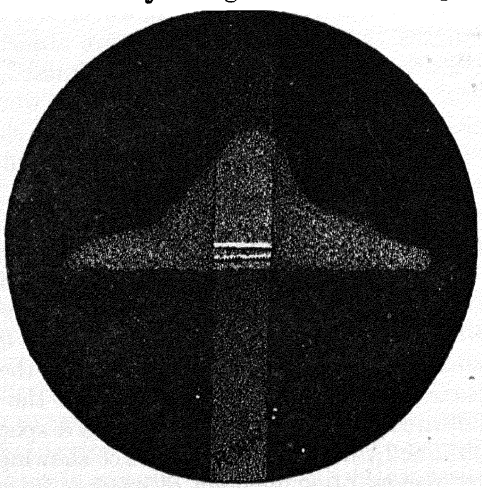


FIG. 86. — FLASH SPECTRUM OF SODIUM PRODUCED BY ANOMALOUS DISPERSION

most exactly in the place of the sodium lines, as is shown in Fig. 86, in which the inverted sodium flame appears on the continuous spectrum. Cutting off the sunlight with a screen caused the instant disappearance of the bright lines. Repeating the experiment it was found that the bright lines came into view on the sides of the sodium lines towards the blue, that is to say, it is light for which the medium has an abnormally low refractive index that is bent around the edge of the plate and enters the instrument. This is precisely what we should expect, for sodium vapor has a refractive index of less than 1 for waves slightly shorter than D_1 and D_2. The rays then will be concave upward in a medium in which the refractive index varies, as in the present case. If the sodium vapor is very dense we see only a single bright line bordering D_2, owing to the complete absorption of the light between the lines.

CHAPTER V

THE ORIGIN OF SPECTRA

Older Theories. — The modern treatment of matters to be considered in subsequent chapters such as the Zeeman and Stark effects requires a knowledge of the quantum theory of radiation and the origin of spectra. A complete treatment of atomic structure, and its relation to spectrum lines would be out of place in this book and the brief account here given is intended only to serve as an introduction to the subject.

The classical theory of a generation ago referred the radiation complex, which the spectroscope resolves into a series of lines more or less regularly spaced, to vibrating electrons, each one remaining fixed in position within the atom, unless disturbed by some outside force, under the influence of which it executed vibrations about its equilibrium position, and radiated light of the same frequency as its own. No satisfactory model of such an atom could be conceived, as the simultaneous presence of an attractive and repulsive force was required to account for an equilibrium position of the electron.

Then came the atom of J. J. Thomson in which the electrons were imagined as imbedded in a uniform sphere of positive electricity in which he showed that they would take up stable arrangements in concentric rings, the number of rings formed increasing with the atomic weight and total number of electrons present.

On adding electrons one by one, theory showed that they arranged themselves in a ring up to the number of five. On adding a sixth, one goes to the centre, when the ring again becomes stable until it contains eight, after which others go to the centre to form an inner ring. The effect can be well shown by steel bicycle balls floating in a dish of mercury over the pole of a magnet, the author's modification of Meyer's experiment with magnetized needles supported in water by small corks. This model suggested the periodic table of the elements, and other chemical properties but was incapable of representing spectroscopic results.

Rutherford's atom came next, and as interpreted by Bohr has formed the basis of the modern theory of spectra. He imagined that the effective mass of the atom consisted of a positively charged

nucleus, very small in comparison to the actual size of the atom, surrounded by negative electrons in orbital motion, like planets revolving about a central sum. This atom he devised to account for his observations on the scattering of alpha particles from radioactive substances. On the classical theory, however, an electron rotating in an orbit should emit radiation of a frequency equal to its own, which at first sight might appear to account for spectral lines, but a little further consideration showed that the radiation of energy would cause the orbit to contract, and the frequency to change, a condition that would give rise to a continuous spectrum instead of one made up of narrow lines.

Moreover, atoms radiate only when excited electrically or by some other means, no light being given off by atoms in the normal state. And we are therefore forced to define the electronic structure of a non-radiating atom as distinguished from that of one which is emitting light.

The simplest of all spectra is that of hydrogen in the atomic modification consisting of a number of series of lines, each series converging upon a limit and spaced at intervals which become uniformly smaller as we proceed from the first line to the "head" or limit of the series. One of these series constitutes what is commonly known as the "line spectrum" or Balmer series of hydrogen, the first four lines being in the visible region and the remainder in the ultra-violet. The first line, H_α, in the red, the second $H\beta$ in the blue-green and the third H_γ in the violet, can be seen in the usual laboratory vacuum tube, while $H\delta$ is visible in the powerful tubes employed by the author in extending the series from the ten lines previously observed in the laboratory, to twenty-two; thirty-three have been observed in certain stellar spectra.

Production of the Hydrogen Spectra. — Prior to 1920 only ten lines of the Balmer series had been observed in the laboratory, while 33 had been identified in the spectrum of the sun's chromosphere, and certain stars. Bohr's explanation of the absence of the higher members in the spectra of hydrogen-vacuum tubes was that it resulted from lack of room for the larger orbits, the atomic diameter for the higher excited states being larger than the mean distance between the atoms. Extremely low pressures appeared to be the necessary condition, and this might obtain in stellar upper atmospheres. An attempt to put this theory to experimental test by the author [1] led to the discovery of the proper condition for extending the series, though, as it turned out, this condition was not ultra-low pressure. An end-on tube one metre

[1] R. W. Wood, "An Extension of the Balmer Series," *Proc. Roy. Soc.*, *97*, 1920.

in length with a bore of 7 mms. was placed at a distance of 3 metres from the spectrograph and an image of the long column of luminous gas focussed on the slit by a lens of 20-cm. focus. It was found, however, that at the lowest pressures at which the discharge could be maintained, fewer lines could be seen than with pressures considerably higher. The best pressure was such as to give a Crookes dark space of about 2 mms. at the negative electrode. The discovery was made that the central part of the tube gave practically a pure Balmer spectrum, uncontaminated by the secondary spectrum and continuous background, and that this was the condition necessary for an extension of the series, as the higher members are extremely faint. The color of the discharge was fiery-purple at the centre of the tube and bluish-white near the electrode bulbs. The tube was accordingly bent at a right angle in two places so that only the central part was used, with the result that 22 lines of the series were photographed. Moist hydrogen from an electrolytic generator flowed continuously through a long and very fine capillary into one end of the tube, and was pumped off at the other end. This kept the spectrum free from impurities. It was operated with a transformer giving 20,000 volts and a current of ¼ ampere. Similar tubes immersed in a water-bath can carry a full ampere of current. With a tube of this description Brackett subsequently found a new series of lines in the infra-red and Pfund observed a single line of still another series. The successful operation of these tubes depends upon the circumstance that the hydrogen in the central portion is in the atomic condition, the return to the molecular state near the ends resulting from the catalyzing action of the electrodes. It was also found that with dry hydrogen admitted through a palladium tube, it was possible to secure the secondary spectrum wholly free from the Balmer series, after operating the tube for several hours with repeated washing out with fresh dry hydrogen. The original paper should be consulted for further details.

A by-product of the investigation [1] was the discovery that the hydrogen remained in the atomic condition after passing out of the region of the discharge and that a fine wire of tungsten mounted n the tube leading to the pump at a distance of 10 cms. from the discharge was maintained at a white heat by the comparatively cold atomic gas, the atoms combining to form molecular hydrogen at the surface of the metal and giving up their heat of combination. This observation led to the invention of the atomic hydrogen welding torch by Langmuir.

[1] R. W. Wood, "Spontaneous Incandescence of Substances in Atomic Hydrogen," *Proc. Roy. Soc.. 102*, 1, 1922.

Early Studies of Spectral Series. — Balmer in 1885 discovered the law governing the spacing of the lines in the hydrogen spectrum of which nine had been obtained in the laboratory, while five more were photographed in the spectrum of Sirius by Huggins. The series was represented by the formula

$$\lambda = 3645.6 \frac{m^2}{m^2-4}$$

in which m takes the successive values, 3, 4, 5, $\cdots$, etc.

The constant term 3645.6 was found from the first four lines, and represents the "head" or point of convergence of the series, as can be seen if we put $m = \infty$ in the formula.

Studies of other spectra by Kayser and Runge, Rydberg and others showed that the lines could be sorted out into overlapping series, similar in appearance to the Balmer series, and like it converging to heads on the short wave-length side. Some of the series had a common point of convergence. The lines of the series may be single, as in the case of helium, double as in the case of the alkali metals or, as in the case of mercury for example, there may be series of single lines, and also series of triplets, the latter in this case being so widely separated as to be unrecognizable by inspection (mercury 5461, 4358 and 4046 for example).

Balmer's formula expressed in wave-numbers $V = 1/\lambda =$ number of waves per cm. takes the form $V = R[(m^2-4)/4m^2]$. The constant R is named for Rydberg, since he showed that it occurred, not only in the formula for the hydrogen spectrum but in the formulae for all spectral series. Lyman discovered a hydrogen series in the ultra-violet, working with a vacuum spectrograph (for these radiations are absorbed by air) and Paschen found another series in the infra-red.

All three series were represented by the Rydberg formula

$$V = R\left(\frac{1}{m^2} - \frac{1}{n^2}\right)$$

in which $R = 109{,}678 cm^{-1}$ and m and n are integers.

For the three series we have:

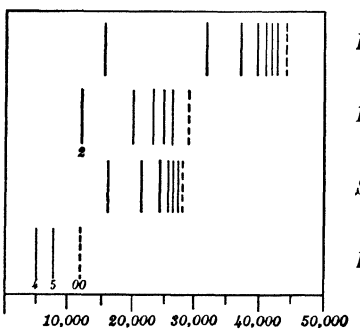
Lyman series	$m=1$	$n=2, 3, 4, \cdots$
Balmer series	$m=2$	$n=3, 4, 5, \cdots$
Paschen series	$m=3$	$n=4, 5, 6, \cdots$

and the more recently discovered Brackett series $m=4$ $n=5, 6, 7, \cdots$ This means that each line of the entire spectrum of hydrogen can be represented as the *difference between two terms of the form* R/m^2. If we assign the proper integer to m, R/m^2 repre-

sents the wave-number of the head of each series. A physical significance has been given to these terms by Niels Bohr, whose theory will be given presently.

For other elements than hydrogen the work of Kayser and Runge and of Rydberg showed that there were, in each case, a large number of series, generally more or less superposed, each series resembling a hydrogen series when sorted out from its overlapping neighbors.

Four chief series have been recognized, the Principal, Diffuse, Sharp and Fundamental (or Bergman) designated by the letters P, D, S and F. These four series for lithium are shown in Fig. 87.



All of the lines shown are close doublets, or, as in the case of the diffuse and Bergman series composite doublets which appear as triplets. Similar series are shown by sodium, the first doublet of the principal series being the yellow sodium lines D_1, D_2. This series is the most complete ever observed. The 8 members previously recognized in the emission spectrum, were

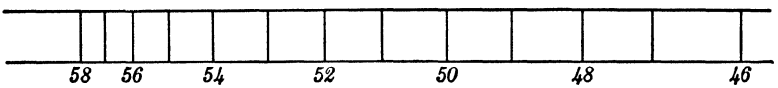


FIG. 87

raised to 48 by the author [1] (1909) by photographing the absorption spectrum of the vapor of the metal in a long steel tube, and the number still further increased to 58 by Wood and Fortrat [2] (1913). The last line observed is only 1.2 Ångström units from the theoretical head. A diagram of the last twelve lines is shown at the bottom of Fig. 87, the head lying to the left at a distance from the 58th line equal to the distance between the 46th and 58th lines. If we continue plotting the series on the same scale, the *D* lines will lie on our right at a distance of 348 metres or more than a thousand feet! The last 32 members, or more than half of the series, cover a spectral range no wider than the distance between the *D* lines.

We will now see how the Rydberg formula fits the case of these spectra. Referring to Fig. 87 we observe that the diffuse and

[1] Wood, *Phil. Mag.*, 1909; *Astrophysical Journal*, *29*, 97, 1909.

[2] Wood and Fortrat, *Astrophysical Journal*, *43*, 73, 1916 (pub. delayed by War).

sharp series converge to the same head. The formula for hydrogen requires modification in the case of the other elements. As given by Rydberg, it is of the form:

$$V_m = A - \frac{R}{(m+\mu)^2} \qquad (1)$$

in which A is the head or limit of the series, R the Rydberg constant for hydrogen and V_m the wave-numbers of the lines which are found by taking successive integral values of m. μ is usually a decimal part of m, though sometimes larger than unity.

Each series is represented by a limit, minus a series of variable parts or "terms." The four series are given by

Principal	$P_m = P_\infty - R/(m+p)^2$
Sharp	$S_m = S_\infty - R/(m+s)^2$
Diffuse	$D_m = D_\infty - R/(m+d)^2$
Fundamental	$F_m = F_\infty - R/(m+f)^2$.

In these expressions P_m, for example, means the mth line of the Principal series, and p the value of μ in (1).

This may be illustrated by taking the case of lithium regarding the very close doublets as single lines

$$P_m = 43{,}487 - 109{,}721/(m+0.959)^2$$
$$S_m = 28{,}601 - 109{,}721/(m+0.595)^2$$
$$D_m = 28{,}598 - 109{,}721/(m+0.997)^2.$$

If we compare this formula with the one for hydrogen we see that for large values of m, *i.e.* for lines near the head of the series, the two are practically identical, since the correction terms p, s, d and f, usually less than unity, have less and less significance, as the integral values of m increase. We shall see presently the physical cause of these correction terms, and why they have different values for the different series.

The important thing to bear in mind is that the frequency or wave-number of each spectrum line of a series is represented as the difference between two terms, one fixed, the other variable, to which a real physical significance was first given by Niels Bohr in 1913.[1]

Bohr's Theory of Atomic Radiation. — Bohr made the bold hypothesis (contrary to the supposedly established laws of electrodynamics) that an electron in orbital motion is non-radiating, and that the terms corresponded to the energy of the atom with its electron in orbits of different diameters. He considered that radiation occurred only when the electron jumped from an orbit of large, to one of smaller diameter, the frequency of the radiation depending on the difference between the energies in the two orbits.

On these assumptions he developed an expression for the fre-

[1] Bohr, *Phil. Mag.*, *26*, 1476, 1913.

quencies of all lines in the hydrogen spectrum, and found that it was identical in form with Balmer's empirical formula, the Rydberg constant being expressed in terms of the mass and charge of the electron, Planck's constant h and c the velocity of light.

We shall now develop expressions for the diameters of the electronic orbits in the allowed states, and the atomic energies in these states, and show that the expression for the difference in energies of the two states concerned in a transition is the equivalent of the right-hand member of the Rydberg equation, the constant appearing as a term composed of natural constants which have all been measured.

Consider first the case of an electron of charge $-e$ moving in a circular orbit of radius r about a nucleus which has a positive charge Ze (Z being the atomic number, or 1 for hydrogen), the attraction for the electron following Coulomb's law, *i.e.* varying inversely with r^2. The energy of the electron is in part kinetic, $\frac{1}{2}mv^2$, if m is its mass, and in part potential, depending on its distance from the nucleus. The total energy is the quantity with which we are concerned. Now it is impossible to give absolute values to the potential energy. Relative values only can be determined, *i.e.* we must fix on some position of the electron at which the value zero is arbitrarily assigned to the potential energy. It is customary to take infinity as the zero point, *i.e.* to consider the ionized atom as having zero potential energy. This will really be the position of maximum potential energy, and the energies at finite distances become negative quantities on this convention.

The potential energy at a distance r from the nucleus is then defined as the work required to bring the electron from infinity, where the potential energy is, by definition zero to the distance r. This amount of work is a negative quantity. The attractive force is Ze^2/r^2, the integral of which from r to infinity is Ze^2/r, consequently we have for the potential energy $-Ze^2/r$.

It is important to note that as r increases, the value of the potential energy increases, *i.e.* it is greater for a larger orbit. The total energy of the electron is therefore

$$E=\frac{mv^2}{2}-\frac{Ze^2}{r}. \tag{1}$$

This introduction of negative potential energies does not affect the final result which involves only energy differences. If the electron runs in a stationary circular orbit the centrifugal force must balance the attraction, therefore we have

$$\frac{mv^2}{r}=\frac{Ze^2}{r^2} \text{ which gives } \frac{Ze^2}{r}=mv^2. \tag{2}$$

This expression makes it possible to express the total energy E in terms either of the orbital velocity v of the electron or of the orbit's radius r

$$E=\frac{mv^2}{2}-\frac{Ze^2}{r}=-\frac{mv^2}{2}=-\frac{Ze^2}{2r} \tag{3}$$

which shows us that the kinetic energy is always one-half of the potential energy $-Ze^2/r$. The total energy is always negative, on the convention adopted, that is when referred to the ionized atom. Decreasing the orbit's radius r increases the kinetic energy but decreases the *total* energy as shown by either of the above expressions on account of the negative sign.

Bohr postulated further that only orbits of certain definite size are permissible, for which the angular momentum of the electron is $nh/2\pi$, n being any integral number and h Planck's constant defined in the Chapter on Radiation. These orbits he termed "Stationary States." The frequency of the emitted radiation is determined by the difference between the energies of the electron in its two orbits (or the two stationary states) as expressed by

$$h\nu=E_n-E_{n'}$$

or the energy difference is equal to the product of the frequency and Planck's constant.

Monochromatic radiation occurs when the electron passes from an outer to an inner orbit, and at no other time.

Equating angular momentum to $nh/2\pi$ gives

$$2\pi mrv=nh$$

or $r^2=\dfrac{n^2h^2}{4\pi^2m^2v^2}$ by substitution of which for r^2 in Eq. (2) gives

$$r=n^2\frac{h^2}{4\pi^2mZe^2} \text{ while } v=\frac{2\pi Ze^2}{nh} \tag{4}$$

which shows that the diameters of the allowed orbits are as the squares of n, the total quantum number, *i.e.* as 1, 4, 9, 16, etc., that of the outer orbit concerned in the production of the 22nd line being 484 times that of the innermost. This was the circumstance that lead Bohr to the idea that very low gas pressure was necessary for the appearance of the higher members.

By combining the expressions for r and v we can calculate the frequency of the electron in its orbit, which is found to be inversely as the cube of n

$$\omega=\frac{4\pi^2mZ^2e^4}{n^3h^3}.$$

For the above-mentioned orbit this is 1/10648 of the inner orbit frequency. We next calculate the energies of the various orbits thus: Calling W the total energy of the atom we have from (3)

$$W = -\frac{mv^2}{2} = -\frac{Ze^2}{2r} = -\frac{2\pi^2 mZ^2e^4}{h^2} \cdot \frac{1}{n^2}. \tag{5}$$

The absolute values of W are inversely as the square of n, which means a diminishing value of the negative energy with increasing orbital diameter (or an increase of the energy if considered as a positive quantity). This means that the larger orbits have the greater energy values.

Writing ΔW for the change of energy during the orbital transfer, we have for the frequency emitted $V_T = \Delta W/h$ or wave-number $\nu = \Delta W/hc$, c being the velocity of light and the wave-number the reciprocal of the wave-length or the number of waves per centimetre. For hydrogen $Z = 1$, and therefore falls out of (5), and the wave-number emitted is represented by

$$\nu = \frac{2\pi^2 me^4}{ch^3}\left(\frac{1}{n''^2} - \frac{1}{n'^2}\right) \tag{6}$$

in which n' and n'' are the total quantum numbers of the outer and inner orbits, the change of sign resulting from the circumstance that we are dealing with negative energies. This is the equivalent of the Balmer formula and shows us that the Rydberg constant is

$$R = \frac{2\pi^2 me^4}{ch^3} \tag{7}$$

a quantity made up of already determined constants and equal to $109{,}800 cm^{-1}$. The value determined from the hydrogen spectrum is 109,678 and this close agreement (within the limits of experimental errors in the determination of the values of m, e and h) constituted the first great triumph for the Bohr theory. The transitions involved in the emission of the lines of the various series of the hydrogen spectrum are shown in Fig. 88, the orbits having their correct relative size. At the right of the figure the same transitions are shown on a diagram of energy levels (not drawn to scale, however), expressed for convenience in wave-numbers since we are concerned chiefly with frequencies in the case of spectral lines. For a diagram with the levels given to scale see Fig. 89, next section. It should be noted that while the distance between adjacent allowable orbits increases as we go out from the nucleus the energy levels crowd together. The lowest term, corresponding to the smallest orbit, is the Rydberg constant for hydrogen, $R = 109{,}800$, the successive terms above it being R/n^2, or $R/4$, $R/9$ . . . , etc. The levels, here defined in wave-numbers, can be

given values expressed in energies, by multiplying by hc. The lowest level has the value 2.15×10^{-11} ergs, which means that 2.15×10^{-11} ergs are required to ionize the atom with its electron in the smallest orbit. The values for the next two orbits are .539 and .239.

In the unexcited state the hydrogen atoms have their single electrons in the innermost orbit. By electrical excitation the electrons are thrown out to the outer orbits, from which they

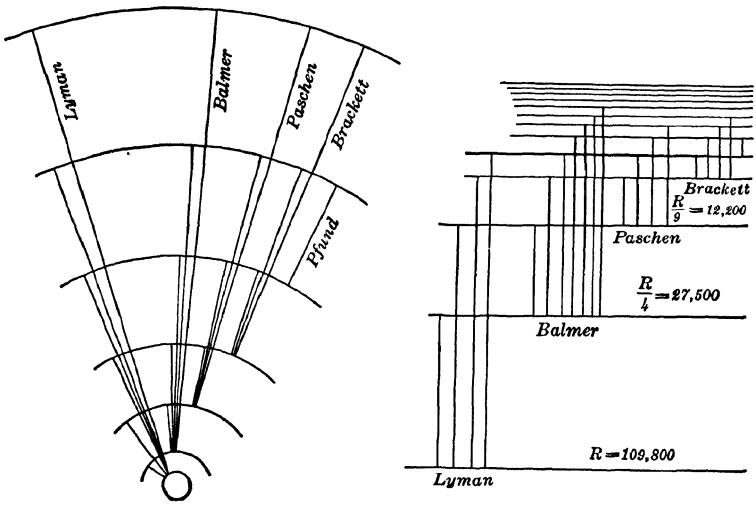


FIG. 88

return to lower orbits and eventually to the lowest. Those returning to orbit $n=1$, give the lines of the Lyman series in the remote ultra-violet, while the ones which land in orbit $n=2$, are responsible for the visible or Balmer series. The infra-red series discovered by Paschen and by Brackett are associated with returns to orbits $n=3$ and $n=4$. One line of the next series for $n=5$ was observed by Pfund, its wave-length being 7.4 μ. It is to be noted that an atom can radiate only a single frequency at any one time. If we could observe the spectrum of one atom under continuous repeated excitation, we should find that only one line at a time appeared in the spectrum, its position changing at random. This is a great advance over the old idea that each atom radiated simultaneously the hundreds or thousands of frequencies which made up its complete spectrum, a phenomenon that would require an atom of enormous complexity.

The Spectrum of Ionized Helium and Doubly Ionized Lithium. — Bohr found a further remarkable verification of his theory in the spectrum of ionized helium.

The normal helium atom consists of a nucleus with a mass four times that of the hydrogen nucleus and of charge +2, with two electrons in orbital motion. If the atom is ionized, which may occur with heavy electrical discharges, one of these electrons is lost and the atom becomes similar to the hydrogen atom except for the double charge, and quadruple mass of its nucleus.

Long before the advent of the Bohr theory, Pickering had observed a series of absorption lines in the spectrum of the star ζ Puppis which appeared to converge to the same limit as the Balmer lines of hydrogen, a relation characteristic of the "sharp" and "diffuse" series of other elements, as we have seen. This suggested they were due to hydrogen in some modification not obtainable in the laboratory. Rydberg calculated a principal series for hydrogen on this assumption, and found values $\lambda=4688$ and 2734 for the first two lines. The former agreed closely with a strong line observed in ζ Puppis and certain nebulae and general acceptance was given to the theory of a special form of "cosmic" hydrogen. The other members could not be verified, as the earth's atmosphere cuts off all wave-lengths below 2900. These lines were finally found in the laboratory by Fowler in vacuum-tubes containing a mixture of helium and hydrogen. In addition to the series predicted by Rydberg, he found an intermediate series of which the first line was $\lambda=3203$, as well as three lines of the Pickering series. All of these lines were subsequently shown by Bohr to be due to ionized helium. The first line of the Pickering series is at 10,123 in the infra-red, the 2nd, 4th, 6th and 8th lines are almost in exact coincidence with the hydrogen lines H_α, $H\beta$, H_γ, etc., while the 3rd, 5th, 7th, etc., are midway between them.

If now we make $Z=2$ in place of $Z=1$ in equation (5), equation (6) becomes

$$\nu=\frac{8\pi^2me^4}{ch^3}\left(\frac{1}{n'^2}-\frac{1}{n^2}\right)=4R\left(\frac{1}{n'^2}-\frac{1}{n^2}\right)=\frac{R}{\left(\frac{n'}{2}\right)^2}-\frac{R}{\left(\frac{n}{2}\right)^2}.$$

The series are as follows:

	Observer
$\nu=4R\left(\frac{1}{1^2}-\frac{1}{n^2}\right)$ in which $n=2, 3, 4, \cdots$	Lyman
$\nu=4R\left(\frac{1}{2^2}-\frac{1}{n^2}\right)$ in which $n=3, 4, 5, \cdots$	Lyman
$\nu=4R\left(\frac{1}{3^2}-\frac{1}{n^2}\right)$ in which $n=4, 5, 6, \cdots$	Fowler
$\nu=4R\left(\frac{1}{4^2}-\frac{1}{n^2}\right)$ in which $n=5, 6, 7, \cdots$	Pickering

For even values of n'' and n' the helium lines are practically in coincidence with hydrogen lines, while odd values give lines which occupy intermediate positions.

If we put R in place of $4R$ in the above expressions we have the Lyman, Balmer, Paschen and Brackett series respectively. Now R for hydrogen was $109{,}800 cm^{-1}$; for ionized helium $4R$ is therefore 439,200, and in a similar way the other hydrogen energy levels may be transformed into ionized helium levels by multiplying their values by 4. This is nearly, but not quite exact, for the Rydberg frequency R is slightly larger for helium than for hydrogen, owing to the greater mass of the nucleus. This factor comes in if we consider the motion of the nucleus (neglected in the treatment of hydrogen) for in reality nucleus and electron rotate about their common centre of gravity, and the heavier helium nucleus will be more nearly stationary. This motion of the nucleus is taken into account in the more complete treatments given in works on atomic structure and spectral theory. The energy levels for hydrogen and ionized helium are shown on the same scale in Fig. 89, the wave-number values at the left expressed in thousands. The lower level for helium should be very much further down than shown in the diagram. The energy necessary to remove the second electron is four times that which is required for the ionization of hydrogen.

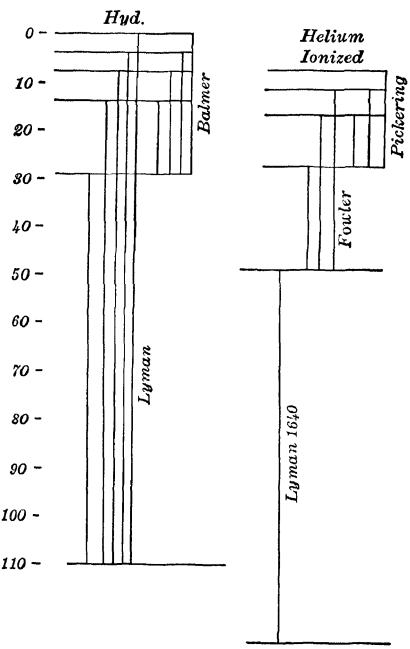


FIG. 89

The lithium atom is made up of a nucleus $Z=3$ with two electrons revolving close to it and a third at a somewhat greater distance. The two inner electrons form the K ring of X-ray spectroscopy, while the outer electron is the one displaced to higher orbits in the formation of the lithium spectrum, which as we shall show later on consists of series of double lines. If this outer electron is driven off we have an atom resembling helium, except for the greater mass and charge of the nucleus, and the spectrum of singly

ionized lithium resembles that of normal helium. Doubly ionized lithium we should expect to give a spectrum resembling that of hydrogen and this has recently been observed.

Elliptical Orbits and Two Quantum Numbers. — Bohr, in his earlier treatments, had considered the case of elliptical orbits and showed that the energy was inversely proportional to the major axis, being the same as for a circular orbit of the same diameter. He considered ellipses of all degrees of eccentricity to be possible, but was obliged to introduce a second quantum number for cases in which the attractive force exerted by the nucleus on the electron was not strictly proportional to the inverse square of the distance. Such a departure from a Coulomb field will cause the orbit to precess in its own plane, and a second periodicity is thus introduced, which must also be quantized. As we shall see, this is the case for the great majority of atoms, for the inner shells of electrons partially screen the outer electron from the full attractive force of the nucleus, but if the orbit penetrates this shell, the screening action is more or less completely removed and the force increases more rapidly than by the inverse square law.

Now observations with spectroscopes of high resolving-power had shown that the lines of hydrogen and ionized helium were in reality double, the separation of the components of H_α being about 0.3 of an Ångström unit or one-twentieth of the distance between the D lines of sodium.

Sommerfeld developed a theory explaining this phenomenon by considering the change of mass of the electron which accompanied its change of velocity as it traversed its elliptical orbit. This periodic change of mass, called for by the theory of relativity, will be greater for orbits of large eccentricity, since the difference between aphelion and perihelion velocity increases with the eccentricity.

This change of mass is accompanied by a change in the energy of the orbit. The electron in an elliptical orbit has two degrees of freedom, r the radius and θ the angular coördinate, and Sommerfeld quantized both, calling n_a and n_r the azimuthal and radial quantum numbers, the sum of which equals the total quantum number n. In place of a single circular orbit for each value of n as in the preceding treatment, we now have n orbits of varying eccentricity as shown by Fig. 90. The total quantum numbers are shown to the left and the azimuthal numbers below. For hydrogen and hydrogen-like atoms the ellipses for a given n lie practically on a horizontal line, which line represents the energy level. These should, of course, be spaced as were the levels in Fig. 88. They are represented here as crowded together to reduce the size of the fig-

ure. We here get our first notion of the physical significance of the *S*, *P*, *D* and *F*, terms which figured in empirical spectroscopy before the advent of the Bohr theory. They were not employed in the case of hydrogen (as has been done here) for reasons which will be immediately apparent. It had been found that the wave-numbers of the lines could be expressed as differences between certain "terms" (corresponding to the modern energy levels). Taking the differences between *all* the terms in pairs showed that

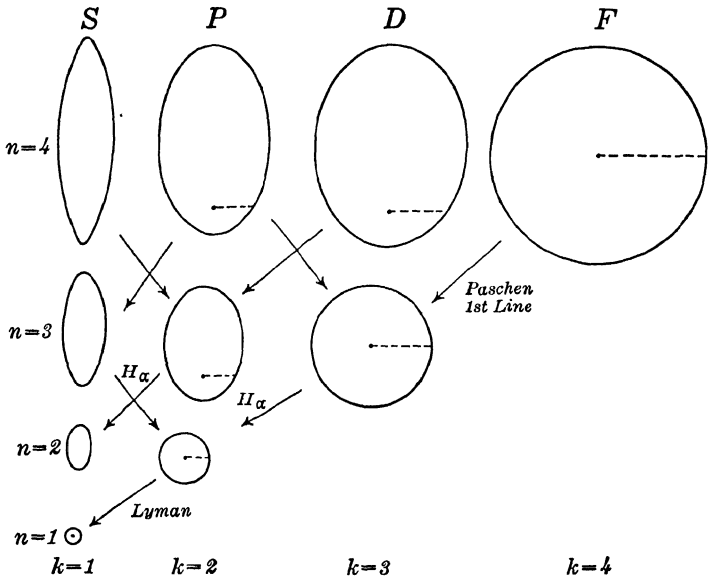


FIG. 90

only certain pair differences gave spectrum lines. It was thus possible to sort out the terms into families, the members being arranged in parallel vertical columns, spectrum lines being as a rule given by the difference between terms in adjoining columns but by no others. For example we might consider Fig. 88 as illustrating this principle, if we substitute "terms," expressed in wave-numbers, for the orbits. But it will be better to take the sodium spectrum diagram, in which the orbits do not lie along horizontal lines as in Fig. 88 which is for hydrogen. The spectrum of sodium is made up of a number of overlapping series of close double lines, which for the present we shall consider single (Fig. 91). The terms were expressed in wave-numbers; and those in the second column designated by the letter *P*, since their differences taken against the lowest *S* term gave the wave-numbers of the lines of the

"principal" series (the first member being the D lines in the yellow). Differences between the lowest P term and the S and D terms gave the "sharp" and diffuse series respectively, while differences between the F terms and lowest D gave the fundamental series, as it was formerly called, as the lines were of low frequency. (It is now called the Bergman series.) Other combinations of terms do not, in general, give observable lines, *i.e.* transitions do not take place between upper D and lower S levels. Bohr's selection principle expresses this by stating that transitions are restricted to those in which the azimuthal quantum number changes by ± 1.

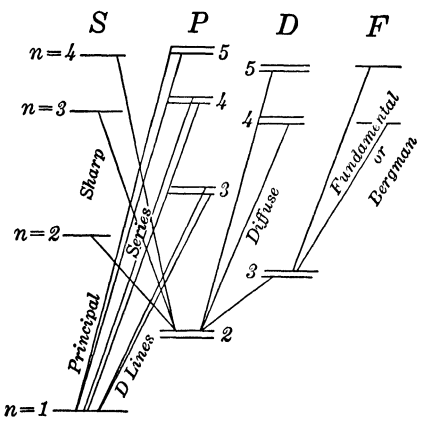


FIG. 91

As a matter of fact such transitions do occasionally take place, especially in strong electric fields. Datta found absorption lines of potassium corresponding to the transition from the lowest S level to the two lowest D levels. Here there was no possibility of an electric field.

Now it is clear from Fig. 90 that each line of the Balmer series can be emitted as a result of the three allowable transitions shown by the arrows for H_α, while the Paschen series lines may result in five different ways as shown by dotted arrows for the first line. In other words, H_α could be regarded as three superposed lines, due to three different transitions, on Bohr's original idea that the energy of the elliptical orbit was dependent on its major axis alone. But by Sommerfeld's introduction of the change of mass of the electron as a result of the high perihelion velocity, the ellipses and the circle corresponding to any given value of n, do not lie exactly on a horizontal line, *i.e.* there is a slight difference in the height of the energy levels. Recent work has shown that electron spin must be considered as will be shown presently.

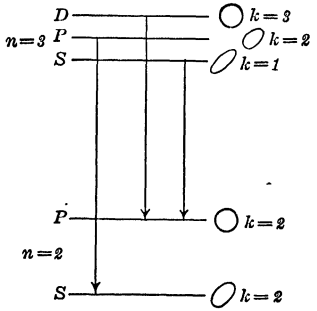


FIG. 92

The lower level for H_α, $n=2$ splits into two representing the

energies of the elliptical and the circular orbits respectively. The upper level, $n=3$ splits into three, which are much closer together than the two lower levels, as shown in Fig. 92, the separation of the levels being enormously magnified. On this scale of separation the $n=2$ levels should lie 300 metres below those for $n=3$. The transitions are indicated by arrows, and it is clear that we have three lines in all, two of which have very nearly the same wave-length, the double hydrogen line resulting from the energy difference between the two lower levels. The failure to find the third line formed a serious objection to Sommerfeld's theory, but the more recent theory involving electron spin calls for the third line between the two main components, instead of to one side, and recent work by Hansen and by Kent has given evidence that such a line exists, though the three have not been resolved, and the presence of the third line is only inferred from the asymmetrical form of the intensity curve of the H_α line. The theory of the electron spin will be given in the Chapter on the Zeeman Effect. The lines of the Lyman series, formed by transitions from the P levels to the lower S level are single as only one transition is possible for each line.

Sommerfeld also calculated the expected fine structure of the lines of ionized helium, and his predictions were verified in a remarkable manner by Paschen's photographs of the 4686 and 3203 lines of the Fowler series which, as we have seen, correspond to the Paschen series of hydrogen, the former having five components.

Spectra of Alkali Metals. — Next to hydrogen in point of the simplicity of their spectra come the alkali metals, lithium, sodium, potassium, rubidium and caesium.

As we have seen the spectra of these metals consist of more or less superposed series, each one of which resembles, at first sight, one of the hydrogen series, but the spacing of the lines follows a different law. The higher terms in fact give lines spaced very nearly according to the law governing the spacing of the hydrogen lines, the departure becoming more pronounced as we proceed toward the lines of longer wave-length. This is shown by the energy diagram (Fig. 93), on which the upper levels are practically in coincidence with the levels for hydrogen. The Rydberg equations for the alkali metals are:

Principal Series $\nu = A - R/(m+p)^2$
Sharp Series $\nu = B - R/(m+d)^2$
Diffuse Series $\nu = B - R/(m+s)^2$
Bergman Series $\nu = C - R/(m+1)^2$.

These differ from the hydrogen formula by the presence of Rydberg correction terms p, d and s.

In these equations A and B and C are constants, R is the Rydberg constant for hydrogen, m takes successive integral values, and

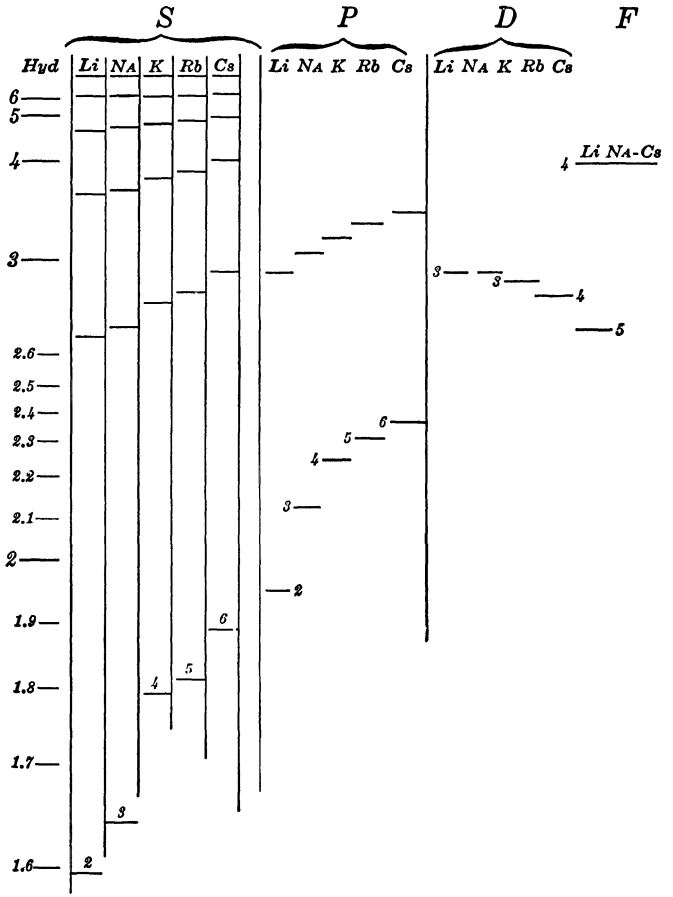


FIG. 93

p, d and s have values generally less than one, and are attached to m as decimals.

The constants A, B and C can be expressed in terms of p, d and s.

$$A=\frac{R}{(1+s)^2} \qquad B=\frac{R}{(2+p)^2} \qquad C=\frac{R}{(3+d)^2}.$$

The S, P, D, F energy levels of these elements are shown in Fig. 93. It is to be noted in the first place that the lowest or ground level which represents the energy of the unexcited atom, as it exists in the case of the vapor formed by heating the metal in vacuo to

a moderate temperature, has not the total quantum number $m=1$, but numbers which increase progressively as we pass along the series, $n=2$ for lithium, 3 for sodium, 4 for potassium, etc., in both the S and P sequences, while in the D sequence the lowest terms have values, 3, 3, 3, 4 and 5 in the F sequence all lower levels have $n=4$. The positions of these energy levels have been determined from spectroscopic data, and the levels for hydrogen are given on the same scale at the left of the figure.

The assignment of these higher total quantum numbers to the ground orbits, or lowest energy levels of the S, P, D, F sequences comes about in the following way.

The lithium atom consists of a nucleus with a charge of 3, and three electrons. Spectroscopic data show that two of these occupy $1S$ circular orbits, corresponding to the $1S$ orbit of hydrogen, while the third occupies the next larger or $2S$ orbit, which, as we see from Fig. 90 is elliptical. In the case of sodium we have two electrons in $1S$ orbits, two in the $2S$ orbits, 6 in $2P$ orbits (circular) and one, the valence electron, in the $3S$ orbit, so that the lowest energy level for sodium has a total quantum number of 3. The orbital distribution of electrons for the five alkali atoms is given in the following table:

	K	L		M			N								
	1S	*2S*	*2P*	*3S*	*3P*	*3D*	*4S*	*4P*	*4D*	*4F*	*5S*	*5P*	*5D*	*5F*	*6S*
Li	2	1				X									
Na	2	2	6	1		X				0					
K	2	2	6	2	6	X	1			0					
Rb	2	2	6	2	6	10	2	6	X	0	1				
Cs	2	2	6	2	6	10	2	6	10	0	2	6	X		1

The K, L, M, N rings of X-ray spectroscopy are given above. The orbital structure for sodium is shown in Fig. 94 and by comparing this with Fig. 90 and the above table, the meaning of the table will be clear and the reader will have no difficulty in constructing similar diagrams for the heavier atoms.

The table shows clearly why the lowest energy levels of Fig. 93 for the S and P sequences have total quantum numbers increasing from 2 to 6 as we pass from lithium to caesium. Spectra arise from transitions of the valence electron which occupies, in the normal or unexcited atom, an S orbit of one quantum number higher than that of the highest S orbit containing its full quota of two electrons. The $2P$ level, or orbit, for the valence electron of lithium is the lowest, while in the case of sodium the $2P$ orbit has its full quota of six electrons, and the valence electron, in falling back from a higher orbit cannot land here, the next unoccupied P orbit having quantum number 3.

That the lowest D orbits have quantum numbers 3, 3, 3, 4 and 5 is shown also by the table in which they are designated by an X, while the lowest F orbits are all of value 4, designated by 0 in the table. In view of what has been considered thus far, we might infer that the lowest level for lithium would coincide with the $2S$ level for hydrogen, but our diagram (Fig. 93) shows us that this is not the case, the lithium levels being well below the corresponding ones of hydrogen. This depression comes about as follows.

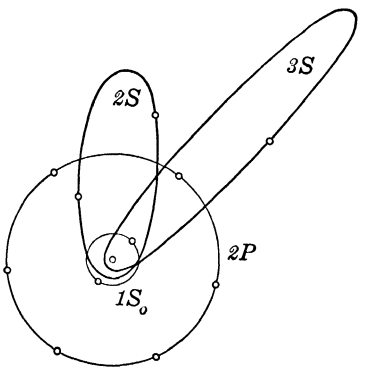


FIG. 94

Penetrating Orbits. — The explanation of the peculiarities in the spacing of these energy levels rests upon theories developed by Sommerfeld, Bohr, Schroedinger and others. In brief we no longer have the orbital electron moving in a Coulomb field, since the nucleus is surrounded by one or more shells of electrons outside of which shell the single valence electron rotates in quantized orbits. The charge on the nucleus is equal to the atomic number of the element, 3 for lithium, 11 for sodium, 19 for potassium, etc. The outer electron, whose orbital transitions give rise to the spectrum lines is more or less completely shielded from the attraction of the heavily charged nucleus by the screening action of the intervening electrons. Taking sodium as an example we have a nucleus with a positive charge, eleven times that of the hydrogen nucleus, surrounded by ten electrons, two in an inner shell, and eight in the outer. This is termed the core of the atom. Rotating about this is the single valence electron subjected to the attractive force of the core which may be regarded as having a positive charge of unity, or $11-10$. If the orbit is elliptical and not too large, the electron penetrates the core during its perihelion passage, and having passed through the screening ring of electrons, comes under a more powerful attractive force. This will increase the potential energy in the same way that the increased nuclear charge of the ionized helium atom makes the energy of the ground orbit greater than in the case of hydrogen. This, as we saw depressed the energy level on our diagram and in the same way the penetration of the electron into the core, by increasing the potential energy, depresses the lowest levels of the alkali atoms, below the hydrogen levels of corresponding total quantum number. If the orbit of the valence electron is large, it may lie

wholly outside of the core, in which case it is approximately elliptical.

If smaller, or of great eccentricity, it may penetrate one or more of the electron shells in which case it is no longer elliptical but of the form shown in Fig. 95 in which the core is represented by the dotted circle. Outside of this circle the orbit is approximately elliptical, but the electron on penetrating the core experiences an increased pull by the nucleus and is deflected from the elliptical path which it would have followed in a Coulomb field (shown dotted), and making a closer approach to the nucleus, passes out of the core into the second elliptical orbit. We thus have an orbit precessing in its own plane. Whether or not an orbit is a penetrating one depends upon the size of the core, and the diameter and eccentricity of the orbit.

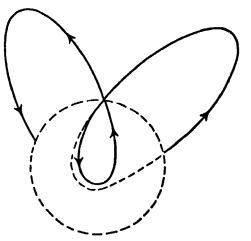


FIG. 95

In Fig. 96 are shown three orbits of say the S sequence, for which the azimuthal quantum number is $k=1$. Their total quantum numbers are 1, 2 and ∞, the latter a parabola of course, corresponding to the upper limit of the series of ellipses. All have the same parameter c which is also the radius of the circular orbit $n=1$, at the centre of which is the nucleus. It is clear that the electron's nearest approach to the nucleus is $c/2$ for the parabolic orbit. If now the radius of the core of the atom is greater than c all of the orbits penetrate it, the circular one lying wholly within it. If r is less than $c/2$ no orbit penetrates, while if r is intermediate between $c/2$ and c, orbits of high quantum number will penetrate while those of low will not. A core of such diameter is shown by the dotted circle.

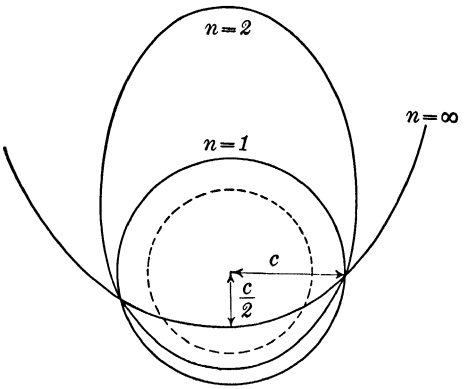


FIG. 96

It is clear that, in general, orbits are most likely to penetrate if they belong to the S series, and least likely if they belong to the F series, for the diameter of the circular orbit and major axis of the ellipses increases with increasing azimuthal quantum number

(see Fig. 90). Also they are more likely to penetrate as the size of the core increases, as in passing from lithium to caesium.

Referring back to Fig. 93 we see that, for lithium the $2S$ level is less depressed below the $2S$ level of hydrogen, than in the case of caesium with its lowest level $6S$ far below the $6S$ hydrogen level.

For the P series, with its larger orbits, $2P$ for lithium is almost in coincidence with $2P$ for hydrogen, while $6P$ for caesium is greatly depressed owing to its large core diameter. The D series terms are only slightly depressed and the F terms practically not at all. An example of a case in which orbits of higher total quantum number penetrate while the lower orbits do not, was found by Wentzel in the $1P$ sequence of terms for mercury. The mercury atom has two valence electrons, however, which complicates matters somewhat.

The Effective Quantum Number. — We have seen that for the hydrogen terms $\nu = R/n^2$, or $\sqrt{R/\nu} = n$ in which n is a whole number defined as the total quantum number. The deviations of a given term ν for the corresponding hydrogen term was taken care of by the Rydberg correction but it can also be expressed by the so-called "effective quantum number," designated by n^* defined by the equation

$$\nu = \frac{R}{n^{*2}} \qquad \text{or} \qquad n^* = \sqrt{\frac{R}{\nu}}.$$

In Fig. 93 the total quantum numbers for hydrogen, at the left have been subdivided into tenths, which are not equally spaced of course, as the distance between the integral quantum numbers decreases as we ascend. The effective quantum numbers for the various S, P, D terms of the alkali metals are given by noting where they fall on this scale, for example the $2S$ terms for lithium has $n^* = 1.59$ while for the $2P$ terms $n^* = 1.96$. In the former case the Rydberg correction $s = 0.6$ in the latter $p = -0.05$.

Absorption of Light on Bohr's Theory. — Bohr regarded absorption as the reverse process of emission, *i.e.* he attributed it to a transition from a lower orbit or level to a higher one. Atoms in the normal or unexcited state will absorb only radiations of wave-lengths corresponding to that of lines emitted as a result of transitions from the upper levels to the ground level of the atom. The energy necessary to carry the electron from the inner to the outer orbit is furnished by the absorbed light quantum, which must have exactly the right magnitude (*i.e.* frequency) otherwise no absorption occurs. The electron in some cases returns to its original orbit reëmitting the absorbed light. This is resonance radiation. Atomic hydrogen in the unexcited state is capable of

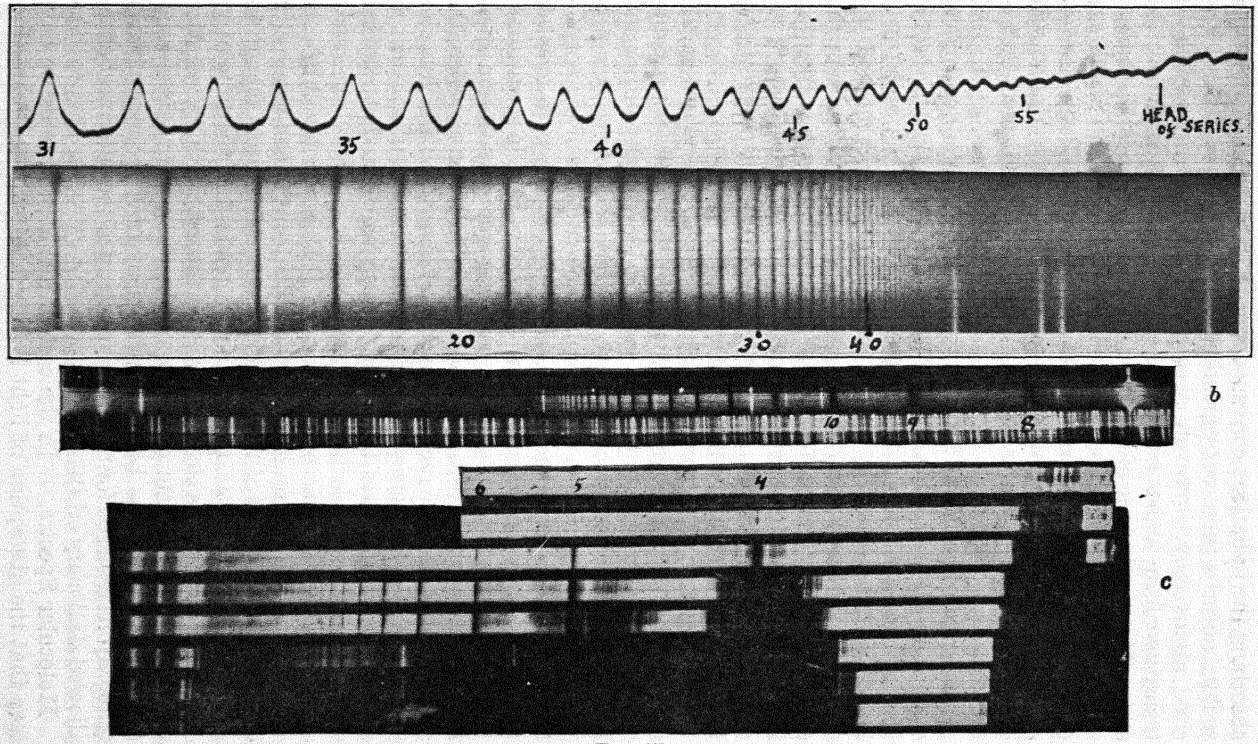


FIG. 97

absorbing the radiations corresponding to the Lyman series, while sodium absorbs those corresponding to the principal series. Excited atoms may, however, absorb radiations corresponding to transitions from higher levels to the level representing the excited state. Thus in the case of hydrogen excited by electrical discharges there are atoms with electrons on the $2S$ level and these are capable of absorbing radiations corresponding to the Balmer series.

Numerous interesting examples of this type of absorption will be given in the Chapter on Absorption and Resonance Radiation.

Absorption Spectra of Sodium. — We will consider one typical case of an absorption spectrum in its relation to Bohr's theory, that of the vapor of sodium obtained by heating the metal in a long steel tube closed by quartz windows. By this method the author raised the number of lines of the principal series from the seven previously known to 48, and in a subsequent investigation in collaboration with Fortrat to 56. The only series of such an extent previously known was the hydrogen Balmer series of certain star spectra which showed 32 members.

A photograph of a portion of this series is shown in Fig. 97, the number of each unresolved pair being indicated. This figure is a reproduction of a photograph made in 1933 by the author and H. Straub, with a 21-foot concave grating, the source being an "end-on" hydrogen tube. A microphotometer curve of the portion near the head is also reproduced. Lines up to the 60th member were counted on the original plate. Just below this, in spectrum *b*, from the author's first paper we have the same absorption spectrum, with an iron comparison spectrum. The short wave-lengths are to the left, and the remarkable general absorption which begins at the head of the series and extends to extreme ultra-violet is shown. Spectrum *c* was taken with denser vapor, and shows that each pair of the series is immersed, so to speak, in a band spectrum, analogous to the one accompanying the *D* lines which will be considered in detail in another chapter. The continuous absorption beginning at the head of the series is of especial interest though its great theoretical importance was not recognized at the time of its discovery. Its cause is clear on the Bohr theory, for this absorption represents the complete expulsion of the electron from the atom (ionization) which can be accomplished by any frequency higher than the one required for raising the electron to the outermost of the allowed stationary orbits.

Molecular Spectra. — In the case of monatomic gases we have seen that the absorption of light is associated with the transfer of electrons from lower energy levels to higher, the emission of light

occurring when the reverse process takes place. In the case of diatomic or polyatomic gases, in which the emission or absorption is produced by molecules, the process is of a different type. We will begin by considering the case of a diatomic polar molecule, that is, a molecule in which the two atoms are charged, one positively and the other negatively, constituting a dipole. If the two atoms forming this molecule are in vibration along the line joining them, we should, on the classical theory, expect an electromagnetic radiation of the same frequency as the frequency of the oscillation, while if the molecule were in rotation we should expect a radiation of frequency corresponding to that of the rotation frequency. As long ago as 1892 Lord Rayleigh [1] pointed out that if an oscillator with frequency ν_0 is rotating with a frequency ν_r about an axis perpendicular to the direction of the oscillation it should emit two frequencies, one $(\nu_0+\nu_r)$ and the other $(\nu_0-\nu_r)$. A rotating linear oscillator will emit plane polarized light along its axis of rotation, the plane turning with the same frequency as the rotation frequency and in the Chapter on Polarization we shall see that this is the equivalent of two opposed circular vibrations of different periods. In a direction perpendicular to the rotation the intensity will fluctuate from zero (when the line of vibration is parallel to the ray) to a maximum (when it is at a right angle to the ray) the equivalent of the interference beats of two sources of different frequencies.

As Lord Rayleigh pointed out, the velocities of rotations of gas molecules are of the order of magnitude to produce a widening of the spectrum line far greater than that actually observed, and that this failure of the rotation to manifest itself as a broadening of the line constituted a grave objection to the theory of spectra. In 1912 a paper was published by Bjerrum [2] in which it was pointed out that the width of the infra-red absorption bands of gases was of the order of magnitude to be expected on the assumption that they were due to vibration in combination with molecular rotation. As a matter of fact the infra-red emission band of carbon dioxide at 4.3 μ had already been resolved into a doublet by A. Trowbridge and R. W. Wood [3] with a vacuum spectrograph and echlette grating. This work was published two years before the appearance of Bjerrum's paper although the significance of the result was not recognized at the time. In 1913 the doublet structure predicted by Bjerrum was found also for the absorption bands of hydrochloric acid, hydrobromic acid and carbon monoxide by

[1] *Phil. Mag.*, *34*, 410.
[2] Nernst, *Festschrift*, Halle, 1912.
[3] *Phil. Mag.*, Nov., 1910.

Burmeister [1] and by Eva von Bahr.[2] From the Maxwell-Boltzmann law governing the rotations of molecules in thermal equilibrium Bjerrum calculated that the separation of the doublet on the frequency scale should be represented by $1/\pi\sqrt{2kT/I}$, I being the moment of inertia of the molecule and k the gas constant. This formula furnishes us an optical means of determining the moment of inertia of the molecule. There is in addition a region of absorption in the extreme infra-red, which is regarded as a pure rotation spectrum, and its frequencies are to be identified with the frequencies ν_r of the molecular rotation. It should extend from the extreme infra-red to a frequency represented by half the frequency distance between the components of the doublet. This region is very difficult to investigate experimentally but von Bahr showed that the theory was in agreement with such experimental data as were available. The next step in the development of the theory of band spectra came with von Bahr's [3] discovery that the near infra-red band not only had a doublet structure, but under high resolving-power could be subdivided into a large number of smaller bands, which appeared to be conclusive evidence of the quantization of the molecular rotation. Bjerrum considered that the observed fine structure of the infra-red band was a proof that the angular velocity of a diatomic molecule is an integral multiple of a certain unit of angular velocity. In other words, that the molecule rotates at certain definite velocities, intermediate values being absent. In the nomenclature of modern theory this is spoken of as quantized rotation, and Nernst [4] and Lorentz [5] made the hypothesis that the energy of rotation was equal to an integral multiple of Planck's constant h, multiplied by its rotational frequency, which leads to the formula $\nu_r = mh/2\pi^2 I$ in which ν_r is one of the possible frequencies of rotation, I is the moment of inertia, while m assumes successive values 0, 1, 2, 3, etc. It was later shown by Kemble that this energy should be set equal to an integral multiple of $\frac{1}{2}h\nu_r$ and hence that $\nu_r = mh/4\pi^2 I$. It is obvious that we must attribute the fine structure of the bands to variable velocities of the molecular rotation, since the diatomic molecules can have but a single period of vibration. As we shall see later on certain phenomena show that the energy of vibration must also be quantized, but this applies only to the amplitude of the vibration which may increase by certain finite amounts, the frequency, however, remaining the same.

[1] *Verh. deut. phys. Ges.*, *15*, 589, 1913.
[2] *Verh. deut. phys. Ges.*, *15*, 710, 731, 1913.
[3] *Verh. deut. phys. Ges.*, *15*, 1150, 1913.
[4] *Zeit. Electrochem.*, *17*, 265, 1911.
[5] *Rep. Solway Conf.* 1912, p. 447.

We will now consider more in detail the production of an absorption or emission spectrum by a rotating polar molecule such as HCl or CO. Bohr's theory applied to this case calls for an absorption or emission of energy only when the velocity of rotation increases or diminishes, rotation at constant velocity in a stationary state being unaccompanied by radiation. The energy in these stationary states was shown by Schwartzschild [1] to be for the (mth) state $E_m = m^2h^2/8\pi^2I$, I being the moment of inertia and m taking the value 0, 1, 2, 3, etc. Further development showed that m must be replaced by $m+\frac{1}{2}$ giving $\frac{1}{2}$, $1\frac{1}{2}$, $2\frac{1}{2}$, etc.

If we represent these energies by levels, as we have done in the case of atomic radiation, it is clear that they will have values proportional to $\frac{1}{4}$, $2\frac{1}{4}$, $6\frac{1}{4}$, etc., and the energy differences between adjacent levels which determine the frequencies of the absorbed or emitted radiation corresponding to the change in rotation velocity will be represented by $2\frac{1}{4}-\frac{1}{4}$, $6\frac{1}{4}-2\frac{1}{4}$, etc., or 2, 4, 6, etc. During the transition m can change by one only and we have the frequencies given by $\nu = 2mB$ in which B is a constant representing $h/8\pi^2I$. This formula represents a spectrum of lines evenly spaced on the frequency scale $2B$, $4B$, $6B$, $8B$, etc. The frequency difference between adjacent lines is constant and equals $2B$, or $h/4\pi^2I$. By measuring this interval the moment of inertia of the molecule may be calculated. A series of absorption lines of this type has been measured by Czerny [2] in the remote infra-red in the case of hydrogen chloride, seven absorption maxima being found between the wave-length 40 μ and 100 μ the maxima forming an approximate arithmetical progression. The frequencies were represented by the formula $\nu = 20.794m - .00164m^3$. The small cubic term which gives a slight convergence to the period is due to change of the moment of inertia of the molecule with the angular velocity. More recent papers from the laboratory of the University of Michigan, where a great amount of work has been done in the remote infra-red with echelette gratings, are listed below,[3] of especial interest being one by N. Wright and H. M. Randall[4] on NH_3 and PH_3 in the region between 50 and 125 μ, very high resolving power being employed in this case. The transitions involved in the case of pure rotation spectra, both absorption and emission, are shown at the left of the diagram, Fig. 98, upward pointing arrows representing absorption, downward pointing, emission. The lower set of horizontal lines are the energy levels of rota-

[1] *Sitzb preuss Akad. Wiss.*, *1*, 548, 1916.

[2] *Zeit. für Phys.*, *34*, 227, 1925.

[3] Meyer and Levin, *J. O. S. A.*, *16*, 137; Meyer and Bennett, *Phys. Rev.*, *32*, 888; Meyer and Levin, *Phys. Rev.*, *34*, 44; Randall and Weber, *Phys. Rev.*, *40*, 835.

[4] Wright and Randall, *Phys. Rev.*, *44*, 391.

tion of molecules with zero vibrational energy. The heights of the levels, 0, 1, 2, 3, 4, $\cdots$ above the zero line are $\frac{1}{4}, 2\frac{1}{4}, 6\frac{1}{4}, 12\frac{1}{4} \cdots$. As we have seen, a molecule rotating in the steady state represented by line number 2, is capable of absorbing energy of frequency corresponding to the energy difference between lines 3 and 2 by which its velocity of rotation is increased until it possesses energy represented by the line 3. It can then emit radiation of frequencies represented by the energy differences between levels 3 and 2, 2 and 1, 1 and 0, the emission occurring in three successive steps. The lines would lie in the remote infra-red as the energy differences involved are small which means low frequency.

We will now take up the case of the so-called vibration-rotation band in the production of which transitions occur both in the vibrational and rotational energy of the molecule. If the force of restitution when the atoms are displaced from the position of equilibrium were proportional to the displacement, the vibration would be harmonic and a monochromatic radiation would be emitted by the oscillating dipole according to the classical theory, the intensity of which would decrease as the amplitude of the oscillation decreases. If the force is not proportional to the displacement we have what is known as the anharmonic oscillator and overtones will be emitted in addition to the fundamental. On the quantum theory the energy of the vibration is quantized, and radiation occurs only when the state of vibration changes from a level of high energy to one of the next lower energy. With a harmonic oscillator the distances between the levels is constant, consequently the frequency of the emitted vibration is the same on the quantum as on the classical theory. For the anharmonic oscillator the levels for the higher quantum number of vibrations crowd closer together. By studying the position of the lines in the spectrum we can tell at once whether it is emitted by an harmonic or an anharmonic oscillator. Two vibration energy levels for a diatomic polar molecule, each one quantized for rotation are shown in Fig. 98. Suppose the molecule to be in the state represented by the point above X, that is, on the lower vibrational level and not rotating. If radiation of the same frequency as the normal frequency of vibration of the atomic dipole strikes it, its vibrational energy will be increased by one quantum, the transition being represented by the upward pointing arrow above X which terminates on the zero level of the upper vibrational state. If the radiation has a slightly higher frequency, the molecule may receive one quantum of rotational energy in addition, this transition being represented by the adjacent arrow terminating on line 1 of the upper level. Each of the two transitions will give

rise to an absorption line but the former is not usually found. In other words, this line corresponding to the natural frequency of the oscillator is absent and is spoken of as the missing line. This means that a transition in vibrational energy is usually accompanied by a change in rotational energy. Suppose now that in a gas, the molecules are rotating with different velocities but without vibration, that is lying on the various rotational levels of the lower vibration level. By the absorption of suitable frequencies, that is suitable amounts of energy, they will be carried from these levels to the rotation levels of the upper vibration level as shown in Fig. 98, the right-hand group of arrows representing transitions in which the rotation is increased by one quantum, the left-hand group transitions in which the rotation is diminished by one quantum. These absorbed frequencies correspond to an equidistant series of lines symmetrically spaced to the right and left of the missing line Q. In many cases we may have an emission spectrum corresponding to the absorption spectrum just described in which the transitions take place in the opposite direction.

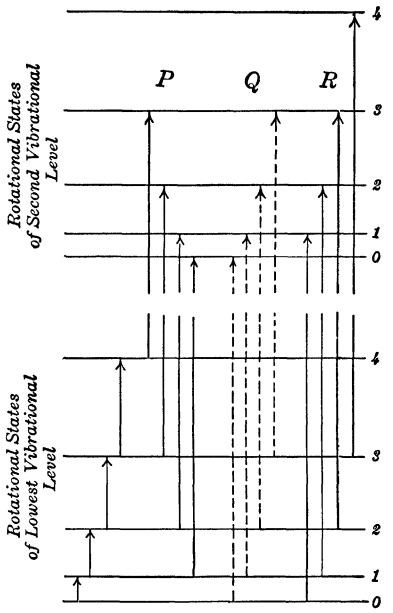


FIG. 98

It is to be noticed that, if the absorption transition is accompanied by a *decrease* of rotational energy, the corresponding emission process involves an *increase*.

If m' is the quantum number in the upper state, and m'' in the lower state then transitions for which $m'-m''=+1$ give lines on what is called the positive or R branch, while $m'-m''=-1$ transitions give lines on the negative or P branch. Some confusion has resulted from definitions based on initial and final states.

The frequencies of the lines constituting these spectra are represented by the formula $\nu=\nu_n+2Bm$ in which ν_n is the frequency of the missing line, that is the natural frequency of vibration of the diatomic molecule. $B_0=h/8\pi^2I_0$ while m assumes positive or negative integral values yielding the two branches

of the band previously alluded to. The line obtained by putting $m=0$ may be called the zero line of the series. It may be remembered that the zero line is never an actual line but is always missing. The frequency difference between adjacent lines is equal to $2B$ as in the case of the pure rotation spectrum. The arrangement of the lines in the two branches is shown in Fig. 99, the transition involved being indicated below each line. This figure, however, is obsolete and represents no actual band. If the vibration is anharmonic the energy difference between adjacent vibration levels is not equal, and we have harmonics of the fundamental vibration which give rise to other bands also having fine structure due to the rotational transition. The doublet of the Rayleigh-Bjerrum theory is formed by the unresolved positive and negative branches of the bands just discussed. A typical double band of this nature is that of hydrochloric acid which in addition to the pure rotation bands in the remote infra-red discovered by Czerny has a strong band at 3.4 μ shown in Fig. 100. This was first resolved into fine structure by von Bahr and investigated more thoroughly by Brinsmade and Kemble,[1] who found in addition the first harmonic of the band, resolving it as a doublet but not getting the fine structure which was subsequently found by Imes [2] who used the grating with which he was able to get in addition the fine structure of the absorption of bands of HBr and HF. The second harmonics of CO and HCl and the first harmonic of HF have been found by Schaefer and Thomas.[3] The number of lines forming the band will obviously increase if we increase the rotational energy of the molecule, for this will occasion absorption corresponding to higher initial values of M, the rotational quantum number. In 1923, Colby, Meyer

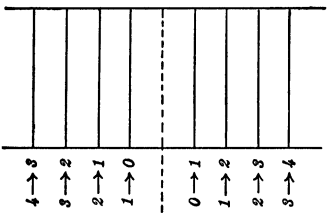
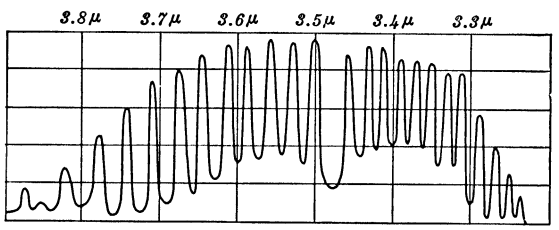


FIG. 99



HCl Band (by Imes)

FIG. 100

[1] *Proc. Nat. Acad. Sci., 3*, 420, 1917.
[2] *Astrophysical Journal, 50*, 251, 1919.
[3] *Zeit. für Phys., 12*, 330, 1922.

and Bronk [1] measured the 3.4 μ band of hydrochloric acid with the gas at 300° Centigrade and were thus enabled to extend the band on both sides to the 20th member giving 40 lines in all. They also discovered five new lines which apparently belong to a weaker overlapping band. Colby states that these new lines were observable only at the higher temperature, and must therefore be believed to have an initial state very infrequent under ordinary conditions. This would be the case for the vibrational transition 1 to 2 which moreover should fall close to the fundamental band 0 to 1.

The rotation and vibration of polyatomic molecules present considerably more intricate problems as Colby remarks in his report on Band Spectra, and, as one might expect, the bands are on the whole much more complicated. The work is of the greatest importance in the study of molecular models as revealed by their normal modes of vibration. Considerable work has been done already with vapors of this class, H_2O, CO_2, NH_3, CH_4 and HCN having been very thoroughly investigated by Sleator and E. R. Phelps using an echlette grating designed for this region. The most recent investigations have been made at the Physical Laboratory of the University of Michigan, by means of echlette gratings, preliminary analysis of the light being made with a spectrometer furnished with five rock-salt prisms. The further study of these absorption bands has already thrown much light upon the subject of the grouping of the atoms in the molecule just as the study of the line spectra of the elements has furnished most of our information about atomic spectra.

Electronic Bands. — Another type of band occurs if the electron system of the atoms making up the molecule is disturbed. This requires in general more energy abstraction from the incident light, which on the quantum theory means that the absorption will be for waves of higher frequency, that is in the visible or ultraviolet region. These electronic bands involve the electron frequencies as well as the vibration rotation frequencies, and the quantum theory attributes the bands to simultaneous changes in rotational, vibrational and electronic quantum numbers.

In the case of monatomic gases we have seen that the line absorption spectra result from the jumps of electrons from lower to higher energy levels. If this process occurs in an atom which is united to another atom forming a molecule, simultaneous changes in the vibration and rotation energy will occur giving rise to a band consisting of two or more branches as we shall see presently. In the case of absorption these bands usually lie in the ultra-

[1] *Astrophysical Journal*, 57, 7, 1923.

violet, iodine and bromine being exceptions, while in the case of emission, where we have to do with excited states in which the energy levels are close together and the electron frequencies low, the band may extend up into the visible or even the near infra-red. The cyanogen bands of the carbon arc and the fluted bands of metallic oxides seen in the arc are examples of this type. The lines of the electronic band differ from those of the infra-red bands in that they are not spaced at equal distances on the frequency scale but crowd together at a head, resembling somewhat the series lines of the elements, such for example as the Balmer series of hydrogen. The law of convergence, however, is quite different from that which obtains in the case of the spectra of atoms. The following account of the electronic bands is quoted from Page's article on Quantum Dynamics in the *Bulletin of the National Research Council* report on the "Molecular Spectra in Gases."

"In the pure rotation spectrum the motion of the molecules consists solely of a rotation about an axis through the center of mass. In the rotation-vibration spectra this rotation is accompanied by vibration of the two ions along the line joining them. To account for the ordinary bands of the visible and ultra violet it is necessary to assume that a change in the electronic configuration of the atoms accompanies the changes in the motions of rotation and vibration which occur during a transition. In fact, in the case of a non-polar molecule, an electric moment exists only by virtue of this electronic configuration, and therefore the correspondence principle requires it to change during a transition. An alteration in electronic configuration will change the forces holding the atoms together and therefore change the equilibrium distance between them. Hence it is no longer legitimate to assume the moment of inertia I_0 in the final and initial states to be the same. Let the term in the frequency of the emitted radiation due to the change in energy of the electronic configuration be denoted by ν_e. Then if we neglect the mutual energy of rotation and oscillation $\nu=\nu_e+\nu_n+hm'^2/8\pi^2I_0'-hm''^2/8\pi^2I_0''$.

"For the positive (R) branch, if the electronic angular momentum vanishes, $m'=m''+1$ and hence

$$\nu=A_1+2B_0m+Cm^2 \qquad \text{(a)}$$

where m is put for m'', and

$$A_1\equiv\nu_e+\nu_n+\frac{h}{8\pi^2I_0'}$$

$$B_0\equiv\frac{h}{8\pi^2I_0'}$$

$$C \equiv \frac{h}{8\pi^2}\left(\frac{1}{I_0'}-\frac{1}{I_0''}\right)$$

from which it is evident that C is generally small compared to B_0. C may be either positive or negative. The negative (P) branch, due to transitions in which the energy of rotation increases, is obtained by putting $m'=m''-1$ giving

$$\nu = A_1 - 2B_0 m + Cm^2. \qquad \text{(b)}$$

"The molecule may, however, have an electric moment at right angles to the plane of rotation. Hence m may remain unaltered during a transition. This gives rise to the zero (Q) branch

$$\nu = A_0 + Cm^2 \qquad \text{(c)}$$

where $A_0 \equiv \nu_e + \nu_n$.

Equations a, b and c are parabolas in m.

"In Fig. 101 ν is plotted against m for the three branches, the points for which m is an integer being projected to the diagrammatic spectrum in the lower part of the figure. The zero branch

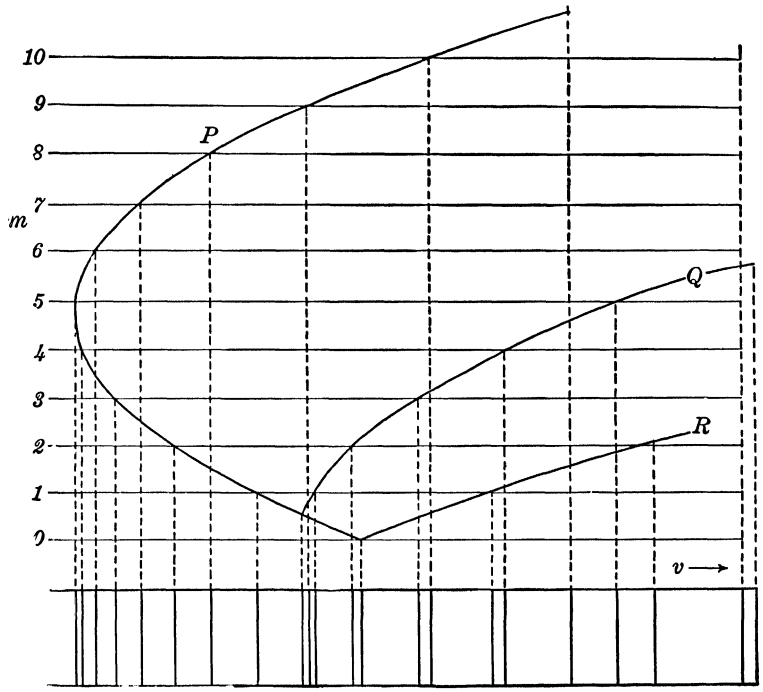


FIG. 101

meets the axis of abscissae at the frequency $\nu_e+\nu_n$, while the point of intersection of the other two branches lies at a distance $h/8\pi^2 I_0'$ to the right of this frequency. The head of the band is the line nearest to the vertex of the parabola representing the negative branch in the case in which C is positive."

Heurlinger pointed out in 1917 that certain electronic bands resembled in structure the infra-red bands of the diatomic gases which were due to rotation-vibration. Treating the moment of inertia of the molecule as variable owing to the elastic nature of the inter-atomic bond, the fundamental equations of the Bjerrum theory led to a frequency formula of the type $\nu=\nu_0\pm 2Bm+Cm^2$. In this formula ν_0 denotes either the frequency of the atomic vibration or a combination of the frequency of an electron with that of the atomic vibration. For an absolutely rigid molecule, in which the moment of inertia is the same in the two states the term in the parenthesis (page 155) becomes equal to zero and the quadric term in the expression for the frequency disappears. If there is a slight difference in the moment of inertia in the two states, C will be small in comparison to B, and the lines of the bands will be nearly equidistant in the vicinity of the frequency ν_0 as is the case with the absorption bands of hydrochloric acid in the infra-red. The existence of the small quadric term causes the spacing between the lines of the band to increase in one branch and decrease in the other as we recede from ν_0 in the positive and negative direction. If on the other hand C is relatively large as will be the case if there is considerable difference between the moment of inertia in one state and that in another, the lines in the negative branch will (for $+C$) crowd together as m increases and the band will eventually double back on itself forming a head, there being two lines in the band for each value of m, the rotational quantum number. This is more clearly seen by plotting the frequencies against the rotational quantum numbers N, a graph first employed by Fortrat for the representation of band spectra. Such a graph is shown in Fig. 101, the position of the lines of the bands being immediately below the intersections of the parabola with the energy levels m. In addition to the negative or P branch and the positive or R branch, certain bands show a third branch designated as the Q branch. The lines of this branch are formed by transitions in which there is no change of the rotational energy of the molecule. The frequencies of the lines in this band are given by the formula $N=A_0+Cm^2$.

CHAPTER VI

INTERFERENCE OF LIGHT

Thus far we have treated single disturbances only, and have not considered the effect at a point when two or more trains of waves act on it simultaneously. We know from observation that two rays of light will cross each other without interference in so far as can be seen in the region beyond the crossing point. The feeble rays from a faintly illuminated object will cross a region traversed by rays of great intensity without being influenced in any way so far as we can see. In this respect then light does not interfere with light. When two light-waves act simultaneously on a point at the same time, the displacement is the algebraic sum of the displacements that would be produced by the waves acting separately. This is known as the principle of superposition. It was stated by Huygens in 1678 as follows. "The displacement, due to a source of small vibrations, is the same whether it acts alone or in conjunction with other sources, provided the displacements are small." This is the fundamental principle which underlies the whole subject of interference. The resultant amplitude depends not only on the separate amplitudes of the two waves, but on their relative phase. If the phases agree, or the waves reach the point simultaneously, the amplitudes are added: if they are half a wave-length apart the smaller amplitude is subtracted from the larger. We shall presently determine the amplitude for any specified phase relation. If, however, zero illumination is produced at a point, there is always increased illumination at some other point, since energy cannot be destroyed. Interference then produces only a redistribution of the light. Now the energy in wave-motion exists partly as kinetic and partly as potential, that is, we have displaced particles at rest but possessing potential energy in virtue of their displacement from their position of equilibrium, and particles moving across the line of equilibrium which possess kinetic energy only. Other particles on the wave possess both potential and kinetic energy, and it can be shown that the total energy of the wave is equally divided between potential and kinetic. Let us now determine the relation existing between the energy and the amplitude.

Average Kinetic Energy of a Vibrating Particle. — The displacement of a particle at any time t is given by the equation

$$y = a \sin(\omega t - \alpha).$$

Its velocity at any moment then will be $v = dy/dt = a\omega \cos(\omega t - \alpha)$ and its kinetic energy $1/2mv^2$, where m represents the mass of the particle.

The velocity varies from 0 to $a\omega$, as is clear from the above formula, and the mean energy during a complete vibration of periodic time T is

$$\frac{1}{T}\int_0^T \frac{1}{2}mv^2\,dt = \frac{ma^2\omega^2}{4T}\int_0^T 2\cos^2(\omega t-\alpha)\,dt = \frac{ma^2\omega^2}{4T}\int_0^T \{1+\cos 2(\omega t-\alpha)\,dt\}$$

$$= \frac{ma^2\omega^2}{4T}\Big|_0^T \left\{t + \frac{1}{2\omega}\sin 2(\omega t-\alpha)\right\} = \frac{1}{4}ma^2\omega^2, \text{ in which } \omega = \frac{2\pi}{T}.$$

The average energy is therefore $ma^2\omega^2/4 = m\pi^2a^2/T^2$, which can be taken as the measure of the intensity, if we define intensity as the energy in unit volume of the vibrating medium. It can be proved that the total energy is evenly divided between kinetic and potential, and since we have only considered the kinetic energy in the above treatment the total energy will be double the amount calculated. We can also define intensity as the quantity of energy transmitted in unit time across unit cross section of a plane perpendicular to the direction in which the energy is travelling. In this case the velocity of propagation enters as a factor, and we must multiply the quantity calculated above by $v = \lambda/T$, which gives us $m\pi^2a^2\lambda/T^3$.

The important thing to notice is that the intensity varies directly as the square of the amplitude, and inversely as the square of the periodic time. The first is of importance in the study of interference, the second in considering the laws of radiation which will form the subject of a subsequent chapter.

If we are dealing with two sources of light which emit monochromatic radiations of the same periodic time or wave-length, their intensities are in the ratio of the squares of their amplitudes.

In comparing the intensities when the periodic times are different, we cannot use the eye, for it is impossible to judge accurately of the equality between two different colors. Moreover the eye cannot directly determine the true intensity, for, as we know, the true intensity or energy of the extreme red end of the spectrum is far greater than that of the yellow, while the eye is more strongly impressed by the latter. In comparing the intensities of two sources which do not emit similar radiations, we must resort to some measuring instrument which reduces them to energy of the

same type, for example the thermopile or bolometer, which measures their heating power. Since the intensity of radiation varies as the inverse square of the distance from the source, as can be proved by the most elementary methods, it follows that the amplitude varies inversely as the distance.

Composition of Vibrations. — If we have a point moving in a circular orbit with a uniform velocity, the projection of this point on any diameter of the circle moves with harmonic motion, just as does a particle vibrating under the influence of a force directly proportional to its distance from its position of equilibrium. The point moving in a circle has an acceleration V^2/r (directed towards the centre), where V = the orbital velocity, and r = the radius of the circle. This acceleration can be resolved into two components parallel and at right angles to the given diameter AA'. The one parallel to the diameter is $V^2/r \times x/r$, where x is the distance of P, the projection of the point on the diameter, from the centre of the circle (Fig. 102).

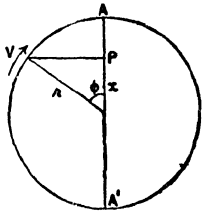


FIG. 102

The acceleration of P is then $V^2/r^2 \cdot x$, directed always towards the centre, and proportional to its distance from the centre. This acceleration is similar to that which the particles of an elastic body receive when moved out of their position of equilibrium, and we assume the ether particles acted on by a force of a kindred nature. The velocity with which the point P moves on the diameter is $v = V \sin \phi$, where ϕ represents the phase.

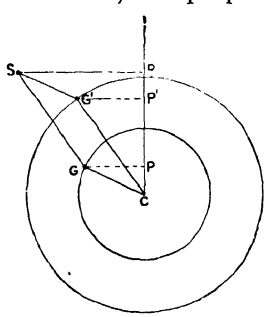


FIG. 103

Suppose now we require the effect on a point of two harmonic motions of equal periods and different amplitudes and phases. We can represent their motions by constructing two concentric circles with radii proportional to the amplitudes (Fig. 103).

The two harmonic motions will be represented by the projections on a diameter of two points G and G', which move around these circles with equal angular velocity. P will then represent the position of the particle at a given time as due to the motion represented by G alone, while P' will represent its position at the same time as due to the motion represented by G' alone. If both these motions are impressed simultaneously, the position of the particle will be represented by R, so situated that $RC = PC + P'C$ (by the

principle of superposition). The phase difference between the two vibrations is the angle GCG', which of course remains constant. If we complete the parallelogram $GCG'S$, R will represent the projection of S on the diameter, and as the parallelogram turns with G and G', the motion of R, the projection of S, will represent the resultant motion. The diagonal of the parallelogram is evidently the amplitude of the resultant vibration, and its square measures the intensity. Consequently if a and a' are the amplitudes of the component vibrations, and e the phase difference between them, the resultant intensity will be

$$I = a^2 + a'^2 + 2aa' \cos e.$$

Distribution of Illumination. — If we have two similar sources of light, which are vibrating in unison, the value of e in the expression which we have just deduced will vary from point to point. Let us consider the distribution of illumination along a line, perpendicular to the direction in which the two sources lie. In this case we will consider that $a = a'$ since the sources are similar, and we will consider the sources as lying on each side of the axis of ordinates. Taking distances along the other line as abscissae, and representing the illuminations as ordinates, we have the illumination due to one source represented by a straight line parallel to the axis of abscissae, the ordinate of which is a^2. With both sources acting together the amplitude will vary from point to point; on the axis of ordinates, where the disturbances arrive in the same phase, we have the amplitude $2a$ and the intensity $4a^2$. We can express our abscissae in terms of the phase difference. If this is 90°, or the waves arrive a quarter of a wavelength apart, cos $e = 0$ and the illumination is $2a^2$, or double that due to one source acting alone. For $e = 180$, cos $e = -1$ and the illumination is zero. Intermediate points can be determined by assigning different values to e, a curve similar to that shown in Fig. 104 being the result.

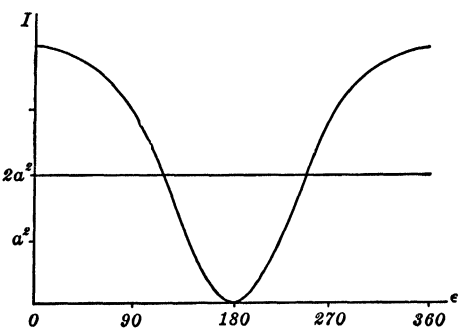


FIG. 104

If now there be no loss of energy the total illumination must remain the same; we can represent this by the area comprised between the curved line and the axis of abscissae. The total in-

tensity due to the two sources acting without interference, which would be the case if they did not vibrate in unison, would be $2a^2$. This is true of course only when we consider the average illumination for a time which is long in comparison to the time between certain assumed abrupt changes in the phases of the vibrating sources.

If no energy is lost the area between a line parallel to the axis of abscissae of ordinate $2a^2$, and the two ordinates erected at $e=0$ and $e=360$ should be equal to the area of the curve within the same limits.

The total illumination along a distance Δx on the axis of abscissae is

$$I_1=2a^2\Delta x,$$

if we assume no interference.

With interference taking place the total illumination is $I_2=\int_x^{x+\Delta x}(2a^2+2a^2\cos e)\,dx$, in which x is the value for which $e=0$.

Since e is a linear function of x we can write $e=K\Delta x$, in which K is a constant, and if Δx represents the distance from $e=0$ to $e=360$

$$K\Delta x=2\pi.$$

Integrating we have

$$I_2=2a^2\Delta x+\int_x^{x+\Delta x}2a^2\cos Kx\,dx,$$

$$I_2=2a^2\Delta x+\frac{2a^2}{K}[\sin K(x+\Delta x)-\sin Kx],$$

$$I_2=2a^2\Delta x=I_1.$$

It must be clearly understood at the outset that to have *permanent* interference, the phase relation between the two sources must remain constant, or they must be similar; their periodic times of vibration must be the same, and any changes of phase which occur in one must occur also in the other. The only way in which this condition can be attained experimentally is by making one source the image of the other, or by dividing the bundle of rays which issue from a single source into two portions, either by reflection or refraction, and then reuniting them.

Resultant of a Large Number of Disturbances of Arbitrary Phase. — In the case of illumination by actual sources of light, say a sodium flame, we have countless radiating sources in all possible phases, and the disturbances from all at a given point will produce an illumination depending, at any given point and time, upon the phase relations between the different disturbances. If

there are n radiating sources and the phases at the point happened to agree, the illumination would be n^2: at another moment there might be complete cancellation of effects and zero illumination. If only a few sources were operating the illumination would fluctuate. The problem was first solved rigorously by Lord Rayleigh who showed by the theory of probabilities, that the average illumination, due to a large number n of disturbances of arbitrary phase is n times the illumination due to a single one of the disturbances.

Interference of Light. — Grimaldi, who was the first to observe accurately and describe diffraction, or the bending of light around the edges of obstacles, described as early as 1665 an experiment which he believed proved that darkness could be produced by the addition of light to light. He admitted sunlight into a darkened room through two neighboring pin-holes, and received the light on a white screen. Each pin-hole casts on the screen a circular image of the sun surrounded by a feebly illuminated ring. By placing the screen at such a distance from the pin-holes that the outer rings overlapped, the outer edge of the ring formed by one of the holes being tangent to the outer edge of the sun's image formed by the other, he observed that the edge of the ring was less brilliant in the overlapping portion than at other places. We shall see presently that interference could not have occurred under these conditions, for two sources of light, in order to produce permanent destructive interference at a given point, must be similar — that is, must be vibrating in unison with similar amplitude and period — and two pin-holes illuminated by sunlight would not fulfil these conditions unless they were less than 0.05 mm. apart, as will be proved later on.

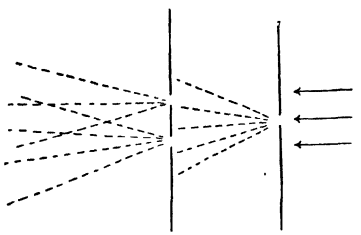


FIG. 105

A century later this experiment was modified by Young, and true destructive interference of light observed. Young passed the sunlight through a pin-hole, and then received the diverging cone on two other pin-holes (Fig. 105). From each one of these a divergent cone of light spread out, and where these two cones of light overlapped on a screen, he observed dark and light bands. In this experiment, the two pin-holes lie on the wave-front of the disturbances coming from the first hole, consequently they are always in the same phase. The dark bands are the loci of points situated at distances from the two pin-holes, differing by an odd number of half-waves. The fringes in this

experiment, being produced by diffracted light, did not prove that two streams of ordinary light could destroy one another at a point. Diffracted light was not well understood at the time; some modification was supposed to have taken place, and the fringes might be due in some way to this modification.

Fresnel realized the importance of producing two streams of light, capable of interfering and containing no diffracted light. The streams must come from two similar sources, and not pass the edges of any obstacles. This was accomplished by Fresnel by reflecting the rays from a point source of light from two mirrors inclined very slightly towards one another. Two virtual images of the point were thus formed behind the mirrors, separated by a very small distance, depending on the angle between the mirrors. Two mirrors of silvered or black glass, receive light from a point source at S (Fig. 106).

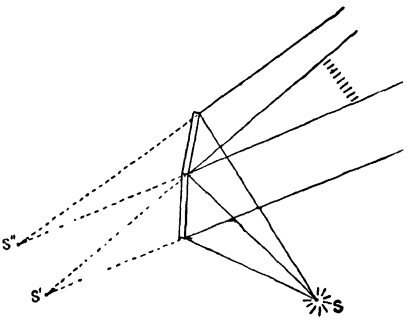


FIG. 106

The light reflected from the two mirrors comes then from two virtual images S' and S'', which lie very close together if the angle between the mirrors is small. We thus have rays coming from the two similar sources, S' and S'', and within the region where they overlap interference takes place. The light, instead of being uniformly distributed, is collected, as it were, into bright lines with dark spaces between them. The dark bands are the places where the waves from the two sources arrive half a wave-length apart and destroy each other: at the bright bands the waves arrive together, and we have reënforcement. It is evident that as the angle between the mirrors increases, the two virtual sources S' and S'' approach, coalescing when the angle equals 180.

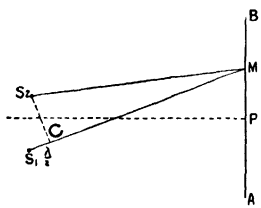


FIG. 107

Let us now examine the form and position of the fringes.

Inasmuch as we can consider the virtual sources S' and S'' as if they were real points of light, we will suppose the mirrors removed, and consider the illumination on a screen placed at a distance "a" from the sources. Let AB be a section of the screen (Fig. 107). At P, which is on a line perpendicular to the line joining the sources

at its middle point, we shall have a maximum illumination, since P is equidistant from the sources, and the waves starting together reach it in similar phase and reënforce. Going away from P we shall find a point M half a wave-length nearer S_2 and S_1, and here the waves will arrive half a wave-length apart, and destroy one another. If we advance a little further along the line AB we shall reach a point where the path difference is a whole wave-length, and we shall have another maximum.

If a is the distance of the screen at P from the sources and S the distance between them, the distance x of any bright or dark fringe is given by $x=(a/s)n(\lambda/2)$, odd values of n corresponding to dark bands, even values to light.

It is clear from the diagram that the point P will be a maximum for light of any color or wave-length. If the source of light is white this central band will also be white. The positions of other maxima being a function of the wave-length, it follows that the spacing between the bands will be different for the different colors, consequently there will be an overlapping, and instead of white fringes with dark spaces between we shall have colored fringes, the dark minima being absent except in the immediate vicinity of the central white band.

We will now take up a more complete investigation of the distribution of the maxima and minima in space. The locus of all points equidistant from two points is a plane perpendicular to the middle point of a line joining the points. The first maximum is then a plane lying between the two sources. The second maximum is the locus of all points in space so situated that the differences between their distances from the sources is one wave-length. Points fulfilling this condition lie on a hyperboloid of revolution, the sources being the foci, for by definition an hyperboloid is a surface generated by the movement of a point in such a way, that the difference between its distances from two fixed points is a constant. The locus of the second maximum will be another hyperboloid with a constant difference of 2. The loci of the maxima and minima in space form a system of confocal hyperboloids, and the fringes formed on a screen intercepting them will be hyperbolae.

In the case of our interference experiments the luminous points are so near together, and the screen so far removed, that its intersections with the hyperboloids are approximately straight lines.

Very satisfactory Fresnel mirrors can be made of modern mirror glass, or even of thin plate glass, unsilvered. Silvered glass is preferable owing to its greater reflecting power. The varnish can be dissolved from the silvered surface with alcohol, and the metal film polished. If glass of this description cannot be procured, a piece of

thin plate glass can be chemically silvered. Two pieces measuring about 2 cms. along each edge are laid side by side on a second piece of plate glass, the outer edge of one being raised slightly by means of a narrow strip of thin paper. The edges of the plates should be in contact and both should be pressed against the supporting plate. They are then fastened in this position with a little sealing wax. The angle between the plates should be such as to make the reflected images of an illuminated slit (situated at a distance of 40 or 50 cms.) appear about 3 mms. apart. The slit should be backed with a bright sodium flame and the mirrors mounted about 30 cms. from it. The dividing line between the mirrors should be adjusted accurately parallel to the reflected images, which lie on either side of it, and the field examined at a distance of 20 or 30 cms. from the mirrors with an eye-piece or pocket magnifying-glass. The eye-piece should be held at the point at which both of the reflected images are seen. If the fringes do not appear at once they can usually be brought into view by readjusting the mirror for parallelism with the slit, the field being watched with the eye-piece. The distance of the nth fringe from the centre of the system is given by

$$x = \frac{a+b}{2a\omega} n \frac{\lambda}{2},$$

a and b being the distances of the slit and the plane in which the fringes are seen from the mirrors, and ω the very acute exterior angle between the mirrors. If we measure this angle, which we can do with a spectrometer, and the distance between the fringes, we can determine roughly the wave-length of the sodium light.

Lloyd's Single Mirror. — Even simpler than the Fresnel mirrors is the device employed by Dr. Lloyd. Here the light streams from the source and its reflected image are made to interfere. The experiment is easily repeated with a strip of plate glass thirty or forty centimetres long and three or four wide, mounted in a clamp-stand with its surface vertical. The illuminated slit is placed a little beyond the further end, and one or two millimetres in front of the plane of the surface. If the eye is now brought up to the opposite end, the slit and its reflected image are both seen, and the fringes are easily found at this point with an eye-piece. Dr. Lloyd found that the centre of the system did not lie on the plane of the surface, as might be expected, but was displaced by the width of half a fringe. This is due to the phase change which the light experiences on reflection. As the mirror is turned slowly about a vertical axis, the distance between the fringes changes. With the images close together they are broad and very easily seen; with the images farther apart they are very fine, and only seen with difficulty.

This piece of apparatus is the easiest of all to work with, it being almost impossible to miss finding the fringes at the first attempt.

Fresnel's Bi-Prism. — In this experiment the beam of light is divided by refraction by means of a prism of very obtuse angle, as shown in Fig. 108.

The rays, originally emanating from a source at s, after refraction have directions as if they came from the two sources s' and s''. The illuminated slit should be parallel to the edge joining the two opposed prisms.

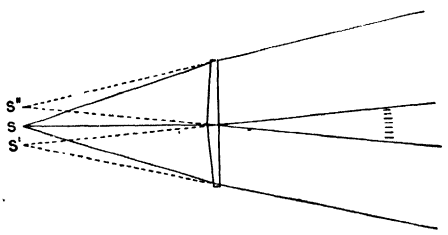


FIG. 108

The wave-length of the light can be approximately determined with the bi-prism.

If a is the distance of the source from the prism, b is the distance of the plane in which the fringes are observed, and c the distance between s_1 and s_2, we have, if we call δ the angle of deviation produced by each half of the prism,

$$c=2a \sin \delta=2a(\mu-1)\epsilon,$$

in which μ is the refractive index of the glass and ϵ the prism angle. The distance of the nth fringe from the centre of the system is given by

$$x=\frac{a+b}{c}n\frac{\lambda}{2}=\frac{a+b}{2a(\mu-1)\epsilon}n\frac{\lambda}{2},$$

which shows us that the bi-prism is equivalent to a pair of Fresnel mirrors inclined at an angle $(\mu-1)\epsilon$. A bi-prism can be easily made in the following manner. Heat a little Canada balsam in a watch glass over a small flame until a drop becomes nearly solid on cooling. Cut two pieces of thin plate glass measuring 1×2 cms., and cement them, with the long edges in contact, to a second piece of plate glass with a little of the balsam, pressing the outer edges into contact with the supporting plate, and allowing the inner edges to be slightly raised (0.5 mm. is about right) by the balsam layer. A prism made in this way works almost as well as those supplied by opticians, which are made of a single piece of glass.

In using the prism, it should be mounted at a distance of about 40 cms. from the illuminated slit, and the dividing line between the plates made parallel to it. The fringes can be found with the eye-piece in the same way as in the case of the mirrors, a little readjusting of the prism being perhaps necessary. The fringes obtained by all of the devices thus mentioned are modified by diffraction

effects, due to the fact that the waves which diverge from the two virtual sources are not complete, but are abruptly cut off at the point of union of the mirrors, or at the obtuse angle of the prism.

The Corresponding Points of the Sources. — It is clear that continuous interference can result only between streams of light which come from corresponding parts of the two sources. Our slit is backed by a sodium flame, and even if we make it extremely narrow, the phase of the vibration will by no means be constant across its width. We must remember that the sodium flame contains countless vibrating sources of light, and continuous interference can result only in the case of rays emitted by one of these and its image, or between the two images of the same vibrator. A large number of these sources will be comprised by the width of the slit, consequently interference takes place between streams of light which come from corresponding parts of the images. In the case of the Fresnel mirrors and the bi-prism the corresponding parts lie on the same side of the images, while in the case of Lloyd's single mirror they are on opposite sides. In the latter case the axis of symmetry, or the position of the central fringe, for which the path difference is zero, is the same for all the corresponding points; in the former case it is different for each pair of points. The fringes obtained with Lloyd's mirror are therefore more sharply defined, and a wider slit can be used. This will be clear by reference to Fig. 109, in which A, B, C are corresponding points, and s the axis of symmetry, which in the case of the bi-prism sources is seen to have a different position for each pair of corresponding points.

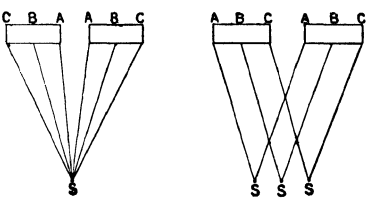


FIG. 109

Limit to the Number of Fringes. — Very interesting conclusions regarding the vibrations of the molecules in the flame can be drawn from the number of fringes which can be counted. At the first dark fringe it is clear that we have destructive interference between vibrations which left the corresponding points at the same instant. At the 100th dark fringe we have interference between a vibration from one source with a vibration from the other which left the source 100 T earlier, if T is the time of the vibration. At the 1000th dark fringe we have interference between vibrations which left the sources 1000 T apart. If now we consider that the atoms execute only about 1000 vibrations without any abrupt phase change, it is clear that we cannot have more than 1000 dark fringes; for, with a difference of path greater than 1000 wave-lengths, we shall

have a train of waves meeting another train which left the source under different phase conditions, and which may therefore reënforce, instead of destroy, one another. The number of fringes which can be observed gives us therefore information regarding the number of regular vibrations performed by each atom before an abrupt change occurs. Fizeau counted as many as 50,000 fringes in the case of sodium light, while improved apparatus and methods of modern times have raised the number to a million in the case of certain kinds of monochromatic light, from which we conclude that under favorable conditions as many as a million vibrations can be performed before any abrupt change takes place.

Shift of the Fringes by Introduction of Thin Transparent Plate. — If a thin plate of some transparent substance is put in the path of one of the interfering streams of light, the optical path will be increased owing to the retardation of the waves in the medium. If the refractive index is μ and the thickness ϵ, the increment of path is $(\mu-1)\epsilon$, in which there are $(\mu-1)\epsilon/\lambda$ waves. Increasing the path by one whole wave-length will cause a bright or dark fringe to shift into the position of its neighbor, therefore in the above case the shift will be n fringe widths; $n=(\mu-1)\epsilon/\lambda$ if we call a fringe width the distance between two bright fringes.

When the fringes are formed with white light the introduction of the plate produces a somewhat more complicated effect. This case will be discussed presently. It is clear that we can determine the refractive index of a thin plate if we know its thickness, and measure the shift of the fringes. The above formula holds only for monochromatic light, and with light of this description the fringes are similar in appearance, and the shift cannot be determined if it exceeds one fringe width, unless it can be produced gradually, as by introducing a gas slowly into a tube, the ends of which are closed with glass plates, and watching the drift of the fringes. More will be said on this subject after we have taken up the subject of the interference of white light. If we require the actual distance through which the central fringe is shifted we can easily deduce the expression $x=(\mu-1)\epsilon(a/s)$, in which s is the distance between the sources and a the distance of the screen.

Interference Fringes with White Light. — We have thus far considered our sources as sending out light of a single wave-length only. If we illuminate the slit with white light and examine the fringe system, we find that only a few rainbow colored bands are visible. The cause of this is at once apparent. The formulae which we have deduced for the distances between the fringes show us that this distance is a function of λ the wave-length, which occurs in the numerator of our expression for x, the distance of a given fringe

from the centre. It is thus clear that the shorter the wave, the closer together will the fringes lie. If we assume white light to be made up of waves of various lengths, the fringes will be out-of-step at every point save on the axis of symmetry. The central bright fringe will coincide for all the colors, but since the red bands are about twice as broad as the violet, the bands soon get completely out-of-step, and we have practically uniform illumination.

The condition of things is shown in Fig. 110, in which the red fringes are represented by the unbroken line, the violet by the dotted line. The first dark fringe on each side of the central bright one will therefore be tinged deeply with violet. The other colors will produce other systems of bands of intermediate spacing, and it is clear that at points a short distance from the centre, we shall have maximum illumination for a large number of wave-lengths regularly distributed throughout the spectrum. The resultant illumination cannot be distinguished from white light by the unaided eye, and the field therefore appears uniformly illuminated. Something resembling interference is taking place, however, in this region just as before, as we can readily prove by substituting a small spectroscope for the eyepiece, when the spectrum will be found to be crossed by dark bands corresponding to the wave-lengths, for which the position of the slit of the instrument is a position of zero illumination. We can get a better idea of the state of affairs if we consider what happens if we place the slit of the spectroscope on the central bright band and then move it slowly out into the fringe system. At the central bright band we have all colors present, and consequently see a continuous spectrum. On moving the instrument the slit enters presently into the first dark fringe for violet, and the violet of the spectrum disappears. As we move the slit along the other colors disappear in turn, a dark band moving up the spectrum. By the time that we reach the first dark band for red, we are again in a region of maximum illumination for violet, which, therefore, appears again as the dark band in the spectrum nears the red end. It is clear that owing to the difference of spacing of the fringes, the dark bands will enter the spectrum at the violet end more rapidly than they leave it at the red end; they will consequently accumulate in the spectrum, the number increasing as we move the spectroscope farther and farther away from the central band. The experiment can be easily performed by means of Lloyd's single mirror, illuminating the slit



FIG. 110

with sun- or lamplight, and substituting a small pocket spectroscope for the eye-piece. The instrument should be mounted in a clamp-stand and pointed towards the double source, the slit being close to the edge of the plate.

Fizeau and Foucault were enabled in this way to detect interference with a path difference of 7000 waves. This result was at the time interpreted as indicating that the elementary components of white light must consist of periodic wave-trains, several thousand regular vibrations being executed without abrupt change of phase. Lord Rayleigh has shown, however, that we can infer nothing whatsoever about the regularity of the vibrations of the source in this case, the limit of the number of bands seen in the spectroscope depending solely on its resolving-power. The interference, in point of fact, does not take place until after the light has passed through the spectroscope.

We now know that white light consists of irregular pulses, and that the prism or grating of the spectroscope spreads them out into long wave-trains of approximately monochromatic radiation. It is the interference between these wave-trains (which we may regard as Fourier components of the pulse), which give rise to the dark bands seen in the spectrum. This case will be more fully discussed in the section on white light.

The Flow of Energy in a System of Interference Fringes. — The interference minima formed by two similar sources of light form a system of confocal hyperboloids, and the question of the flow of energy in this case, or any similar case, was discussed by the author many years ago. Energy is obviously flowing out from both sources at its normal rate, but the direction of flow is perhaps not quite obvious. Suppose the minima equal to zero, which is nearly correct at the centre of the system. Energy evidently cannot cross a plane along which there is no disturbance.

In stationary waves, if the nodes are absolutely at rest, which is the case if the two wave-trains are of equal amplitude, we cannot speak of a flow of energy across them. A node may be considered as having the properties of a perfect reflector, that is to say the point acquires the power of reflecting as a result of the arrival of a wave travelling in the opposite direction, the reflection being accompanied by a phase change of π as we shall see presently. We are thus forced to the conclusion that the flow of energy in the case of the interference fringes must be along the hyperboloids, that is along curved paths. We can show this experimentally by means of ripples in mercury excited by two needles mounted on the prong of a tuning-fork. If we view the mercury surface through a narrow slit opened and closed by the vibrations of another fork

slightly out of tune with the first, we see the waves (stroboscopically) creeping slowly along the surface, and following the lines of the hyperboloids. Two questions now naturally occur to us. How does the energy get into the bright fringes, if the dark fringes are supposed to act as barriers? and what is the nature of the wave that is travelling along a bright fringe? In regard to the first question: the dark fringes are never absolutely black, as no one of them is equidistant from both sources. The amplitudes are therefore slightly different, and there will be a flow of energy across the minimum in the direction of the disturbance having the larger amplitude. Though it may be very slight at any given point, it is ample to account for the flow along the hyperboloid. We can take as an analogous case two parallel sheets of cloth tightly stretched, and very close together. Consider water forcing its way into the space between the two sheets from both sides. A very small flow across unit cross section will give us a large flow across unit section taken perpendicular to the sheets.

We may, however, have a fringe which is absolutely black, for there is nothing to prevent us from considering the sources as vibrating with a difference of phase of 180°. This makes the centre of the system dark, and equal to zero, and it must act as a barrier to the flow of energy from both sources. In other words, the central fringe can be considered as acting as a perfect mirror, and we can regard the fringes as formed by the interference of the reflected waves with the direct. If the flow of energy is along the hyperboloids, it is evident that in the region between the sources the flow is in a direction nearly perpendicular to the rays. We can watch this flow with the mercury ripples and tuning-forks. (A ring of castor-oil or glycerine poured around the edge of the mercury surface prevents troublesome reflections from the walls, and to a large extent waves due to jars from the table. It is analogous to painting the walls of a room black in optical experiments.) The bars of light perpendicular to the line joining the vibrating sources slide out sideways, each one of them forming one of the waves which travel along the hyperboloid. We can perhaps get a better idea of what happens if we consider what is going on in a bright fringe outside of the region between the sources. What is the type of wave, and is it capable of showing us both of the sources if it alone is allowed to enter the eye? If we regard the dark fringes as absolutely dark, that is, as perfect reflectors, we must regard the waves as travelling between them as between two silver walls. The incidence is very oblique, *i.e.* the wave is nearly perpendicular to the reflecting plane; and if we consider the wave as a portion of a sphere with its centre at one of the sources, the wave after re-

flection from the interference plane will be a portion of a sphere with its centre of curvature at the other source. This process will repeat itself over and over again, a given portion of the wave-front appearing to come first from one source and then from the other.

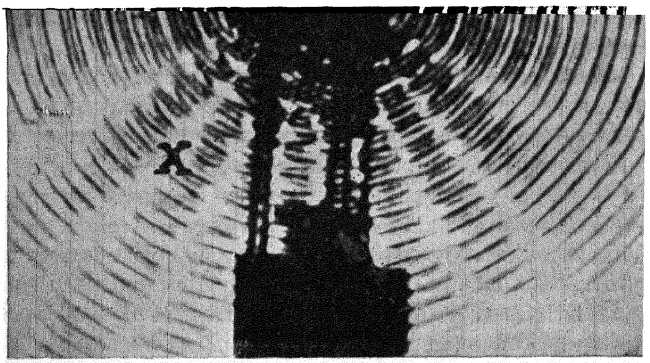


FIG. 111

The bright fringe will then contain two groups of wave-fronts inclined to each other at a small angle. These can be seen with the tuning-fork waves, in some of Mr. Vincent's photographs, published in the *Philosophical Magazine*. An enlargement of a portion of one of these photographs is reproduced in Fig. 111, the inclined wave-fronts showing especially well above the point marked *X*.

In Fig. 112 we have intersecting wave-fronts described around the two similar sources *A* and *B* (at a greater distance than shown). It is to be noted that wave-front 1, a crest centred on *A*, is reflected from the right-hand interference minimum as trough 2 centred on *B*, and trough 3 is reflected as crest 4, that is, a phase change of 180° occurs at reflection. This accounts for the fact that the inclined wave-fronts in Fig. 111 do not join, as only the crests are photographed, the condition being as if 1 and 4 of Fig. 112 were photographed.

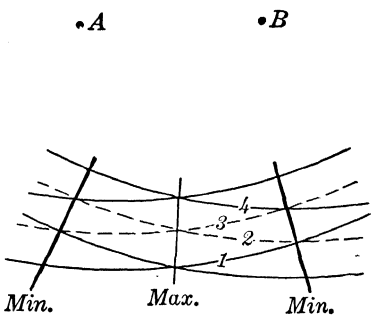


FIG. 112

Though each bright fringe contains two wave-fronts, we cannot "resolve" the sources with them, for calculation shows that their width will be insufficient. In other words, if we screen off the other bright fringes, passing the waves in the one through a slit, the slit

width necessary turns out to be just what is needed to prevent resolution, for, as we shall see when we take up diffraction, two small sources close together cannot be seen separated or resolved, if viewed through a sufficiently narrow slit. In the case of light reflected at 45° from a mirror as in Fig. 134 the flow of energy in the region in which the direct and reflected beams overlap is parallel to the surface of the mirror.

A Simple Interference Refractometer. — A very simple and inexpensive interference apparatus, which can be used for measuring the refractive index and dispersion of a gas, and for showing and measuring anomalous dispersion at the absorption bands of a gas, can be made of a slit and a long focus lens. This type of apparatus was used by Fizeau in determining the velocity of light in a rapidly moving medium (see Chapter on Relative Motion of Matter and Ether), and it has since been employed by Lord Rayleigh in determining the refractive indices of the rare gases of the atmosphere.

As usually employed its chief disadvantage lies in the fact that the fringes are very narrow, and invisible except when viewed through a very high power eye-piece. Lord Rayleigh used a glass cylinder 1 mm. in diameter, as magnification in one dimension only is required, and a much brighter image is obtained in this way.

This trouble can also be overcome by the use of long focus lenses, which lessens the convergence of the interfering rays. A lens of about 2 metres focus (spectacle lens) is covered by a screen furnished with two vertical slits a millimetre or two in width, and separated by a distance of from .5 to 3 cms., depending on the required separation of the interfering beams. If we are to measure the refractive index of a gas, we use two parallel tubes, with their ends closed by pieces of thin plate glass cut from the same piece, and the beams must be far enough apart to enable us to pass one down each tube.

The slits can be made by pasting strips of black paper across an aperture cut in a card. The illuminated slit should be well made, as it is necessary to make it very narrow. The slit of a small spectroscope can be used, and it should be mounted at a distance of 4 metres from the lens. An image of the sun or the crater of the arc should be focussed on the slit, and the lens with its two apertures placed in such a position that it receives the light from the slit. Hold a card behind the lens, and move back until the two lines of light transmitted by the slits fuse together into a single image at the conjugate focus. This is the position for the eye-piece, and we shall find the image traversed by a beautiful set of vertical interference fringes. Their distance apart will increase, as the dis-

tance between the two slits is made less, and their sharpness will increase as the first slit is narrowed.

If it is desired to receive the fringe system on the slit of a spectroscope, the three slits should be mounted in a horizontal instead of vertical position and the slit of the spectroscope substituted for the eye-piece. The spectrum will then be found to be traversed by horizontal dark bands which are closer together in the blue than in the red. This is the arrangement commonly used for the study of anomalous dispersion of gases by the interference method.

The shift of the fringes by the introduction of a transparent plate can be shown by holding a very thin flake of mica over one of the slits. If the flake is too thick the fringes will be shifted out of the field and disappear. Or we can put two pieces of thin plate glass, one in front of each slit, and by slowly turning one of them cause the fringes to move, the result depending on the increase in the path through the glass with increasing angle of incidence.

The Rayleigh interferometer was improved by Haber and Loewe who substituted for the cross hairs a second system of fixed inter-

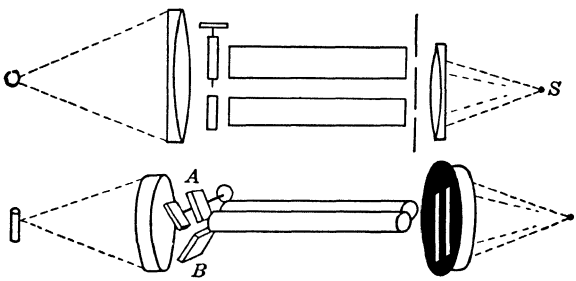


FIG. 113

ference fringes in coincidence with the first, which served as a fiduciary mark. Much higher accuracy is obtained by bringing two systems into coincidence, than by setting a cross hair on the centre of a fringe. Moreover any distortion of the framework due to temperature changes or other causes, affects both systems alike and hence introduces no error. The arrangement is shown in Fig. 113. S is a vertical slit (perpendicular to the paper) from which the light, rendered parallel by the lens passes through the two slits and the tubes to a second lens after passage through two thin glass plates at 45°, one of which, A, can be rotated on an axis, thus varying the optical path of the rays passing through it, and causing a lateral shift of the fringes. That portion of the light which, coming from the lower portions of the slits, passes below the tubes forms a system of fringes superposed on the first, when viewed

with an ordinary eye-piece, but with a cylindrical eye-piece we see the two sets one above the other separated by a dark band due to the lower floor of the tubes. The lower set is brought up into coincidence with the upper by an inclined plane-parallel plate, B, bevelled at the top so as to make the line of demarkation as narrow as possible. It is only with a cylindrical eye-piece that the fiduciary set of fringes can be employed. With a cylindrical lens we are viewing the aperture of the telescope through a simple slit so far as the vertical dimension is concerned, and if we cover the lower half of the lens the lower half of the field vanishes while in the case of a lens eye-piece, covering any part of the objective does not restrict the field of view. This point does not appear to have been brought out in recent text-books, and in one case all mention of the single inclined plate is omitted, both in text and figures.

The width of the first slit plays a very important part, for if it is too wide no fringes will be seen. In general the greater the distance between the slit apertures which cover the lens, the narrower must be the slit which transmits light to them.

That this must be so is clear from the following considerations: We imagine the first slit divided into a number of vertical linear elements. Each one of these will produce a system of interference fringes, but the individual systems will be displaced laterally with respect to each other. The angular shift between the systems due to the edge elements of the slit will be the angle subtended by the edges of the slit as seen from the lens. This angle will be independent of the distance which separates the two slit apertures on the lens. The distance between the maxima and minima of each system, however, becomes less as the distance between the apertures is increased, consequently if this is made too great, the fringes will disappear, as the maxima of one system will coincide with the minima of another. The condition for disappearance of the fringes is, for a slit of angular width e (as seen from the lens) and slit separation s

$$\epsilon = \frac{\lambda}{s}, \frac{2\lambda}{s}, \text{ etc.}$$

that is the maxima of the system formed by the light from one edge of the slit, must coincide with the maxima due to the light from the other edge. This can be seen at once by drawing the displaced intensity curves for the linear elements into which we divided the first slit. Faint fringes reappear when e exceeds λ/s as a new set of maxima and minima are superposed on the uniform field. If we consider the two edges only, or rather if we have two very narrow slits separated by a distance equal to the width of the first slit, disappearance occurs when the maxima due to one, coincide

with the minima due to the other, and this condition is represented by

$$\epsilon = \frac{\lambda}{2s}.$$

It is thus clear that we have here a method of measuring the width of a distant slit or the distance between two slits if we know how far off they are. This is the principle employed in the application of this type of interference apparatus to the problem of measuring the diameters of stars, or the distance between two double stars which are unresolvable by the largest telescopes.

Application of Interference Phenomena to Astronomical Observations. — If the first slit of the interference refractor just described exceeds a certain width the fringes are no longer visible This is due to the fact that each vertical linear element of the widened slit forms its own set of fringes, and these sets are relatively displaced, the superposition of all producing uniform illumination. The greater the distance between the two slits in front of the lens, the finer must be the front slit, in order to have the fringes appear. It is thus clear that we have a means of measuring the angular diameter subtended by a distant source of light. This method was proposed by Fizeau in 1868 for measuring the diameters of the fixed stars.

In 1874 Stefan made an attempt to carry out Fizeau's plan, placing two slits in front of the objective of the Marseilles telescope, the largest available at the time. The fringes remained visible even when the slits were separated by the full diameter of the objective. In 1890 Michelson measured the diameters of the four moons of Jupiter, using the 36-inch telescope of the Lick Observatory. The method can also be used for determining the distance between the components of a double star.

Each star produces its own fringe system, the two being displaced with respect to one another through the angle ϵ, which is the angle subtended by the stars. This angle is, of course, independent of the distance between the slits. If now the slits are close together, the fringes are broad, and the angular displacement ϵ of the two systems does not alter their appearance. As we separate the slits, however, the fringes become finer and presently disappear, owing to the fact that the maxima of one system fall upon the minima of the other. In the case of a disk as the source of light, the expression for the disappearance of the fringes is

$$\epsilon = \frac{1.22\lambda}{s}$$

the factor 1.22 being the same as that employed in calculating the resolving-power of a telescope, which will be studied in the Chap

ter on Diffraction. The centre of the disk contributes more light than the edges.

The theory of the double slit interferometer as applied to the measurement of star diameters and the distance of separation of close double stars is as follows:

In Fig. 114 a lens brings plane waves to a focus at 0. If we consider, however, that secondary waves (diffracted rays) radiate from all points on the wave in lateral directions, we must take them into account in determining the nature of the image at 0. Rays diffracted in the direction indicated by the dotted arrows have a path difference λ as indicated, and will be brought to a focus at the point marked $\frac{1}{2}$, where the illumination will be zero, since the path difference between a ray passing through the edge of the lens and one at the centre will be $\lambda/2$ as indicated, and they will destroy each other. The same will be true for the disturbances coming from other pairs of points, such as 2 and 8, 3 and 9, etc. For more oblique rays the destructive interference is less complete, and we have a faint maximum of illumination at 1, where diffracted rays from the edges with a path difference of $3\lambda/2$ are brought to a focus. We may consider the effect of $\frac{2}{3}$ of the wave-front destroyed, the illumination resulting from the remaining $\frac{1}{3}$. The intensity distribution is shown by the curve. If now we restrict the passage of light to two slits close to the edges of the lens, the path difference between the diffracted rays indicated by the dotted arrows, as before, reënforce each other at the point $\frac{1}{2}$ where zero illumination occurred before, and halfway between this point and the point 0, rays diffracted at an angle equal to one-half of the previous angle will cancel each other, having a path difference of $\lambda/2$. The fringes are thus seen to be twice as close together as before.

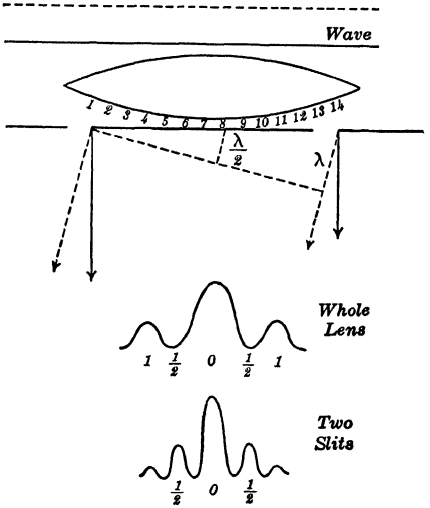


FIG. 114

If the source of the rays is double, say a double star, each component will produce its own set of maxima and minima, shifted with respect to each other through a diffraction angle equal to the

angular separation of the stars. The fringes will disappear when the maxima of one set fall on the minima of the other, and the closer together they are, the less the angular shift necessary for their disappearance, consequently they will vanish (for a given case of star separation) when the slits are employed, and not vanish when the whole lens is employed. By the use of the slits we thus have double the resolving-power of the whole lens. This matter will be better understood after the subject of diffraction and resolving-power has been discussed.

In 1920 Michelson took up the problem of the determination of stellar diameters. Even the great 100-inch telescope of the Mt. Wilson Observatory is not large enough to allow of a sufficient separation of the slits; consequently Michelson designed a "periscopic" arrangement of four mirrors, the two outer ones, twenty feet apart, reflecting the light to two inner ones which in turn reflected the beams down upon the mirror of the 100-inch telescope. The mirrors were mounted on a metal beam attached to the top of the telescope tube. The instrument was constructed in collaboration with F. G. Pease of the Mt. Wilson Observatory.

The bright star Betelgeuse was the first investigated. This star shows evidence of its diameter with the 100-inch telescope if a canvas cover is placed over the instrument, provided with two holes 7 inches in diameter and 94 inches apart, the diffraction disk of the star being crossed with *faint* interference bands. If either hole is covered the bands disappear. If the telescope is pointed at Rigel, however, the bands are clear and strong, showing that its angular diameter is smaller than that of Betelgeuse. With the twenty-foot interferometer the bands disappeared entirely in the case of Betelgeuse when the mirrors were separated by a distance of 120 inches, while Rigel showed very distinct bands. The angular diameter of Betelgeuse was computed as .047 second of arc. From the known distance of the star, its actual diameter was calculated as 250 million miles or greater than the earth's orbit about the sun. Its diameter has been found to vary, however, for at times the mirrors must be separated by a distance of 14 feet before the fringes disappear. Antares was found to be still larger, having a diameter of 400 million miles. The minimum angular diameter measurable with the 20-foot instrument is .024 second of arc.

An instrument has more recently been constructed with mirrors 50 feet apart capable of measuring stars having an angular diameter of .01 of a second which is equivalent to that of a cent seen from a distance of 250 miles! The new instrument is complete in itself,

and not an attachment for the 100-inch, its main mirror being only 40 inches in diameter.

Williams Refractometer. — A very ingenious improvement of the Rayleigh refractometer has been made by W. E. Williams.[1] In his paper he brings out very clearly that more accurate values of the refractive index and dispersion of air are necessary if reductions of wave-lengths to vacuum are not to introduce errors greater than the errors of observation, or as he says

"The ordinary double-slit arrangement of the Rayleigh instrument, with slits of width 'a,' is modified as shown in figure 115. A single slit S_2 of width $2a$ is used and the beam is divided and displaced by an Albrecht prism so that the two beams have the same relative separation W as before. As far as the observing telescope is concerned, the angle between the fringes will still be (λ/w), but the primary slit may now be opened until it subtends an angle $\frac{1}{2}\times 1.43\lambda$ at the collimator objective before the fringes disappear.

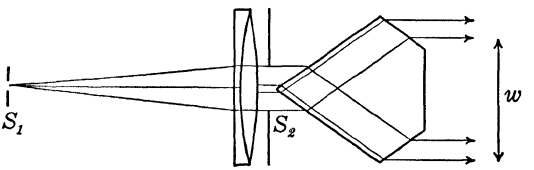


FIG. 115

The arrangement in fact practically amounts to a Michelson star interferometer used in the reverse way, with the difference that the distance between the outer beams is fixed, and that we are using the arrangement for an entirely different purpose — to increase the brightness of the fringes.

"While the same standard of visibility of the fringes is retained, the primary slit can now be opened to a width $0\cdot 715w/a$ of its previous value, so that the brightness of the fringes is increased as the square of this ratio.

"The instrument behaves (apart from the increased brightness of the fringes) as a standard Rayleigh interferometer; the angle between the fringes and the distribution of light among the fringes remain exactly as before. The fringes at the focal plane of the telescope objective are so close together that a very short-focal-length cylindrical rod of glass must be used to magnify them sufficiently. This, not being achromatic, requires refocussing in different parts of the spectrum, so that in spite of the increased brightness of the fringes, the modified Rayleigh instrument cannot be successfully combined with a spectrograph.

"The success of the first separation of the two beams from the centre naturally led to a trial of the reverse process at the telescope

[1] *Proc. Phys. Soc.*, xliv. 1932.

end. When the two beams are recombined by a second rhomb or analogous device, their separation w on entering the telescope objective is reduced to a third. The angle between the fringes increases from $(\lambda/3a)$ to $(0 \cdot 715\lambda/a)$ and the focal length of the eye-piece can be doubled while the same magnification as before is retained. As far as the telescope is concerned, what we now have is a two-plate transmission échelon, and the six or eight fringes obtained with the standard Rayleigh are reduced to the two or three orders seen with an échelon grating.

"The nett overall gain of brightness when two rhombs are used as described has not been carefully measured, but a conservative estimate would be about twenty times. The increase in brightness will be practically appreciated when it is realized that the white-light fringes with a 2-volt pocket-lamp source are brighter than the fringes of the standard Rayleigh with a Pointolite lamp source. While this is of value in making the apparatus more portable, the real advantage lies in the fact that it now becomes possible to carry out refractive-index measurements with the lines of helium, neon, etc., while previously the use of the Rayleigh with monochromatic light was confined to the stronger lines of mercury."

Interference of Waves of Different Lengths. Light-Beats. — When two tuning-forks of slightly different pitch are sounded simultaneously we hear a fluttering sound, the intensity rising and falling. The interference in this case is not continuous in time at a given point in space. If we draw two wave-trains of slightly different wave-length we shall see that they are "in step" and "out-of-step" at periodic intervals. Where they are in step we shall have maximum amplitude, where they are out-of-step we shall have minimum or zero amplitude. As the double wave-trains sweep by a given point it will be in alternation the seat of large and small disturbances.

If we seek for the optical analogy it is easy to see that two sources of monochromatic light, of slightly different period, should give us a moving system of interference fringes, any given point in space being alternately the seat of maximum and minimum illumination. The frequency of the beats being equal to the difference between the two interfering trains, the wave-length of the beat is $\lambda'\lambda/\lambda' - \lambda$.

Light-beats have never been obtained by uniting two streams of light from sources of different color. Righi has, however, performed an experiment in which the frequency of vibration of one of the two streams of light which form a system of fringes can be increased any desired number of times per second by passing the light through a revolving Nicol prism.

It is also worthy of remark that the moving fringes observed in a Michelson interferometer, as the back mirror is advanced, can be regarded as a manifestation of light-beats, the wave-lengths of one of the interfering trains being lessened by reflection from the moving mirror by Döppler's principle.

These cases are not very satisfying, however, as a statical treatment can be given as well as a dynamical, the fringes, observed in motion in Righi's experiment, merely occupying positions depending on the orientation of the Nicol. Rupp has recently described a more dynamical experiment in which a change in the wave-length of the green thallium line was produced by very rapid interruptions of the light beam, by means of a Kerr cell operated by electrical oscillations of half metre wave-length. This experiment will be described later on.

The moving system of interference fringes which constitute beats can be most beautifully shown by means of capillary waves on a mercury surface excited by two fine wires attached to the prongs of two tuning-forks of the same pitch, one fork being thrown slightly out of tune with the other by small lumps of wax attached to the prongs. The tips of the wires are dipped into the surface of the mercury and the hyperboloid fringes will be seen to be in motion, sweeping around in a most beautiful manner. Between the wire points they will be observed to travel from one vibrating point towards the other. If the wax lumps are removed, the fringes immediately become motionless. The phenomenon can be projected on a screen to advantage, by reflecting the light down upon the mercury surface, and thence to the screen through a projecting lens, by means of a pair of mirrors or large reflecting prisms.

Achromatic Interference Fringes. — As we have seen, the fringes obtained with Lloyd's mirror and a source illuminated with white light soon blend into a uniformly illuminated field, owing to the fact that the distance between the maxima and minima varies with the color. If by some artifice we can make the widths of the fringes the same, the system will become achromatic, and we can count a large number of fringes even with white light. This can be accomplished by using, as our source, a short spectrum with its blue end towards the reflecting plate. The blue sources will thus be closer together than the red, and if the adjustments are right the blue sources will give fringes of the same width as the red, which are farther apart.

The condition is best realized by employing a diffraction grating and a lens to form the spectrum. A vertical slit is illuminated with sun or arc light, and a glass grating with two to three thousand

lines to the inch, combined with a lens of four or five inches focus, is so arranged as to form a series of diffraction spectra on a card mounted in the focal plane of the lens. The card should be perforated with a small hole through which the light of one of the first order spectra is allowed to pass. The Lloyd plate is placed in such a position as to furnish a reflected image of the spectrum, the blue end of which is turned towards the original spectrum, as shown in Fig. 116. The reflecting surface should be pointed exactly at the central image formed by the grating and lens, if perfect achromatization is desired. The fringes are viewed as before with an eye-piece, a little adjusting of the plate being all that is necessary to completely fill the field of view with fine black and white lines.

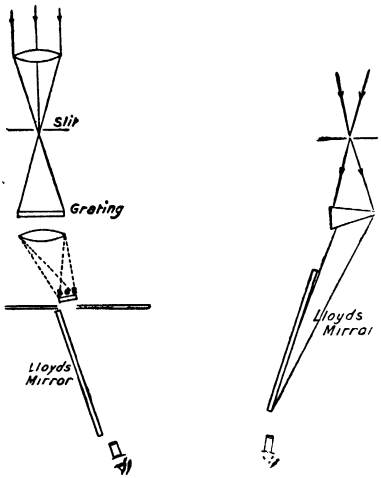


FIG. 116

If the spectrum is formed by a prism of about 20°, which can be made of plate glass, and filled with water, less perfect achromatization is obtained; still a large number of fringes can be seen.

The spectrum in this case can be virtual, *i.e.* no lens need be used, the prism being mounted between the slit and the mirror, as shown in the second diagram (Fig. 116). If the prism is placed between the mirror and the eye-piece no achromatization results, for in this case the two spectra are not opposed.

Introduction of Thin Transparent Plate. — As we have just seen, a plate of thickness ϵ and refractive index μ shifts the central fringe for monochromatic light through a distance

$$x=(\mu-1)\epsilon\frac{a}{s}$$

This distance will be different for the different colored systems, since μ varies with the wave-length, and there will therefore be no point at which the waves of all lengths will arrive in the same phase, or, in other words, there will be no strictly achromatic fringe.

There will, however, be a system of colored fringes with a central band which appears nearly achromatic, the determining condition of which is not that the path difference be equal to zero, but that

the change in phase with change of λ be a minimum. Suppose the introduction of the plate shifts the whole system to the right. Owing to the dispersion of the plate the blue fringes will be shifted through a slightly greater distance than the green, and the green a trifle more than the yellow. It is clear that this difference may close up the rainbow-colored band to the right of the central fringe, into an approximately achromatic band, and expand what was originally the white central band into a rainbow band. If we do not take this into account, we shall make an error of one fringe width in measuring the shift of what *appears* to be the central band. If our plate had a higher dispersion with the same average refractive index, the second or third rainbow band might be achromatized and appear to be the centre of the system. If the distance of the achromatic fringe from the original centre is x the difference of path is $xs/a-(\mu-1)\epsilon$, since the shift is towards the side on which the plate is introduced, and the original short path is lengthened by the introduction of the plate. Now $(\mu-1)\epsilon$ is a function of λ, and we will write it $f(\lambda)$. The difference of phase at the point in question for any value of λ will be, writing

$$D=(\mu-1)\epsilon=f(\lambda) \text{ and } D'=\frac{xs}{a},$$

$$\Delta=2\pi\frac{D'-f(\lambda)}{\lambda}.$$

The required condition of minimum phase variation with λ is given by differentiating this expression with respect to λ and equating to zero; performing the operation we get, putting $f'(\lambda)=df(\lambda)/d\lambda$

$$\frac{-2\pi f'(\lambda)\lambda-2\pi D'-2\pi f(\lambda)}{\lambda^2}=0, \text{ or } D'=f(\lambda)-\lambda f'(\lambda).$$

The central fringe corresponding to wave-length λ is shifted by the plate to the position of the nth fringe given by

$$n=\frac{(\mu-1)\epsilon}{\lambda}=\frac{f(\lambda)}{\lambda}.$$

By our original supposition regarding the position of the shifted achromatic (approximately) fringe it occupies the position of a fringe of order n' given by

$$n'=\frac{xs}{a\lambda}=\frac{D'}{\lambda}=n-f'(\lambda),$$

and is therefore shifted relatively to the central fringe for monochromatic light of wave-length λ by a number of fringes given by $n'-n=-f'(\lambda)$.

The variation of μ with λ is well expressed in the present case by the equation

$$\mu = A + \frac{B}{\lambda^2},$$

which we shall discuss more in detail when we come to the subject of dispersion.

$$f(\lambda) = (\mu - 1)t,$$

$$\frac{df(\lambda)}{d\lambda} = f'(\lambda) = -\frac{2Bt}{\lambda^3},$$

$$n' - n = \frac{2Bt}{\lambda^3},$$

which shows us that the shift of the approximately achromatic fringe obtained with white light, with respect to the central fringe obtained with monochromatic light of wave-length λ, varies inversely as the cube of the wave-length, and directly as the thickness of the plate. We shall have occasion to make use of this formula when we come to the subject of the interferometer.

A remarkable instance of the shifting of the region of fringe visibility far out into the system was observed by the author [1] in studying the dispersion of sodium vapor with the interferometer. The path difference under which it is possible to obtain interference fringes with the D_3 light of a helium tube can be nearly trebled by the introduction of a small amount of sodium vapor into the path of one of the interfering beams. This development of fringes far out in the system by the dispersive action of the vapor is accompanied by their complete disappearance at the centre of the system, where the difference of path is zero.

The introduction of a medium into the path of one of the interfering beams causes a shift of the fringe system as a whole, and if the medium is dispersing, the shifts will be different for the different colors. The red, green and blue fringes, which are out-of-step at a given point, may thus be brought into coincidence by the inequality of their respective displacements. In this case, however, since the systems are shifted as a whole, the fringes will be thrown out-of-step at the centre of the system, consequently we have obtained an increased visibility far out in the system at the expense of visibility at the centre. Now the helium light is very near the D lines of sodium, and sodium vapor in this region of the spectrum has a dispersive power so great that a prism of it giving the same deviation as a 60° glass prism (if it could be formed) would separate two lines only $\frac{1}{20}$ as far apart as the D lines, by an amount as

[1] R. W. Wood, "Achromatization of approximately monochromatic interference fringes by a highly dispersive medium," *Phil. Mag.*, September, 1904.

great as the distance between the red and the greenish-blue of the spectrum yielded by the glass prism. This enormous dispersive power may well be expected to modify profoundly the appearance of the fringe system produced even with light as monochromatic as that of the D_3 line. That a change is produced depends on the fact that no light is absolutely monochromatic, the finest spectrum lines having an appreciable width. We can thus consider the D_3 light as an extremely short spectrum, and apply the same reasoning as in the case of a thin transparent plate introduced into the path of one of the interfering streams of white light.

The treatment will be better understood after a study of the interferometer and the resolution of spectral lines, but it is given here, on account of the identity of the phenomenon with the displacement of the white centre. The helium fringes under ordinary circumstances disappear when the path difference is between 1.5 and 2 cms., there being no recurrence of visibility by further increment of path difference as in the case of sodium light. We must therefore regard the helium (D_3) line as a single line of finite breadth

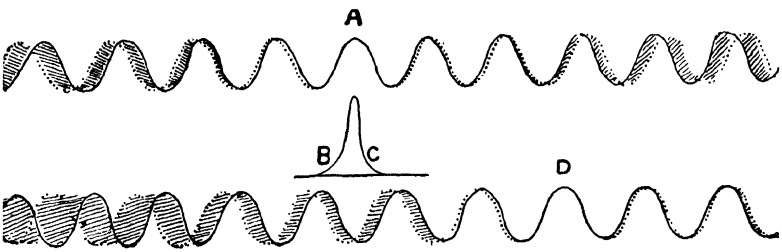


FIG. 117

or a close group of lines. In Fig. 117 let BC represent the intensity curve of the helium light, C being the edge of shorter wave-length. Immediately above we have a schematic representation of the fringe system, with its centre at A. Light from the side B of the D_3 line will produce the fringes indicated by the dotted line, which are farther apart than the fringes formed by the light of shorter wave-length coming from the side C of the line. There will, in addition, be an infinite number of other systems formed by light of wave-lengths intermediate between B and C which are indicated by light shading.

Now suppose sodium vapor to be introduced into one path of the instrument, and the whole system shifted slightly to the left in consequence. Owing to the enormous dispersive power of the vapor, the dotted system (longer λ's) will be shifted more than the other, since the D_3 line lies on the blue side of the sodium absorption

band, and the change in the velocity of the light is greatest for the longest waves, namely, those on the B side of the line. The result of this dispersive action is that the fringes are brought into step at a point D, to the right of the centre, thrown out-of-step at the centre and still more out-of-step to the left of the centre.

The achromatizing action of the sodium vapor is most beautifully shown if we illuminate the interferometer with white light.

Under ordinary conditions only two or three black and white fringes are seen, bordered on each side by perhaps a dozen rainbow-colored bands, which fade rapidly into a uniform illumination. If sodium vapor is formed in one of the interferometer paths, the colored fringes rapidly achromatize, and increase in number, breaking up, however, into groups. As the density of the vapor increases the number of groups increases, each group, however, containing fewer fringes. The position of the centre of the grouped system drifts in the same direction as the point of maximum visibility in the previous experiments.

A fuller treatment with a rigorous mathematical analysis by Lord Rayleigh will be found in an earlier edition.

Distribution of Phase over Small Area Illuminated by Source of Finite Size. — In Young's celebrated experiment, where interference was observed between the diffracted rays coming through two pin-holes illuminated by sunlight coming from another small hole, it is clear that the phase of the vibration must be the same at the two pin-holes. If the source of light were infinitely small, the phase relation between the vibrations passing through the two holes would be permanent, even if the holes were widely separated; but if the source has finite size, as is always the case, this will not be true. The reason for this is clear. The vibration at each hole is the resultant of the disturbances coming from the various points of the source, and this will vary with the position of the holes.

It can be shown (see earlier editions) that the diameter of the area over which we have constant phase is $R\lambda/2r$ in which $2r/R$ represents the apparent diameter of the luminous source (of actual radius r), as seen from the area at distance R, and we can therefore say in general that the phase can be considered constant over a circular area not greater in diameter than the wave-length of light divided by the apparent diameter of the source. In the case of sunlight $r/R = \tan 16' = .005$, and $\lambda = .0005$ mm. and the diameter of the area of constant phase $= .0005/.01 = .05$ mm.

In ordinary sunlight, therefore, the phase is the same over an area measuring only .05 mm. in diameter, or in a square millimetre there are 400 different states of vibration. We can easily apply our formula to Young's experiment. Suppose we form an

image of the sun with a lens of 5 mms. focus. Its diameter will be .05 mm. and from a distance of 1 metre ρ/R will be .00005.

The diameter of our circle of similar phase will be 5 mms., *i.e.* the two pin-holes should not be over 5 mms. apart if we are to regard them as similar sources, which is the condition which we must fulfil if we wish to obtain interference fringes.

Interference Fringes When the Sources Are in Line. — If we consider the hyperboloid fringes formed in space when waves radiate in all directions from two similar sources, to be cut by a plane perpendicular to the line joining the sources, the maxima and minima form circles which have a common centre on the prolongation of the line joining the sources.

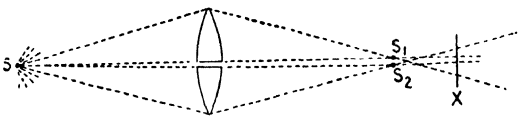


FIG. 118

Fringes of this description were obtained by Meslin [1] by an ingenious modification of the arrangement of the Billet split lens. This instrument consists of a double convex lens, cut in halves, the two portions being slightly separated, so as to form two images s_1, s_2 of the source, the fringes being observed in the plane X (Fig. 118).

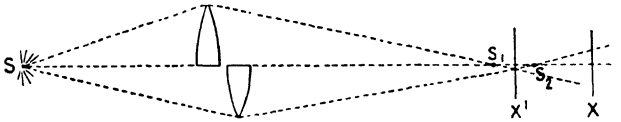


FIG. 119

By displacing one of the halves as shown in Fig. 119 the sources are brought into the line of sight. The fringes will be circles in this case, but they will not be found as before in the plane X, since the beams from the two sources do not overlap in this region, but in the plane X' *between* the sources, where overlapping occurs. This is in reality interference between waves radiating from a source, with waves converging to a similar source.

Interference Fringes along Caustics. — In the Chapter on Reflection we have seen that the caustic surfaces are traced by the cusped wave-fronts. Just within the caustic we therefore have two wave-fronts, which travel obliquely with respect to each other, and which came originally from the same source. Let c, c' be the caustic traced by the cusped waves, the crests of which are represented by solid lines, the troughs by dotted lines (Fig. 120). Along the line A we have crests meeting troughs, and a consequent destructive interference, while along the line B we have similar

[1] *Compt. Rendus*, 1893.

phases, and maximum illumination. It is clear from the diagram that the interference is between a portion of the wave which has passed through a focus with one which is converging to a focus, the condition being similar in some respects to experiments with the Billet split lens, described in the previous section. These fringes are easily seen by concentrating sunlight upon a pin-hole and reflecting the light from an oblique concave mirror. The fringes will be found in the region between the primary and secondary focal lines, *i.e.* along the caustic.

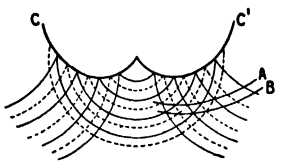


FIG. 120

The Colors of Thin Plates. — The iridescent colors which are displayed by thin films of transparent substances were first investigated by Boyle. In 1665 Hooke devised the method of producing the colors by means of an air film between two lenses of large radius of curvature. He found that the colors were distributed in concentric rings, showing that they depended on the thickness of the film, and that equal thickness gave always the same color. Hooke explained the production of color as follows. A portion of the light is reflected from the upper surface, and a portion penetrates the film and is reflected from the lower surface. This portion has been weakened, and Hooke's notion was that sensation of color depended on successive impacts on the retina of strong and weak impulses. If the stronger preceded the weaker one color was produced. If *vice versa*, then another color resulted. He was right in explaining the color as produced by the union of the light streams reflected from the two surfaces, and being ignorant of the nature of white light and of wave-length as we speak of it, gave what seemed the simplest and most probable explanation of the regular sequence of the colors.

The subject was more carefully investigated by Newton, who made careful measurements of the colored rings (since known by his name) produced by the air film between a lens and a plate of glass.

It remained for Young, however, to give the true explanation that the rings were due to the interference between the wave-trains reflected from the upper and lower surface of the film.

If we place the curved surface of a plano-convex lens of very long focus (the longer the better, say 2 or 3 metres) on a flat plate of glass, and view the reflected image of a sodium flame, we shall see the point of contact surrounded by dark and bright circles of light, produced by interference of the streams of light reflected from the two surfaces. This gives us a very simple means of obtaining inter-

ference under a variable path difference. The diameters of the circles vary according to the same law as that which obtains in the case of the zone-plate, the scale however varies with the wave-length, red light giving us larger rings than yellow or green.

If the source emits two wave-lengths, the bright rings will therefore be in-step at some points, and out-of-step at others. In the

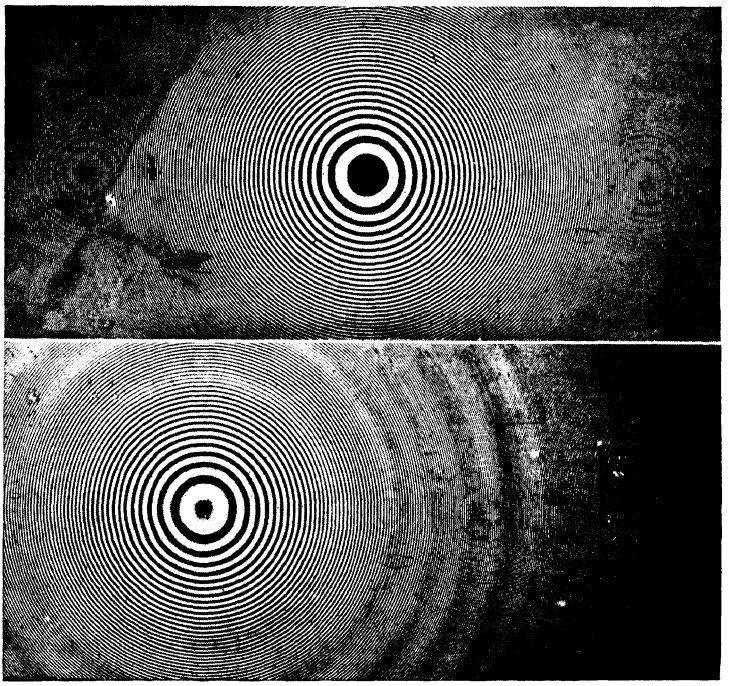


FIG. 121

case of sodium light we have two wave-lengths the difference between which is 1/1000 of their actual values. The fringes are therefore exactly out-of-step, or in "Dissonance" at the 500th ring from the centre. At the 1000th ring they are again in-step or in "Consonance." When in dissonance they are nearly invisible. The alternate appearance and disappearance of the sodium flame rings cannot be well seen with the lens and plate unless we use a strong magnifying-glass to view the fringes.

Their distinctness is much greater if we get rid of the light reflected from the upper surface of the lens, and the lower surface of the plate. This can be done by putting a horizontal slit .5 cm. in width in front of the soda flame. The images reflected from the

different surfaces will appear separated, the rings being seen only in one of them. An easier way of showing dissonance is to use a lithium and thallium flame. This gives red and green rings where dissonance occurs, and yellow (subjective) at the points where the red and green are superposed. If white light is employed we get colored rings.

Photographs of the rings produced by the green light of the mercury arc, and by the green and yellow lines showing dissonance are reproduced in Fig. 121.

Newton's Rings. — The thickness of a film of air at any point between a spherical and a plane surface in contact is easily expressed in terms of the distance of the point from the point of tangency, and the radius of curvature of the surface.

This gave Newton the means of determining accurately the color produced by an air film of any thickness. A lens, the radius of curvature of which is known, is placed on a piece of plate glass and viewed by reflected light. Circular colored rings are seen surrounding the point of contact, the colors being most brilliant where the air film is very thin. We wish to determine its thickness ϵ, for example, where the first yellow ring appears. Let the radius of curvature of the lens be R, and the radius of the yellow ring r (Fig. 122). We have

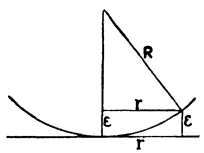


FIG. 122

$$r^2 = R^2 - (R-\epsilon)^2 = 2R\epsilon - \epsilon^2 \text{ or } \epsilon = \frac{r^2}{2R},$$

since ϵ^2 is small in comparison to $2R\epsilon$, an expression which shows us that the thickness of the air film is proportional to the square of the radius of the ring. Newton found that with monochromatic light he obtained bright and dark rings, and that the rings when produced by red light were larger than when produced by blue. With white light, then, we have an infinite number of ring systems superposed, and to the blending of these systems is due the complicated succession of colors observed by Newton.

It may be remarked here that in viewing the interference phenomena produced by thin plates, the eye must be focussed on the film.

Newton's Black Spot. — Interference by reflection from a thin film is illustrated in Fig. 123. The ray AB is refracted at B and reflected at C, emerging at D in the direction of reflection of the ray $A'D$, and interference will occur. At normal incidence AB and $A'D$ will coincide, and the path-difference will be twice the thickness of the film. For a path-difference small in comparison to λ we might expect strong reflection, but Newton found that

in this case the reflection was zero, the film showing a black spot. Such spots can be seen on thin soap films and with the "Rain-Bow Cup" of C. V. Boys (supplied by J. J. Griffin, Kingsway, London) in which the soap film is rotated in its own plane and thinned by centrifugal force, concentric rings of the most brilliant colors forming with a circular black spot at the centre an inch or more in diameter. On reversing the rotation, the colors form wonderful patterns, winding up in spirals and designs similar to those on Persian shawls. The black spot under certain conditions shows different degrees of blackness, the areas being sharply bounded. An interesting paper, illustrated by color photographs has recently been published by André Marcelin [1] which will be referred to presently. The black spot is due to the fact that the two reflections take place under different conditions. At the upper surface the reflection is from a rare medium to a dense, at the lower surface, from a dense to a rare. The waves reflected at the rarer medium are reflected without change of phase, those reflected at the denser medium suffer a phase change of 180°. This is, of course, equivalent to a path-difference of half a wave-length.

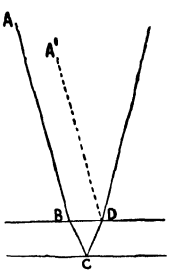


FIG. 123

This explanation was given by Young, who devised a very beautiful experiment in support of it. By using a lens of crown and a plate of flint glass with a film of cassia oil between them, he secured a system in which reflection from the upper and lower surface of the film took place under the same condition, the oil having a refractive index intermediate between that of the crown and flint glass. The ring system formed under these conditions had a white centre, exactly in accordance with his theory.

Under normal incidence we have then the following equations for the thickness e of the film:

$$\left.\begin{aligned} e &= (2n-1)\frac{\lambda}{4} \text{ for a maximum} \\ e &= 2(n-1)\frac{\lambda}{4} \text{ for a minimum} \end{aligned}\right\} \text{by reflection,}$$

where the reflection takes place under opposite conditions, and

$$\left.\begin{aligned} e &= 2n\frac{\lambda}{4} \text{ for a maximum} \\ e &= (2n+1)\frac{\lambda}{4} \text{ for a minimum} \end{aligned}\right\} \text{by reflection,}$$

[1] *J. de Chemie Physique*, Dec., 1931.

where the conditions are the same as in the case of the cassia oil experiment. In these equations λ is the wave-length in the film. If the film has a refractive index μ, and λ is the wave-length of light in vacuum, we must write λ/μ for λ in the equations.

An inspection of the formula for the maxima $e=(2n-1)(\lambda/4)$ shows us that with a given thickness, large in comparison to the wave-length, the formula will hold for a number of different values of λ, n taking different values as we change λ. Thus a given thickness may fulfil the conditions of the equation for maxima for a large number of different colors. The analogy between this case and that of the fringes produced by the Fresnel mirrors with large path-difference is obvious. Suppose the thickness of the film to be .01 m., we then have $.01=n(\lambda/4)$ or $.04=n\lambda$ for a maximum when the value of n is any odd number. If we give λ its value for red, .0007, we find n to be 57, or we have $n=57$ for red, corresponding to the 28th maximum.

With violet light of wave-length .0003 we find $n=133$, corresponding to the 66th maximum. Between these two values we shall have 66.28 other maxima for intermediate wave-lengths. Consequently a film measuring .01 mm. in thickness will reflect 38 different parts of the spectrum and refuse to reflect 38 intermediate parts, or if we examine the light reflected from the film with the spectroscope we shall find the spectrum crossed by 38 dark bands.

Refractive Index and Dispersion of a Thin Plate. — It is plain that we have here a means of determining the thickness of a thin film. By examining the light reflected from it with a spectroscope and counting the number of dark bands between any two known points (Fraunhofer lines) in the spectrum, we can, by making the substitutions in the formula, calculate the thickness. In the formula which we have given we have supposed the incidence normal, and λ to be the value of the wave-length in the material of the film. If we are dealing with films of glass we must, of course, reduce the wave-length values to their values in glass.

The complete formula for determining the thickness of a plate of any substance with light reflected at any incidence is

$$e=\frac{n\lambda_1\lambda_2}{2\mu\cos r(\lambda_1-\lambda_2)},$$

in which n is the number of dark bands between wave-lengths λ_1, λ_2, μ the refractive index of the film, and r the angle of refraction.

If we know the thickness we can determine μ at different points of the spectrum, or measure the dispersion.

Influence of Multiple Reflections. — The theory of thin plates as it came from the hands of Young had an imperfection. The por-

tions of the light reflected from the two surfaces are not equal, since the light which suffers reflection at the second surface has already been weakened by reflection at the first. The two portions should therefore never wholly destroy each other as they do when we employ monochromatic light. Poisson showed that we must take into account the multiple reflections which occur within the film. If the retardation of the ray $A'B'C'B$ on the ray AB is $\delta = 2e \cos r$, the retardation of consecutive rays incident at B'' B''', etc., are $2\delta, 3\delta$, etc. (Fig. 124). We thus know the phases of the components as they arrive at B, and to calculate their joint effects we must

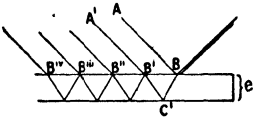


FIG. 124

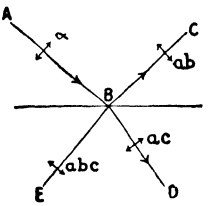


FIG. 125

know their amplitudes. A certain percentage of the incident light will be reflected at the glass-air surface, and a certain percentage at the air-glass surface, but we have no right to assume that the fractional part reflected is the same in each case. The following method of ascertaining the relations between the amplitudes of the reflected rays was used by Stokes.

Let the amplitude of the incident ray be a (Fig. 125), then the amplitude of the reflected ray will be ab, in which b is a fraction, and the amplitude of the refracted ray will be ac, c being in general a fraction larger than b. By the principle of reversibility, if we send these two rays back along their paths, they should give rise to the original ray, reversed in direction of course, with the original amplitude a. If we reverse BC, however, it will give rise to two rays, one along BA of amplitude ab^2, and one along BE of amplitude abc. In reversing DB we cannot obtain the amplitudes of the reflected and refracted components by multiplying its amplitude by b and c respectively, since the reflection takes place under different conditions. We will therefore designate the amplitudes of the reversed components of BD by acf (along BA) and ace (along BE). If the sum of these components is to represent a ray along BA of amplitude a, and we are to have no ray BE, as must be the case if the reversed rays give rise to the original ray only. the following relations between b, c, e and f must hold:

$$acf + ab^2 = a \text{ and } ace + abc = 0.$$

These equations give us

$$cf = 1 - b^2 = 1 - e^2 \text{ and } b = -e.$$

The latter equation shows us that the amplitude of the ray arising from reflection in passing from the upper to the lower medium is equal to the amplitude of a ray of equal intensity which has suffered reflection in passing from the lower to the upper medium. The fact that the sign of b is opposite to the sign of e indicates moreover that there is a relative phase change of half a period between the ray reflected under opposite conditions. This explains the absence of a ray along BE when we reverse the reflected and refracted components, the components along BE having a phase difference of 180 and destroying one another.

The perfect blackness of the interference fringes when monochromatic light is used follows at once from the above equations. The amplitude of the stream reflected from the first surface is ab. The transmitted amplitude is ac, of which abc is reflected from the lower surface, and $abcf$ emerges into the upper medium. The amplitudes emerging into the upper medium owing to the multiple reflections form a series $abcf + ab^3cf + ab^5cf + \cdots$. Complete interference will occur if the sum is equal to ab. This is seen to be the case, for

$$abcf(1 + b^2 + b^4, \text{ etc.}) = abcf \cdot \frac{1 - b^{2n}}{1 - b^2} = ab \frac{cf}{1 - b^2} = ab,$$

since $cf = 1 - b^2$, as we have seen above.

Silvered Films. — A remarkable conclusion is reached if we consider that the two surfaces have a very high reflecting power, as when they are coated with a thin film of silver. Suppose that the metal films are of such a thickness that they reflect 80% of the light and transmit 20%. This condition can never be realized in practice owing to absorption by the silver, but there is nothing to prevent our assuming the possibility of a film with high reflecting power and zero absorption. If the plate had the right thickness before silvering to give zero reflection (owing to interference) it will also show zero reflecting power after the silvering. This is of great importance in connection with a remarkable instrument which will be discussed later on, the interferometer of Fabry and Perot.

At first sight there seems to be a paradox here for if the plate has zero reflecting power (by interference), it must give complete transmission, provided there is no loss by absorption. This can mean only that if a reflecting film, which is so opaque that but 20% of the incident light gets through, is backed by a second film

of equal opacity at the right distance, the combination becomes perfectly transparent, instead of twice as opaque as experience with opaque media such as smoked glass might lead us to expect.

Let us see how this surprising behavior comes about. To apply the equations which have just been developed we must translate our intensities into amplitudes which are the square roots of the intensities. Suppose the intensity of the original light is 100, the reflected intensity will then be 80 and the transmitted 20. The corresponding amplitudes will now be 10, 9 and 4.5 (very nearly).

In our equations a will now be 10, $b=.9$ and $c=f=.45$ and $cf=.2025$ while $1-b^2=.19$. We get our series of emergent rays by multiplying acf or 2.025 by b, b^3, b^5, etc. The first three terms, or the amplitudes of the first three emergent rays are 1.82, 1.48 and 1.2 which add up to 4.5 or one-half of the required amount to destroy the reflected amplitude of 9.

Putting this in the form of a diagram, we see at once how sufficient energy is thrown up against the upper film to destroy its original high reflecting power by interference. Let the incident ray be AB, of amplitude 10 (Fig. 126). This gives a reflected ray BR of amplitude 9. A neighboring ray, A', belonging to the same wave-front, after one reflection from the lower silver film emerges at B also in the direction BR with amplitude 1.82, the ray A'' after three reflections emerges at B with amplitude 1.48, the ray A'' after five reflections emerges with amplitude 1.2 and so on for other rays $'''$, etc., until the sum of all the amplitudes equals 9, which is sufficient to destroy the ray BR. We thus see that when the two films have a high reflecting power, the comparatively small amount of energy which penetrates the first film is beaten back and forth between the film, and we have available a large number of emergent rays at any point, all in the right phase to interfere destructively with a ray incident upon the upper film at the same point.

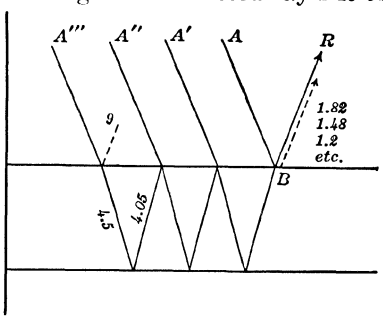


FIG. 126

We will now determine the path-difference between the rays reflected from the upper and lower surfaces.

Interference with Thick Plates. — Consider the plane-wave AB incident on a thick plate (Fig. 127). The ray BS will be reflected to P, while the ray $B'H$ will, after refraction by the plate and reflection from E to S, emerge into the air and follow the path

SP. We require the path-difference between these two rays at S. Draw HK and SR perpendicular to the incident and refracted rays. The phase at R and S will then be the same, and the path-difference required will be $(RE+ES)\mu$ where μ is the refractive index of the plate. Extend RE until it cuts at F the perpendicular dropped from S, and $SF=2t$, t being the thickness of the plate. The angle at F (or the angle of refraction r) we will designate as r. We now have $RF=RE+ES$, and for the path-difference μRF, or $2\ \mu t\cos r$, since $RF/2t=\cos r$. To this we must add a half wave-length, since there is a phase-change of 180 degrees as a result of the reflections taking place from air to glass and glass to air respectively. The reflected intensity will be a minimum whenever $2\ \mu t\cos r+\lambda/2=m\lambda/2$, m being an odd number, or transposing $2\ \mu t\cos r=n\lambda$, n being any integer.

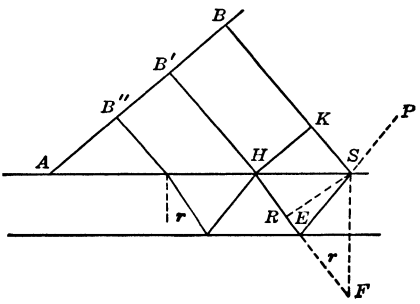


FIG. 127

The path-difference is a maximum for normal incidence and decreases with increasing inclination of the rays. This also is of importance in connection with the Fabry and Perot interferometer, which will be considered in the Chapter on Interference Spectroscopes.

Testing Glass Plates for Flatness and Plane-Parallelism. — If we have an optical glass flat, we can test the flatness of any other plate by laying it upon the flat and observing the reflection of a sodium flame from the surfaces. Flatness is indicated by perfectly straight fringes, parallel and equidistant. If we have no flat plate we can use a method due to Lord Rayleigh. The glass plate is mounted in a small dish, which is then filled with clean water until the fluid covers the plate to the depth of about a millimetre. There should be a margin of surface of 2 or 3 cms. between the edges of the plate and the sides of the dish, to avoid capillary troubles. The whole is mounted upon a stand provided with levelling screws, in a locality free from all tremors. A glass plate should be placed over the dish to prevent air currents from disturbing the surface. The fringes formed by the interference between rays reflected from the glass-water surface and the water-air surface are observed. The two surfaces should first be made as nearly parallel as possible, by observing the reflection of a very small and bright point of light, and bringing the two images together by means of the levelling screws. A small mirror should

be mounted over the surface of the water which reflects sodium light down upon the surface. The surface should be observed by means of a magnifying-glass. At first the fringes will usually be found very close together, but they may be made as broad as desired by levelling. If the water surface is of a certain thickness, they may not appear owing to the fringes formed by D_1 and D_2 being out-of-step or in dissonance. It is safer to use the light of the mercury arc filtered through a green glass, for the green line is very homogeneous. If the fringes are curved the plate is not flat. We must now determine the nature of the surface, that is whether convex or concave, spherical, cylindrical or saddle-shaped.

If we move our eye so as to increase the angle of incidence, the retardation will decrease, since cos i decreases in the formula $2\,\mu e \cos i = n\lambda$ for a fringe of order n which remains constant for a given fringe. The thickness e must therefore increase, and a given fringe move towards a region of greater thickness.

If the surface is convex the fringes will move out towards the edge as the eye is moved down, if concave they will move away from the edge.

It is however so much easier to work with a glass flat, that the water method should be attempted only to standardize the test plate. Suppose we have a small square plate which we know to be flat.

If we place a larger piece of thin plate glass upon it, and press down upon the four corners with the fingers, the under surface will become concave, and we see Newton's rings, which will close in towards the centre as the eye is moved down.

If the surface is cylindrical, and we can bring the plate into this condition by pressing down upon two opposite edges, the fringes will be straight, but not equidistant, lying closer together along the two edges than at the surface.

A saddle-shaped surface gives fringes shaped like hyperbolae. The thickness along a given fringe is a constant, and the fringes can therefore be regarded in the same way as the contour lines on a map.

It is instructive to try the experiment of improving or figuring a poor surface by means of dilute hydrofluoric acid applied with a soft brush over the regions from which material must be removed. The surface should be tested frequently by laying it upon the test plate. A subsequent polishing with rouge upon a surface of pitch will remove the slight irregularities introduced by the etching process.

To test for plane-parallelism we observe the Haidinger fringes obtained with thick plates when the eye is focussed for infinity.

Mount the plate in a horizontal position and reflect the green light from the mercury arc down upon it from above by means of a piece of plate glass at an angle of 45°. The plate must first be made perpendicular to the line of vision (from above) by observing the reflection of the pupil of the eye in it. The fringes appear as concentric circles when the eye is focussed for infinity. The direction of the incident light and the line of vision should coincide as nearly as possible. Fix the attention upon the central ring, and move the plate slowly in its own plane. If the rings expand, new ones opening out from the centre, we are moving towards a region of greater thickness, for the incidence angle i is greater for each successive ring and the formula for a ring of order n is

$$n\lambda = 2\,\mu e \cos i.$$

As i increases cos i decreases, therefore e must increase, that is the ring moves out. To determine the error, count the rings which develop in passing from edge to edge of the plate. If k represents this number, then $k/2$ is the number of wave-lengths by which the plate departs from parallelism, and $(k/2\,\mu)\lambda$ is the actual difference in thickness in millimetres. ($\lambda = .000546$ for the green mercury light.)

Colors of Iridescent Crystals and Opals. — Some very remarkable phenomena connected with the colors of thin films are frequently exhibited in crystals of chlorate of potash. The cause of these colors was investigated by Stokes, and found to be due to the existence of planes within the crystal at which a periodic "twinning" had occurred. The colors are extremely brilliant and pure, much more so than any exhibited by soap-films. An interesting paper by Lord Rayleigh will be found in the *Phil. Mag.*, xxvi, pp. 256–265, 1888. One of the most remarkable facts connected with the phenomenon is that the spectrum of the reflected light is frequently found to consist of a comparatively narrow band. One in possession of the author at a certain angle of incidence reflects yellow light which, when examined in the spectroscope, is found to consist of a narrow band not much wider than the distance between the yellow mercury lines. In the case of a *single* thin film, of such thickness that but a single region of the spectrum is reflected, this region is always of considerable breadth. To account for the reflection of light of such a high degree of purity Lord Rayleigh assumes that the reflection takes place at a number of thin laminae sensibly equidistant, the distance between any two being of the order of magnitude of the light-wave. Quoting from his paper, "In order to explain the vigor and purity of the color reflected in certain crystals it is necessary to suppose that there are a consid-

erable number of thin surfaces disposed at approximate equal intervals. At each angle of incidence there would be a particular wave-length for which the phases of the several reflections are in agreement. The selection of light of a particular wave-length would thus take place upon the same principle as in diffraction spectra, and might reach a high degree of perfection." Lord Rayleigh describes an interesting acoustical analogue, the sound of a bird call, giving a pure tone of high pitch, being most copiously reflected from a number of flat equidistant screens made of thin muslin, stretched upon brass rings at a certain distance apart.[1] The pure colors of some opals are due to the same cause.

This remarkable limitation of the reflected light to a narrow region of the spectrum will be better understood after a study of the diffraction grating. It will be interesting to compare this action of multiple films with the action of the Fabry and Perot interferometer, which will be described in a subsequent chapter. The colored crystals of chlorate of potash are easily prepared by making a hot saturated solution of the potash, and allowing it to cool slowly in a large flat-bottomed tray. On breaking up the crystalline mass, and shaking the tray gently in front of a window, numerous highly colored flakes will appear, which can be lifted out by means of a small bent spatula, made of thin sheet copper or brass. They should be dried on filter-paper, and mounted in balsam, preferably between two glass prisms of about ten degrees angle. This prevents the dilution of the color by white light reflected from the surfaces of the crystal plate.

As we have seen, a plate several wave-lengths thick shows maximum reflection for a number of spectral regions, the number increasing with the thickness. This is the explanation of the fact that the crystals usually show several narrow bright bands by reflection. Their optical properties in the ultra-violet were examined by the author [2] and maxima of higher orders found in the expected positions. The colored flakes were mounted in glycerine between two quartz prisms of small angle, and the light of a cadmium spark reflected from them into a quartz spectrograph.

The position of the maxima in the case of certain crystals indicated that the first order color was in the infra-red region, and a flake showing this peculiarity, for which the calculated position of the first order band was 1.2 μ was examined, at the request of the author by A. H. Pfund.

The radiometer gave a deflection of several centimetres, when illuminated with the red band, immediately dropping back to zero

[1] *Nature*, xl, p. 227, 1889.
[2] *Phil. Mag.*, July, 1906.

as the prism was so turned as to bring the infra-red region upon the slit. At about 1 μ a large deflection was obtained, which at 1.2 μ was "off the scale." The energy curve was then plotted from a large number of readings, the slit having been narrowed in the meantime, and a strong maximum found at 1.25 μ.

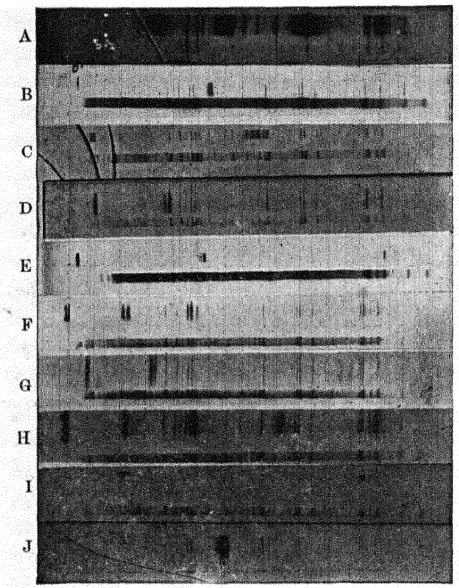


FIG. 128

The series of photographs obtained in the ultra-violet is reproduced in Fig. 128. In some cases a number of bands appear in the ultra-violet, and in other cases only one or two. It is obvious that the thicker the reflecting strata, the larger will be the number of regions selectively reflected in a given spectral range. The approximate values of the lengths of the reflected waves are given in the following table:

	1st Order Color	Higher Order Colors					
A	1.6 μ	.80	.54	.40	.30	.26	.23
B	.617	.31					
C	.53	.28					
D	1 μ	.53	.28				
E	.64	.325	.22				
F	1.26 μ	.663	.435	.32	.23		
G	1.13	.565	.39	.235			
H	1.23	.614	.42	.326	.266	.226	
I	.44	.23					
J	.60	.30					

One flake, measuring about 6 mms. on a side, exhibited total reflection at normal incidence of a region of the spectrum only 10 or 12 Ångström units in width, that is, only double the distance between the *D* lines. The spectrum of the transmitted light exhibited a very black band at the same point, and of the same width. This band was photographed, after having been brought into the vicinity of the *D* lines by suitable inclination of the plate, and the *D* lines

themselves impressed on the plate by holding a sodium flame in front of the slit for a few seconds. The photograph is reproduced in Fig. 129, spectrum "*c*." The position of the *D* lines is marked on the spectrum immediately above this one. (The marks refer to the lower spectrum only.)

By increasing the angle of incidence, the band can be made to move down the spectrum, widening as it moves. When in the green it appears as in spectrum "*b*" and is accompanied by fainter lateral minima. The narrowness of the reflected region has been shown to be due to multiple twin planes, sensibly equidistant. From the width of the reflected region we can form an estimate of the number of laminae present in the crystal plate. Each lamina gives us by reflection a virtual image of the source, these images being in line, one behind the other, at normal incidence. The action is not unlike that of a diffraction grating when the diffracted ray is at grazing emergence. If we are dealing with a first-order spectrum, 1000 lines are necessary to resolve the *D* lines.

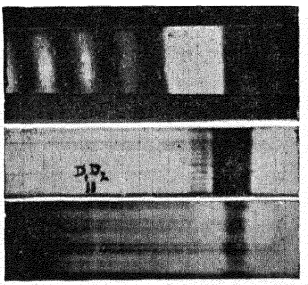


FIG. 129

If now we compare the width of the band in the photograph with the distance between the *D* lines, it is clear that the crystal plate is very nearly able to separate or resolve the *D* lines. In other words, if we incline the plate a little less, causing the band to move up the spectrum, it will reflect D_2 before it reflects D_1. It was, however, not quite able to do this, but would easily separate lines of twice the separation of the *D*'s. From this we may infer that the number of twin planes in the crystal is somewhere between 500 and 1000, say roughly 700. If we are dealing with a first-order spectrum the path-difference between rays reflected from two adjacent twin planes must be equal to the wave-length of the light in the crystal. Assuming no phase-change this makes the thickness of each lamina about .0002 mm., and multiplying this by 700 gives us .14 mm. as the thickness of the crystal plate, which was not very far from the truth.

It is interesting to compare this crystal with the best Lippmann photographs, which show color resulting from the same type of interference.

H. E. Ives has made remarkably brilliant Lippmann photographs with monochromatic light and the spectrum of the band reflected from his best plate is shown, Fig. 129, "*a*," on the same scale of

wave-lengths, *i.e.* taken with the same spectroscope. It has a width about three times as great as the band reflected from the crystal. Ives sectioned a similar film and counted about 250 silver laminae, built up by the stationary waves. This is in good agreement with our estimate of the number in the chlorate crystal.

Effect of a Prism upon Newton's Rings. — It was observed by Newton that, when the colored ring system formed by a lens and a flat plate was viewed through a prism, the number of rings visible was greatly increased on one side of the system, the increase being about twelvefold. If the fringes were equidistant, as is the case with the Fresnel mirrors or the bi-prism, a prismatic shift would bring the fringes of different colors into step at a point far out in the system, but would throw them out-of-step at the centre, so that their appearance would not be much altered. The rings, however, become narrower as we advance out into the system, and if we simplify the problem by supposing that we have only red and blue fringes which are shifted through different distances by the prism, it is not difficult to see how the achromatization results, for the blue arcs, from a portion near the centre, can be made to fit approximately over the red arcs in a more remote region. Now the blue rings are shifted more than the red, consequently the achromatization will occur on the side of the centre towards which the shifts have taken place. A full treatment of the subject will be found in Lord Rayleigh's [1] paper on achromatic interference bands. Fringes can often be found by this means on thin glass bulbs, easily made by blowing out a glass tube; sodium light will give fringes without the prism, but nothing can be seen with white light owing to the thickness of the glass.

Achromatization of the Fringes Formed by a Thin Reflecting Lamina. — An arrangement was devised by Talbot which yielded achromatic fringes of equal widths. The achromatization which we have just considered depends upon the different widths in the different parts of the system, and cannot be applied to the equidistant fringes obtained with a wedge-shaped film.

To obtain achromatization in this case it is necessary to arrange matters so that the scale of the system is the same for the different colors. Now the scale depends on the angle of the wedge (which is obviously fixed) and the angle of refraction. Under ordinary circumstances the angle of refraction is very nearly the same for the different colors, but if we employ an air film between glass plates, with the light incident in such a direction that the angle of refraction is nearly 90°, owing to the powerful dispersion the angle will vary with λ, and since the angle is greater for the blue than for the

[1] *Phil. Mag.*, 1889.

red, the blue fringes may be formed on the same scale as the red, and nearly perfect achromatization result.

To obtain a sufficiently large angle of refraction it is necessary to employ a prism of the form shown in Fig. 130. A right-angle prism answers nearly as well. It should be placed with its hypotenuse on a silvered glass plate or a plate of polished speculum metal. A horizontal slit illuminated with white light is placed in such a position that the light is incident on the second surface at nearly the critical angle. This position can be found by lifting one end of the prism a little and watching the slit image, varying the height of the slit until a number of images appear side by side. On lowering the prism these images will run together. They are virtual images

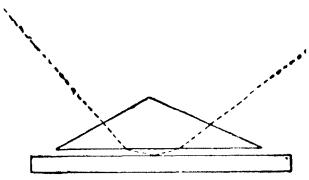


FIG. 130

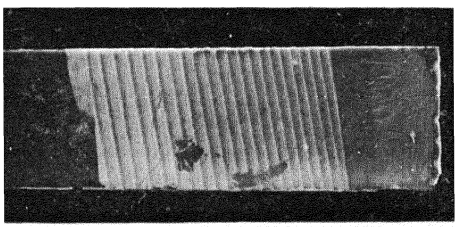


FIG. 131

formed by multiple reflections between the prism face and the metal plate. The colored fringes can now be found with an eye-piece, and by tilting up the edge of the prism which is nearest to the slit it is usually possible to achromatize them at once, 50 or 60 black and white bands appearing in the field. When the prism rests on the plate, the fringes are broad and highly colored. An eye accustomed to the appearance of fringe systems near the central fringe will recognize that there is something peculiar in this case. The appearance is due to the fact that we have a number of virtual sources in line one behind the other. If the slit is illuminated with sodium light, the fringes will present a remarkable appearance. The maxima will appear bright and narrow, with broad minima between them. On *one* side of each maximum a number of fainter maxima will be seen, which gives a corrugated or shaded appearance to the system. A photograph of the system can be obtained by laying an orthochromatic plate on the face of the prism, and exposing it for a couple of hours, taking care to shield it from all light except that which comes through the slit. A picture obtained in this way is reproduced in Fig. 131.

Preparation of Films for the Exhibition of Newton's Colors.—In the case of a thin transparent lamina, such as a soap-film, the amplitudes of the disturbances reflected from the two surfaces are equal,

and consequently completely destroy each other when the phase-difference is 180°. Inasmuch, however, as only a small percentage of light is reflected from each surface, the colors, though saturated, are not as intense as is desirable. If a plate of mica is pressed against a pool of molten selenium on a glass plate, and the whole allowed to cool under pressure, on stripping off the mica, films of mica of variable thickness will be left upon the surface of the selenium, which show Newton's colors of great beauty, arranged in mosaics. The patches of equal thickness being sharply bounded by straight lines, present an appearance similar to that of selenite films under the polariscope. The selenium has a much higher refractive index than the mica, consequently the reflection at each surface is the reflection of rays incident from a rare to a denser medium, and the difference of phase is given by the difference of path alone; *i.e.* we do not have the loss of half a wave-length due to reflection under opposite conditions, as would be the case if the mica films were in air. [1]

Still more brilliant films can be prepared by first thickly silvering the mica, cementing the silvered side to a glass plate with a beeswax rosin cement, and then stripping off the mica. The colors are scarcely visible, owing to the disproportionality between the amplitudes of the two interfering streams of light, but appear as soon as the upper surface of the mica is half-silvered, which can be done by immersing the plate in a silvering solution until the colors reach their maximum brilliancy.

The colors are of great purity and intensity and the fact that the areas of the mosaic are bounded by straight lines, adds to the beauty of the colored pattern. Gradual gradations of color are obtained by an alternative method.

A sheet of glass can be silvered chemically, or procured by removing the varnish from the back of a piece of modern mirror-glass with alcohol. The silver film is then flowed with collodion diluted with three or four parts of ether. As soon as the film dries colors appear, contrary to theory. These colors may be quite brilliant, and are due to diffraction, as will be shown presently. If the plate be now immersed in Brashear's silvering-bath, the colors will instantly disappear, owing to the fact that the collodion film and the solution have nearly the same refractive index. As soon as the silver begins to deposit, the colors reappear and increase rapidly in intensity. The bath should be rocked, the process being similar to the development of a negative. A little experience will enable the moment of maximum brilliancy to be correctly judged, when the plate should be removed from the solution, washed and dried.

[1] R. W. Wood, *Phil. Mag.*, April, 1904.

A cover-glass, sloping at an angle of say 20° to avoid its reflection is provided to exclude dust, and side plates attached with strips of lantern slide binders, the whole forming a prismatic box. These plates make wonderfully beautiful objects for reflectoscope lantern projection.

In repeating the experiment with the mica film on selenium, described above, René Marcelin conceived the brilliant idea of determining the thicknesses of the various colored areas with the idea of finding the smallest difference in thickness that could occur, or in other words to determine the thickness of what may be termed the elementary strata of which the crystal is built up. He used a colorimetric method, matching the color of the patches in the mica mosaic by sliding a quartz wedge of small angle across a slit mounted between two Nicols.

If the thickness of the mica is e and its refractive index n the path-difference of the interfering rays is

$$\delta = 2ne$$

since the phase-change is the same at each surface owing to the high refractive index of the selenium.

For the quartz wedge, with indices n' and n'' (Ord and Extra-ord) and thickness E the path-difference corresponding to the same color as that shown by the mica

$$\delta' = E(n' - n''),$$

or

$$\frac{E}{e} = \frac{2n}{n' - n''}.$$

Since for the quartz $n' - n'' = .009$ and for mica $n = 1.6$ it follows that $E/e = 355$, or that the variation Δe in the mica thickness which determines a certain color change, corresponds to a variation 355 times greater in the thickness ΔE of the wedge, which quantity is measurable. He made a very large number of measurements and found that the quantity was in every case a multiple of 7 Ångström units (.0000007 mm.).

His values are given in the following table in which Δe values are measured differences of thickness and N the corresponding number of unit strata of 7 A.U. thickness.

In A.U.	N	In A.U.	N
8.4	1	42.4	6
12.2	2	71.1	10
20.8	3	90.9	13
28.4	4	98.3	14
34.1	5	111.1	16

The theoretical diameter of the molecule of mica he gives as between 7.2 and 6.6 A.U.

These experiments showed that the cleavage of mica could be pushed to molecular dimensions, and that the molecular diameter could be determined optically. An autochrome photograph of a highly magnified portion of a mica on selenium plate, which accompanies a paper by his brother André Marcelin [1] is reproduced. Plate 4, Fig. 1 (facing this page).

He had previously made similar observations and measurements on soap-films which under special conditions exhibit sharply marked areas of uniform color similar to the mica mosaics, only bounded by arcs of circles instead of straight lines. These he found to be integral multiples of a unit stratum of 43.5 A.U., obtaining for Δe the numbers 43.5, 94, 138, 188, 222, 276, the first being the mean of 120 measurements, the others the mean of 40 each.

We have thus a species of liquid crystalization, and Marcelin showed that the thickness of the elementary stratum is double the length of the molecules of oleic acid deduced from the study of surface films of the acid on water. The condition necessary for the formation of these sharply bounded areas of uniform thickness is the presence of a small amount of fluorescent material in the soap solution combined with the presence of light of suitable wave-length for the excitation of the fluorescence. An autochrome of these soap-film mosaics from the same paper is also reproduced.

In the same paper are reproduced photographs showing the Newton black spot of eight different grades of blackness.

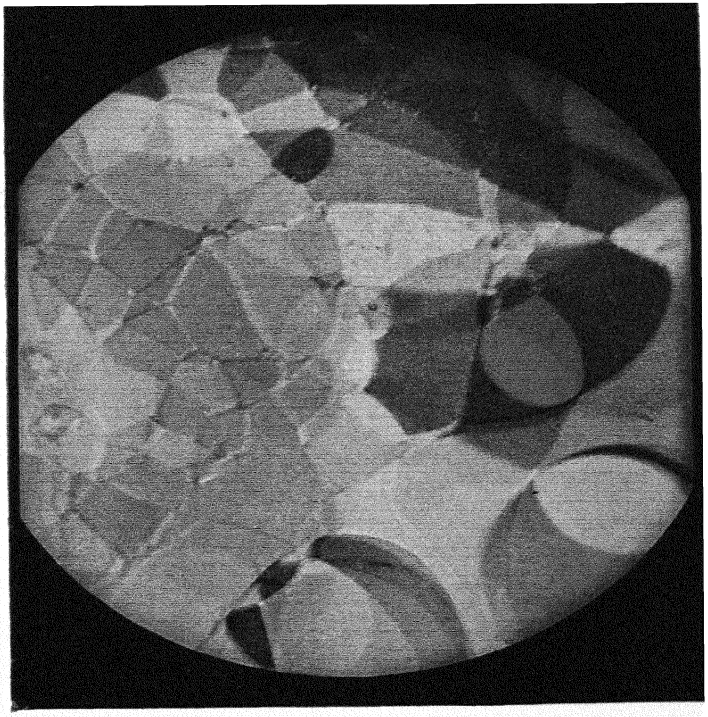
Colors of Frilled Transparent Films on Metallic Surfaces. — Some observations were published by the author [2] on a curious type of interference-colors of a different nature. As Lord Rayleigh points out in his article on "Wave-Theory of Light," a transparent film on a perfectly reflecting surface shows no interference-colors. It was found, however, that a thin film of collodion deposited on a bright surface of silver shows brilliant colors in reflected light. It, moreover, *scatters* light of a color complementary to the color of the directly reflected light. This is due to the fact that the collodion film "frills," the mesh, however, being so small that it can be detected only with the highest powers of the microscope. The collodion should be diluted with five or six parts of ether to which a little alcohol has been added. The addition of a drop or two of water may help the frilling. If chemically pure ether obtained by distillation is used, the film does not frill, and no trace of color is exhibited. Still more remarkable is the fact

[1] *J. de Chemie Physique*, 1931.
[2] *Phil. Mag.*, 1904.



Autochrome by M. Gilbert

FIG. 1. Interference-Colors of Mica on Selenium.



Autochrome by Mlle. Choucroun

FIG. 2. Interference-Colors of Stratified Soap-Film.

PLATE 4. ILLUSTRATING EXPERIMENTS BY MARCELIN

that if sunlight is thrown down upon the plate at normal incidence, brilliant colors are seen at grazing emergence, if a Nicol prism is held before the eye. These colors change to their complementary tints if the Nicol is rotated through 90°, *i.e.* in the scattered light, one-half of the spectrum is polarized in one plane, and the remainder in a plane perpendicular to it.

In the cases of the transparent films with the first surface lightly silvered, the second heavily coated, the waves absent in the reflected light are absorbed by the metal. In the present case these waves are scattered by the granular surface. If a spot on the film which appears purple by reflected light is illuminated with sunlight, it will be found that green light is scattered, not in all directions, but through a range corresponding to the size of the granulation, as in the case of the mixed plates, described in the Chapter on Diffraction.

If the light is incident normally, the scattered light comes off through an angular range included between 10° and 30°, and again at an angle of nearly 90°, the latter being strongly polarized. Conversely, if the sunlight be incident at nearly 90°, strongly polarized light is scattered normally. Considerable difficulty has been found in explaining these colors satisfactorily. They appear to be saturated, *i.e.* certain wave-lengths are completely absent in the reflected light, and until the granulation was detected with the microscope it was impossible to make even a satisfactory hypothesis. Even now the polarization effects are difficult to account for.

Tentative explanations and a fuller description of the experiments will be found in the original paper or in earlier editions of this book.

Curves of Equal Thickness and Equal Inclination. Haidinger's Fringes. — In the case of fringes formed by the reflection of light from a thin film of variable thickness, the thickness of the film along any fringe is a constant. These fringes are therefore called "curves of equal thickness."

The fringes are located at the film, and the eye should be focussed on the film to see them distinctly.

Another class of fringes was first described by Haidinger, in 1849, and subsequently studied by Michelson in 1882 and by Lummer in 1884.

They are produced by interference of light reflected from or transmitted by the two surfaces of a thick plane-parallel transparent plate, when waves of different inclinations fall upon it. These fringes are not located in the plane of the plate, as in the case of thin films, but at infinity, and to see them it is necessary to focus the eye for parallel rays, or employ a telescope previously

focussed for a distant object. This type of interference bands is very important in connection with the interferometers of Fabry and Perot, and Lummer and Gehrcke, which will be discussed in a later chapter. The manner in which the fringes are formed is not quite as easy to follow as in the previous cases, and we will discuss those formed by transmitted light as the diagram is less complicated in this case.

It is clear that in the case of thin films each portion of an interference maximum is seen by the light coming from a definite region of the source.

In the case of the Haidinger fringes, however, the complete ring system is formed by a single luminous point of the source, and this ring system is in coincidence with the rings formed by all of the other luminous points.

In Fig. 132 let B, C, D represent three luminous points of the extended light source, a sodium flame for example. P is a thick plane-parallel plate, L a lens with a focal plane for parallel rays at F, or simply the lens of the eye. Consider first the rays from B, C and D which are incident normally on the plate: they are parallel and the lens will, therefore, bring them together at the focus A, which will be the centre of the ring system. The multiple reflections within the plate of these three rays will also contribute to the illumination of A which will be a maximum or minimum according to the path-difference corresponding to twice the thickness of the plate. It must be clearly understood that the interference is only between the directly transmitted ray and the twice-reflected ray from each source; there is no interference between the rays from source B and source C, of course. Next consider the rays BB' and DD', which are also parallel, but which will be brought together at A'' since they slope to the right. Similar rays sloping to the left will be united at A'. One of these is shown for the source B. The twice-reflected ray is shown for this ray and the ray BB', and the emergent rays, being in each case parallel to the directly transmitted rays, are brought to the same focus points A'' and A'. It is thus clear that if we take a *cone* of rays of the same

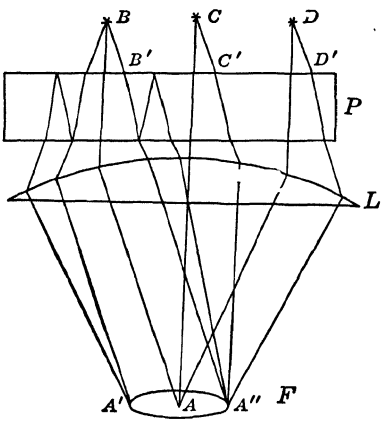


FIG. 132

aperture from each source (of which the rays just considered form a section) we shall have them united in a *ring*, and the ring thus defined will be bright or dark according to the path-difference between the directly transmitted ray and the one resulting from the internal reflections. If in the case of the rays normal to the plate, the path-difference introduced by the double transit through the plate is an odd number of half-waves, the twice-reflected rays will interfere destructively with the transmitted rays, and the illumination at the point A will be decreased. There will be another set of rays in the form of a cone around each normal ray from all points in the source, for which the path-difference is half a wavelength less, since the maximum path-difference is for normal transmission, as we saw in the preceding section. These rays will all come to a focus in a ring surrounding A, and there will be reënforcement here, as the path-difference between the directly transmitted rays and the twice-reflected rays is now a whole number of wavelengths, we thus have a bright ring around a darker centre, and cones of increasing aperture give bright and dark rings in alternation, similar in appearance to Newton's rings, but originating in a very different manner. The actual difference in the illumination will be small as the interference takes place between a feeble twice-reflected ray and the very intense directly transmitted one. Consequently the rings are not very conspicuous by transmitted light. By lightly silvering the two surfaces of the plate the contrast between the bright and dark rings can be made as strong as we please, as we shall see when we come to the Fabry and Perot interferometer. By reflected light the dark rings are practically black, as the intensities of the interfering beams in this case are very nearly equal. To see them by reflected light the best method is to reflect the light from the source (*e.g.* a mercury arc with a green filter) down on to the plate by means of a piece of plate glass at an angle of 45°. The rings are then seen by looking down through the reflecting plate with the eye focussed for infinity. Lummer recommends a concave silvered mirror perforated with a hole, such as is used by oculists, instead of the plate glass mirror.

Influence of Multiple Reflections on Width of Fringes. — If the two surfaces of the plate are lightly silvered so that about 50% of the light is reflected, we have a large number of parallel emergent rays for every incident ray, as we have seen. This circumstance profoundly modifies the character of the fringes. The transmitted system, instead of consisting of light and dark rings of equal widths, is now made up of very narrow bright rings separated by broad regions of complete darkness. This is the principle employed in the Fabry and Perot interferometer, and if two different wave-

lengths are present in the incident light (as with a sodium flame) the D_1 and D_2 fringes, when out-of-step, will not disappear, as with Newton's rings, or the Michelson interferometer, but will form a double system of narrow bright circles. These can be well seen by lightly silvering two small squares of thin plate glass, which are then placed with their silvered surfaces together, but separated by two narrow strips of thick writing paper. If held in front of a sodium flame they will resolve the D lines, *i.e.* show a double set of fringes. By varying the thickness of the paper strips, commencing with thin tissue, the process can be followed. The fringes appear at their best when the eye is normal to the plates, an adjustment best accomplished by bringing the reflected image of the pupil of the eye into coincidence with the flame. The flame viewed by reflection from the opposed silver films shows very narrow dark fringes with broad regions of uniform illumination between them. The transmitted and reflected systems are thus complementary. This experiment illustrates in a striking manner how the reflecting power of a silver film can be destroyed by radiations coming up from below as explained previously.

Referring to Fig. 132 it is clear that if we take incident rays slightly more oblique than the ones giving a bright ring, we shall get a dark ring. Suppose that only two rays are interfering in the transmitted light, as in the figure, and the increment of angle necessary is b. We can represent the illumination as the sum of two vectors, which for a bright ring point in the same direction, thus $\longrightarrow\!\!\longrightarrow$ and for a dark ring in opposite directions thus $\Longleftrightarrow$.

If the plate is silvered and we have a large number of emergent parallel rays due to the multiple reflections of our incident ray the illumination along the bright ring will be represented by vectors thus $\rightarrow\!\!\rightarrow\!\!\rightarrow\!\!\rightarrow\!\!\rightarrow$ while an increment in the angle of incidence very much less than b will now give us zero illumination, since each vector need turn only through a very small angle with respect to its neighbor, to cause the chain to bend around into a closed polygon, which means zero illumination. The next bright ring will occur with an incidence angle such that each vector has rotated through 360° with respect to its neighbor. The intervening region is filled with very faint maxima and minima which are too faint to observe. This treatment will be taken up in greater detail when we come to the diffraction grating, in which the interference fringes produced by light coming through any number of parallel equidistant slits is discussed.

Stationary Light-Waves. — In all the cases of interference which we have thus far examined, the interfering wave-trains have been moving in nearly the same direction. In acoustics we have cases

of interference where the waves are moving in opposite directions. Interference under these conditions gives rise to the so-called stationary waves. If we send a train of waves along an elastic cord, one end of which is fastened, the waves are reflected from the wall, and running backward along the cord, interfere with the direct waves. The cord is at rest at points half a wave-length apart, vibrating in nodes. We should expect something of the kind to occur when light is reflected normally from a mirror, and the possibility of such an action was early recognized. Zenker in his *Lehrbuch der Photochromie* explained the colors sometimes exhibited in photographs of the spectrum taken on silver chloride as due to the formation of layers of reduced silver between the nodal points of a system of stationary light-waves.

The existence of these stationary light-waves was first demonstrated in 1890 by Otto Wiener. When rays of light are incident normally on a polished mirror the reflected rays travel back over the same course. If the light is monochromatic, we shall then have planes of zero illumination half a wave-length apart, parallel to the plane of the mirror.

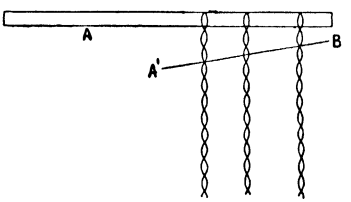


FIG. 133

The presence of these nodal planes was demonstrated by Wiener who employed a photographic film one-twentieth of a wave-length in thickness mounted obliquely in front of the mirror as shown by the line $A'B$ in Fig. 133 which intersects nodes and loops in succession at distances which increase as the angle of inclination is made smaller. The film will be acted on by the light along lines marking its intersection with the loops.

Wiener coated a glass plate with a thin photographic film, placed the film side close to the mirror, at a very small angle, and allowed monochromatic light to pass through the film and suffer reflection from the mirror. On developing the film he found it blackened along lines corresponding to the points where it intersected the loops of the standing waves, while the intervening portions were quite clear. Increasing the angle of course caused the dark lines to become finer, since a greater number of planes were cut in a given distance. The presence of stationary waves can be shown independently of photography by employing a thin fluorescent film in place of the sensitized collodion. This experiment was performed by Drude and Nernst.[1]

Another very beautiful experiment was performed by the same

[1] *Wied. Ann.*, *45*, 460, 1892.

investigators. One-half of a glass plate was coated with a film of silver, and the whole then coated with a fluorescent film only a small fraction of a wave-length in thickness. When this plate was illuminated with monochromatic light obtained from a spectroscope, it was found that the fluorescence was much stronger over the transparent portions of the plate than over the silvered portions. In the former case the film is traversed by a beam of light of which only a small percentage is reflected back through the film, consequently it fluoresces brightly. In the latter case we have stationary waves, and the film is located at the first node, which lies on the surface of the silver, a position in which fluorescence is impossible, since there is practically no vibratory motion at this point.

Wiener also made a study of the interference field formed by light incident at an angle of 45°. In this case the two interfering beams are at right angles to each other, and the nodal planes, parallel to the reflecting surface, are separated by a distance equal to $\sqrt{2}\ (\lambda/2)$ instead of $\lambda/2$ as in the case of perpendicular incidence of the light, as shown by the construction in Fig. 134 in which wave-crests and troughs are represented by solid and dotted lines respectively, the intersections of which give the lines of minima. It should be noted that a crest is reflected as a trough. The experiment was made with polarized light and interference was found when the electric vector was perpendicular to the plane of incidence, as in this case the vectors of the incident and reflected waves are parallel. The corresponding magnetic vectors are, however, at right angles to each other consequently the interference pattern is not represented magnetically except as a system of plane, elliptical and circular vibrations as will be shown presently. With light polarized parallel to the incidence plane the electric vectors are at right angles and cannot interfere, while the magnetic vectors, being parallel, can form a system of maxima and minima. The photographic film recorded interference only in the former case, proving that the electric vector was the one responsible for the photo-chemical change. Drude and Nernst repeated the experiment, using a fluorescent film in place of the photographic plate, and in this case also the electric vector was found to be the one producing the fluorescence, as the maxima and minima were found in the same positions.

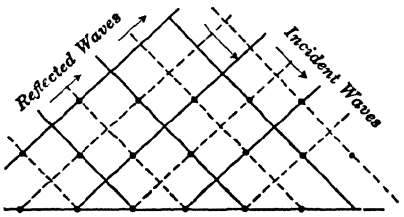


FIG. 134

A very striking and instructive modification of the experiment was made by P. Selenyi who employed a deposit of very minute sulphur particles on a glass plate as a detector of the interference field. These scatter the light only when they are in regions in which there is an electric vibration, as they do not disturb a magnetic field. They are thus luminous in the same regions of activity indicated by the photographic plate and fluorescent film, but differ from these in the following very important respect. The fluorescence and photographic action is a measure of the total intensity of the light, while the scattering is produced only by the component of the electric vector perpendicular to the line of vision, as is shown in the Chapter on Scattering the component parallel to this direction producing no scattering (except in the case of anisotropic particles). With light incident at 45° and polarized parallel to the plane of incidence the electric vectors of the incident and reflected light are perpendicular to each other as shown in Fig. 135*a*. Each can be resolved into two components parallel and perpendicular to the mirror surface. Those parallel to the surface will interfere producing minima and maxima, with the electric vector directed in such a way that particles in the maxima scatter light in directions parallel to the x axis. The other components form a system of interference maxima which cause scattering parallel to y, the two systems being out-of-step, and consequently failing to record photographically. When the plate with its layer of particles is substituted for the photographic plate or fluorescent film, which fail to show maxima and minima in this condition, the maxima and minima are shown by the scattering particles, one set of fringes appearing when the eye views the film along the x axis, the other when the eye is located as shown in the figure.

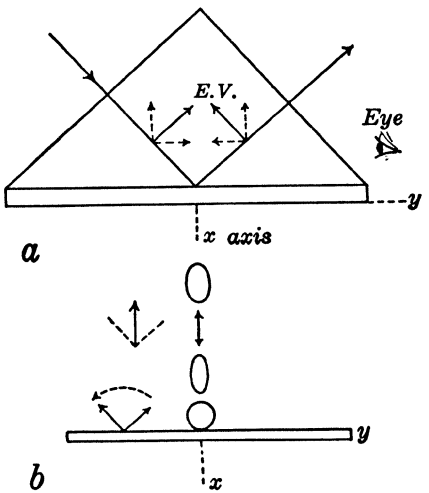


FIG. 135

Or we may consider that the vectors of the incident and reflected light (which are at 90°) compound to a plane vibrations perpendicular or parallel to the surface according to their phase difference (*i.e.*

height above surface) with intermediate elliptical and circular vibrations (page 348). The perpendicular vector causes scattering along y, the parallel one along x. See Fig. 135*b*. (Parallel vector not shown.)

Transition from Stationary Waves to Lloyd Mirror Fringes. — As we have just seen, the distance between the interference maxima and minima in a stationary wave-system increases from $\lambda/2$ to $\sqrt{2}\,(\lambda/2)$ as we increase the angle of incidence from 0° to 45°. The separation increases indefinitely with further increases of angle, as can be seen by drawing wave-fronts as in Fig. 134, but with increased inclination, and in the vicinity of 90° incidence we have the case of Lloyd's mirror, with a fringe separation sufficient for direct observation with a low-power eye-piece.

Lippmann's Color Photographs. — Photographs in natural color were accidentally obtained by E. Becquerel in 1850, by means of standing light-waves, although he was not aware of the part they played. In 1868 Zenker explained the colors, sometimes seen in Becquerel's photographs of the spectrum, as due to standing waves, formed by the reflection of the light from the surface of the silver plate on which the sensitive film was formed. The silver was reduced in the anti-nodal planes forming a system of reflecting laminae, which showed interference-colors in reflected light in the same way as the crystals of chlorate of potash previously described.

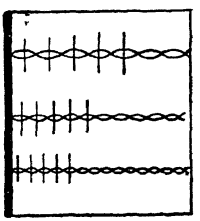


FIG. 136

A process of direct color photography, based upon this principle, was originated by Lippmann in 1891. The photographic plate is placed in the camera with the glass side facing the objective, and the sensitive film backed by a reflecting layer of mercury. This of course requires a special form of plate-holder.

A system of stationary waves is formed in the film as shown in Fig. 136, and the silver compound is acted upon only at the anti-nodes, which form planes parallel to the reflecting surface. On developing and fixing the plate in the usual manner, it is found that the film shows, in reflected light, brilliant colors, similar to the colors which illuminated it. The silver, instead of being reduced in a mass, uniformly distributed throughout the thickness of the film, is laid down in thin laminae, coinciding with the antinodal planes of the stationary light-waves. The distance between the laminae is equal to the half wave-length of the light which formed them, consequently they show the same color by interference in reflected light. The process is not an easy one to carry out, and very

few have been successful with it. Especially prepared plates must be used, as the grain of the commercial plates is too coarse to record the minute structure of the wave-system.

A very complete study of the process was made by H. E. Ives, who gives the following formula for the preparation of the plates:

A. Gelatine 1 gram Water 25 c.c.	B. Gelatine 2 grams KBr 0.25 gram Water 50 c.c.	C. $AgNO_3$.3 gram Water 5 c.c.

A and B are heated till the gelatine melts, allowed to cool to 40°, C added to A and then A to B slowly with stirring, and the whole filtered. After flowing and setting, the plates are washed for fifteen minutes and allowed to dry. The most satisfactory thickness was obtained by flowing the emulsion on and off glass plates at room temperature. The plates are bathed for 10 minutes in a 1/100000 solution of isocol in water and dried rapidly to render them color-sensitive. They are exposed in a plate-holder so arranged that mercury can be introduced behind the plate, and in contact with the film. Exposures with $f/3.6$ on sunlit objects range from $1\frac{1}{2}$ to 5 minutes according to sensitizers, etc.

After development and drying, the pictures are made ready for viewing by cementing a thin prism of small angle to the film side to destroy the disturbing surface reflections, and the back of the glass is flowed with asphaltum varnish. The prism is usually cemented by means of Canada balsam.

The plates bathed in isocol lose their color-sensitiveness in four or five hours, and must therefore be used immediately. If plates of better-keeping quality are required we may introduce 1 c.c. each of a 1/1000 alcoholic solution of a mixture of erythrosine and pinacyanol into the emulsion before pouring it over the plates.

It is of considerable interest to see how nearly we can reproduce a monochromatic color by means of the Lippmann process. If we form the laminae in the film by means of monochromatic light, the light, selectively reflected, will embrace a considerable range of the spectrum unless a very large number of laminae are formed. The case is the same as with the chlorate of potash crystals, which we have already studied.

Mr. Ives made an exhaustive study of this, and found that to obtain a large number of laminae special precautions as to the preparation of the sensitive film and its development must be taken. The film was illuminated by the green line ($\lambda = 5461$) of the mercury arc, and after developing and drying the film, it was used to reflect sunlight into a spectroscope. In this way it was possible to see at once how nearly the film was able to manufacture

green monochromatic light. A number of photographs obtained with the spectroscope are reproduced in Fig. 137. As will be seen the green light manufactured from white light in the case of film No. 6 is highly homogeneous. This was the film compared with the

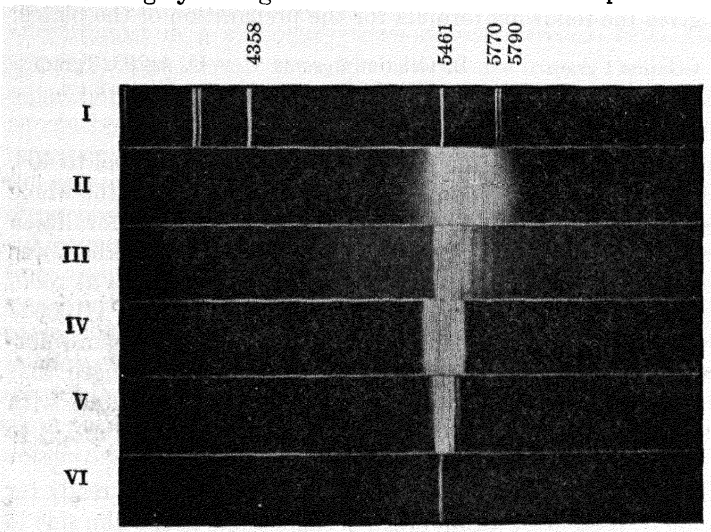


FIG. 137

chlorate of potash crystal previously alluded to. By cutting a section of the film and examining it with a microscope, the laminae thrown down by the standing light-waves can be seen and counted. It was found in this way that as many as 250 laminae could be

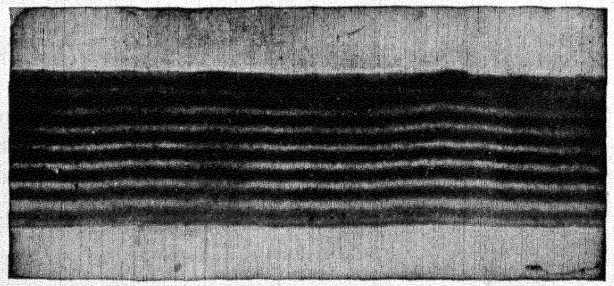


FIG. 138

formed, if the sensitizing dye was introduced into the emulsion and the plate developed with hydroquinone. In the case of bathed plates the sensitizing action only penetrated a short distance, and with pyrogallic acid development the developing action occurred

chiefly near the surface, limiting the possible number of laminae. A microphotograph of a similar section by Mr. Senior is reproduced in Fig. 138.

Neuhauss in 1898 was the first to make thin sections of the film and observe the laminae with the microscope. Since the distance between them is $\lambda/2$ they are at the limit of the resolving-power of the microscope, but the effective distance can be increased by cutting the sections in an oblique direction, or by causing them to swell by the application of moisture as was done by Cajal.

Neuhauss improved the color process by bleaching the pictures with bichloride of mercury; this treatment increases the transparency of the laminae without sensibly reducing their reflecting power, consequently a larger number are able to coöperate, as the incident light is able to penetrate to a greater depth.

CHAPTER VII

DIFFRACTION

The passage of a system of waves past the edge of an obstacle is accompanied by a phenomenon known as diffraction. Primarily the term is applied to the bending of the waves around the edge into the region of the shadow, but in the region beyond the obstacle we find usually a system of maxima and minima of illumination produced by interference and termed diffraction fringes, the study of which is of importance in connection with their bearing on the formation of images by optical instruments and the part played by them in the production of spectra.

Fresnel, who was the first to make a careful study and formulate a theory of these fringes, accounted for them as the integrated effect of the secondary wavelets of Huygens originating on those portions of the wave-front not intercepted by the obstacle. Some years earlier Thomas Young had rightly attributed them to interference, regarding them as produced by the simultaneous action of the unscreened part of the wave and a wave originating at the edge of the obstacle, a viewpoint which has, in comparatively recent times, been confirmed as correct by the very complete mathematical investigations of Sommerfeld. As we shall see later on we can see and photograph the source of this secondary wave, the edge appearing, under suitable conditions of observation, as a highly luminous line.

In our study of the phenomena of diffraction we shall first make use of Fresnel's method, employing the zone construction which he introduced, and which has already been fully described in the Chapter on The Rectilinear Propagation of Light, and later on consider the effects from the viewpoint which regards them as due to the interference between two sets of waves, those passing by, and those originating at the edge of the obstacle.

The older theories accounted for the distribution of light in the vicinity of the geometrical edge of the shadow, but to fully account for the effects observed at large angles of diffraction it is necessary to specify the optical constants — refractive index and absorption coefficient — of the material forming the obstacle, as was shown by the work of Gouy [1] and Wien.[2] The

[1] *Comptes Rendus*, *96*, 967, 1883.
[2] *Berlin Sitz. ber.*, 1885, p. 817.

matter is still further complicated by the state of polarization of the light.

The diffraction of light was first observed by Grimaldi about the middle of the 17th century. Admitting sunlight into a darkened room through a very small aperture, he observed that the propagation of the light by the edges of objects, did not obey strictly the laws of geometrical optics. The edges of the shadow were bordered by several rainbow-tinted fringes, while in the case of very small objects similar fringes were found within the geometrical shadow.

Newton repeated and improved upon the experiments of Grimaldi, using light of different colors, and found that the distance between the fringes decreased as the refrangibility of the light increased, and increased as the screen was removed to a greater distance from the object casting a shadow. He explained this phenomenon on his corpuscular theory as due to attractive or repulsive forces, which the edges of the obstacle exerted on the flying corpuscles.

The first attempt to bring the wave-theory to bear upon the subject was made by Young, who regarded the fringes as due to interference between the rays passing close to the edge, and rays reflected at grazing incidence. The internal fringes he explained as due to the interference of inflected rays, without attempting to explain how the inflection took place, and in this he was in part correct, for the internal system we can regard as a set of interference-fringes produced by two similar sources of light situated at the edges of the obstacle.

Fresnel made a series of experiments with slits having polished and blackened edges, and showed that the intensity of the fringes was independent of the nature of the edge.

He was the first to give a rigorous mathematical treatment of the phenomena, regarding the maxima and minima as the result of the interference of the hypothetical secondary wavelets diverging in all directions from those portions of the wave-front not blocked off by the opaque screen.

Fresnel Class of Diffraction Phenomena. — In commencing the study of diffraction the following simple experiment will be found useful in forming an idea of the actual magnitude of the effects. Arc, or preferably sunlight, is concentrated by a lens on a pinhole in a thin metal plate, the diverging beam from this minute source being received in a dark room on a large sheet of white paper, placed at a distance of two or three metres. Diffraction fringes will now be found to border the edges of obstacles held midway between the source and screen. Beautiful effects can be

secured by shaking a little lycopodium dust in the air, the shadow of each particle being surrounded by colored rings.

Fringes formed in the above manner, in which the source and screen are at a finite distance from the diffracting edge, belong to the so-called Fresnel class. The mathematical treatment of the phenomena is simplified if the screen and source are at infinity. This can be accomplished, on a small scale so to speak, by employing parallel rays from a lens, and mounting the screen at the focus of a second lens placed close to the obstacle on the side facing the screen. The fringes thus formed belong to the Fraunhofer class.

If, in our first experiment, we are dealing only with the fringes bordering a straight edge, the pin-hole may be replaced with a narrow slit, giving enormously increased illumination.

Diffraction by a Straight Edge. — We will begin with an experiment. Illuminate a narrow slit with concentrated sun or arc light, or back the slit with a mercury arc, arranging matters so that the room is in comparative darkness. The light from the slit is rendered parallel by a lens and allowed to fall upon a razor blade mounted with its edge parallel to the slit. If we bring the pupil of the eye into the region of the blade's shadow, at a distance of a metre or two from the blade, we observe that the edge of the blade shines like an independent source of light, *i.e.* light is being sent into the region of the shadow from the edge of the blade, the wave in this region having its centre of curvature at the blade's edge. Now, holding a short focus lens in front of the eye, we examine the distribution of intensity over a plane in front of the eye, and find a uniformly illuminated field at some distance from the edge of the shadow but near the edge very pronounced maxima and minima, colored fringes if we employ white light, or nearly monochromatic fringes if we put a green filter in front of the slit. The fringes, however, are not equidistant, as in the cases studied under Interference, but become less widely separated as we retreat from the border of the shadow. Just at the edge of the shadow there is a bright fringe, conspicuously brighter than the uniform field, which shows that this region is more powerfully illuminated than by the unobstructed or total light from the slit. Within the shadow the illumination falls off without exhibiting maxima and minima.

We have seen that light apparently radiates from the razor edge into the region of the shadow as from a real source. Are we now justified in assuming that the wave from the edge also radiates into the illuminated field, and by interfering with the plane-waves which pass by the edge, gives rise to interference-fringes? If this were the case, we could class the maxima and minima as produced by interference between intense parallel rays from the slit and fee-

ble divergent rays from the razor edge, *i.e.* two sources of light, and a treatment by Sommerfeld has shown that the position and intensity of the fringes calculated in the most rigorous manner is equivalent to a system originated as described above. The distribution of the maxima and minima in space is shown by Fig. 139 in which the intersecting plane and cylindrical waves have been drawn, the minima being given by the intersections of the solid and dotted lines (crest intersecting with trough). Coming back from the edge the fringes trace out parabolic arcs as shown in the figure.

The question as to why the relative intensities of the maxima and minima remain practically constant at all distances from the diffracting edge, appeared troublesome, since the waves radiated by the edge must decrease in intensity with increasing distance, while the plane-waves advance without loss of amplitude. This matter was submitted to Professors Herzfeld and Sommerfeld who supplied this extremely interesting explanation: While it is true that the cylindrical wave from the edge decreases in amplitude in proportion to $1/\sqrt{r}$ *along a line* corresponding to a given angle of diffraction, the amplitude decreases also with increasing angle from the normal or direction of propagation of the plane-waves. But the maxima and minima do not lie on straight lines, but on parabolae, a given fringe moving to smaller angles of diffraction with increasing distance. The intensity of the diffracted cylindrical wave is thus independent of the distance r from the edge *when* measured along the parabolae, the decrease due to increasing r being compensated by the increase due to the fact that the parabola along which we are measuring the intensity corresponds to decreasing angle of diffraction, which means an increase in the amplitude of the diffracted wave.

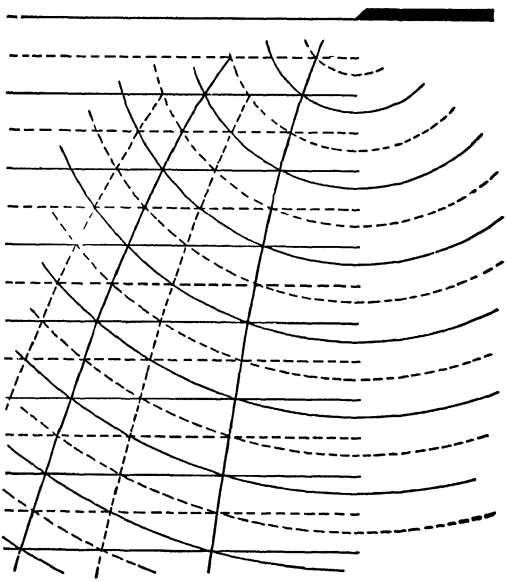


FIG. 139

That the edge diffracts light into the illuminated region, as well as into the shadow can be shown by the following simple experiment: A long focus lens is placed behind and close to the razor blade. This brings the plane-waves to a line focus while the waves radiating from the edge are scarcely affected by the lens on account of its proximity. If we screen off the brilliant image of the slit, the luminous edge is easily seen from directions well outside of the geometrical shadow as shown in Fig. 140 in which the focal length of the lens is shown much reduced. In carrying out these experiments one gets the impression perhaps that the intensity of the diffracted light from the edge is too small to account for the fringes. The same impression is also given by the numerical values given in the following table from Schuster's *Optics*, for a screen at 1 metre from edge.

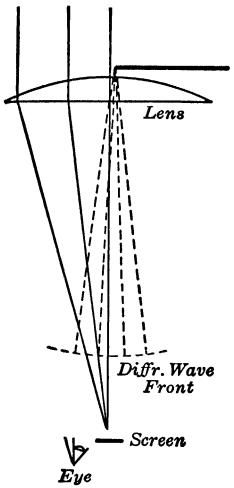


FIG. 140

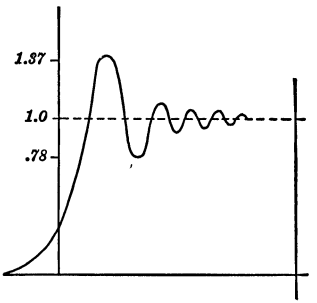


FIG. 141

DISTANCE FROM EDGE IN CMS.	INTENSITIES	
	Outside Geometrical Shadow	Inside Geometrical Shadow
.061	1.3748	.0298
.094	.7774	.0140
.117	1.1995	.0091
.137	.8492	.0067
.154	1.1509	.0053
.170	.8718	.0044

The intensity of the field beyond the fringe system = 1. We have now as the intensities of the first bright and dark fringe 1.37 and .777 to be accounted for by the interference of light of unit intensity with light of intensity .0298 and .014 respectively. If we take the square root of the two intensities, to convert them into amplitudes, and add the first to, and subtract the second from, unity — we get the amplitudes in the bright and dark fringes, and if we square

these numbers we get almost exactly the numbers given for the intensities in the table. The intensity curve is given in Fig. 141.

Treating the diffraction by a straight edge by the method of Fresnel zones, the usual method, we proceed as follows. In Fig. 142 we consider that the source of light is at a distance behind the paper and that we are looking at the edge from the position of the screen on which the fringes are formed. We describe Fresnel zone systems on the wave-front, as we did to account for the rectilinear propagation of light. For a point c, distant bc inside the geometrical edge of the shadow, the zone system will be placed as at the top of the figure, the portions of the zones covered by the edge, and hence inoperative, being dotted. The illumination will be due to the exposed zones, *i.e.* equal to a series of form $m_1 - m_2 + m_3 \cdots$, etc. which will be a fraction of m_1 as the zones decrease in width. m_1 (the amplitude due to the first uncovered zone) will decrease in value very rapidly as we move the point c to the right, *i.e.* as we pass further into the region of shadow. The illumination will therefore fall off rapidly without passing through maxima and minima. At the edge of the shadow we have the zone system as at the middle of the figure, centred on the edge. One-half only of the wave operates and the amplitude is consequently reduced to $\frac{1}{2}$ and the intensity to $\frac{1}{4}$. Outside of the edge at distance ba, the amplitude is due to one-half of the wave plus such portions of the zones as are not covered by the edge. If an even number of complete zones is uncovered, we have a minimum of illumination, *i.e.* the half-wave plus the effect of the two zones — or amplitude —

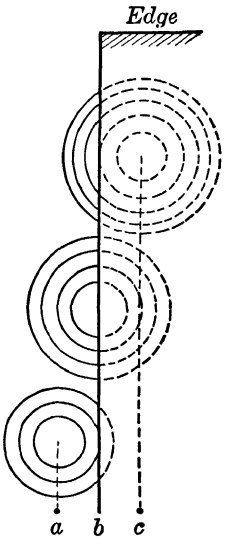


FIG. 142

$$\tfrac{1}{2} + m_1 - m_2.$$

This is the case shown in the figure. If our point is so situated between a and b that only one zone is uncovered, the amplitude will be a maximum, $\frac{1}{2} + m_1$ in which case m_1 is greater than $\frac{1}{2}$ owing to the partial screening of the other zones, which, if allowed to operate, would reduce the effect due to the right-hand half of the central zone. This explains why the bright fringes are more luminous than the field due to the unobstructed wave. As we move away from the edge of the shadow we uncover more zones and the maxima and minima are accordingly spaced as the radii of the

zones, *i.e.* their distances from the edge are proportional to the square roots of 1, 2, 3, etc.

Circular Disk and Aperture. — The zone method is better adapted to the discussion of the distribution of the illumination produced by a small circular disk or aperture, since, if we confine ourselves to the axis of the conical shadow, the disk exactly coincides with zones described on the wave-front with respect to points lying on the axis.

This is the celebrated problem of Poisson, who was led by theoretical considerations to the remarkable conclusion that the illumination along the axis of the shadow of a small circular disk is the same as if the disk were removed, a prediction which was verified experimentally by Arago. It is not very difficult to repeat this experiment, provided it is tried on a large scale. The experiment has been described already in the Chapter on The Rectilinear Propagation of Light.

Applying the Fresnel construction to this case we see that the illumination on the axis is due to the action of the entire wave with the exception of the zone or zones covered by the disk. These reduce by interference to approximately one-half of the effect of the zone bordering the disk. Since the effect of the zones becomes less as we recede from the pole, it is clear that increasing the size of the disk, other things being equal, will reduce the illumination at a given point on the axis. In the case of the aperture the illumination along the axis is due to the zones lying within the aperture. Suppose the point to be so situated that the aperture contains only the two central zones. The disturbances from these will completely destroy each other at the point, and the illumination will consequently be zero. If we bring the point a little nearer to the aperture the scale of the zones will be reduced, and the aperture will contain say three. The two outer ones will annul each other, and we shall have an illumination due to the outstanding central one. We thus see that the illumination is a maximum or minimum according as the aperture contains an odd or even number of zones.

When the illumination is a maximum, as in the last case considered, the amplitude is double that produced by the unobstructed wave, since the whole of the central zone is operative, while the unobstructed wave gives an amplitude due to one-half of the central zone. The intensity is thus increased fourfold by restricting the area of the wave to that of a circular aperture containing one Fresnel zone. At points of zero illumination on the axis the light is found in circular rings surrounding the dark spot.

Diffraction by a straight edge, two apertures, one containing one Fresnel zone and giving a bright point, the other two zones

giving a dark point surrounded by a bright ring and the bright spot at the centre of the shadow of a circular disk, can be simultaneously demonstrated with the following arrangement of apparatus.

Two holes are drilled, near the edge of a small square of sheet brass, with drills 47 and 39 (American gauge) (0.2 and 0.25 cm.). The diameters of these holes will be .079 and .099 inch respectively. Just beyond the edge of the sheet, and between the holes, which should be about 1 cm. apart a thin circular disk (made by turning a brass rod to a diameter of about 6 mms., and then cutting off the end) is mounted on a thin wire or glass fibre, as shown in Fig. 143. It is very important to have the disk accurately circular, with a very smooth edge. A minute hole is punched in a sheet of tin-foil or black paper with the point of a fine needle. Only the point should be used, as the full diameter of the needle will be too great. This is mounted over a hole in a light tight box, and backed by a small mercury arc or automobile lamp. The light from the needle hole is made parallel by a lens of about 50-cm. focus and the perforated screen and disk mounted in the path of the parallel rays. The diffraction effects are examined with a small hand magnifying-glass at a distance of about 2 metres from the screen, the correct point being found by trial. Fringes border the straight edge. A diffraction disk surrounded by rings, and a diffraction ring with dark centre are given by the holes, and a minute spot of light is seen at the centre of the disk's shadow. A small steel ball makes an excellent substitute for the disk. If the source of light is a minute triangular hole in a thin sheet of metal, placed in front of an arc, an image of the triangle can be seen with an eye-piece on the axis of the shadow. We thus have a steel sphere imitating the action of a lens!

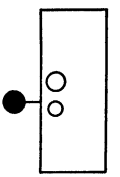


FIG. 143

Graphical Solution of Diffraction Problems. — The graphical representation of the resultant of a large number of vibrations, of continuously varying phase and amplitude, was employed by Cornu in the discussion of diffraction problems. The resultant effect of a number of disturbances of different amplitude and phase can be represented graphically as the closing side of a polygon, the sides of which are proportional in length to the amplitudes produced by the disturbances acting separately, and make angles with a fixed line equal to the phases of the disturbances. For example, in Fig. 144, let OA, AB, BC and CD be the amplitudes produced at a point by four disturbances which arrive simultaneously with phases represented by the angles α. β. γ and δ.

The resultant amplitude will be represented by the closing side *OD* of the polygon, the phase of the resultant being the angle *DOX*.

We can consider the effect of a complete wave at a point in front of it as made up of a large number of small amplitudes, of variable phase and of constantly decreasing magnitude. Each point on the wave-front will produce at the point in question a disturbance of a certain amplitude and phase.

For example, in the lower figure 144 we have a source of light *S* illuminating a screen at *X*, and consider the amplitude at *X* as due to disturbances coming from the pole of the wave and the points *A*, *B*, *C*, on the wave-front. The pole contributes the vector *OA* (which should have been drawn in coincidence with *OX*) while the points *A*, *B*, *C* contribute vectors *AB*, *BC* and *CD*, of decreasing length with phases of increasing amount. The amplitude at *X* is equal to *OD*, the closing side of the polygon. In reality we have an infinite number of points on the wave-front. Let us consider the first Fresnel zone as divided into eight elements, each one of which produces at *X* unit amplitude. The effect of all acting simultaneously can be found by employing the following construction. The first element, which is next to the pole, will produce the amplitude *AB*, the second *BC*, the third *CD*, etc. (Fig. 145); the eighth element will produce the amplitude *HI*, the phase having turned through 180°, since by Fresnel's construction the edge of the zone is half a wave-length farther away from the illu-

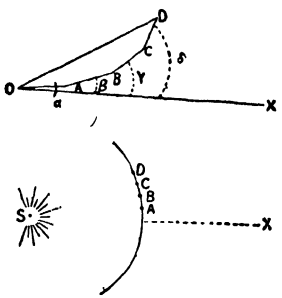


FIG. 144

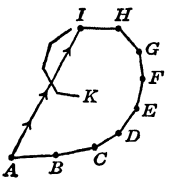


FIG. 145

minated point than the pole. The resultant effect of the first half zone will therefore be represented by *AI*. The effects of the second Fresnel zone can be represented by continuing the construction from the point *I* to the point *K*, and as the amplitudes due to the successive elements decrease rapidly owing to the obliquity, the broken curve will assume the form of the spiral. If now we consider an infinite number of points on the wave-front, the phase will no longer change abruptly, but will vary continuously in passing from each point to the next. The amplitudes due

to the successive elements being very small, the broken curve will now be smoothed out into a continuous one, as shown in Fig. 147.

Fresnel discussed only the diffraction patterns produced by screens bounded by straight lines of infinite length, such as wires, edges and slits. He first showed that the relative intensities at different points on the projection screen along a line perpendicular to the diffracting edge could be determined by considering only the secondary disturbances coming from a section of the wave-front, the problem reducing to the determination of the resultant of an infinite number of secondary disturbances from a limited portion of a linear circular wave. In Fig. 146, 0 is the source of light, AB the screen and F the edge. We have to determine the effect at P of disturbances coming from M, M', M'', etc., on the circular wave-front. If the displacement at A is proportional to $\sin 2\pi t/T$, that at P contributed by an element ds at the edge will be $\sin 2\pi(t/T-b/\lambda)ds$ while an element at M will contribute a displacement represented by $\sin 2\pi[t/T-(b+\delta/\lambda)]ds$, in which $b+\delta=MP$.

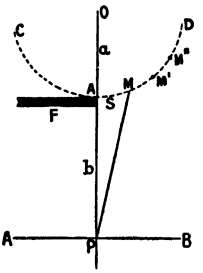


FIG. 146

The displacement at P due to the simultaneous action of all the elements ds of the circular arc will be

$$\int \sin 2\pi\left(\frac{t}{T}-\frac{b+\delta}{\lambda}\right)ds,$$

and the intensity

$$I=\left(\int \cos 2\pi\frac{\delta}{\lambda}ds\right)^2+\left(\int \sin 2\pi\frac{\delta}{\lambda}ds\right)^2,$$

in which we have resolved each disturbance into two rectangular components, which are separately added.

If we can confine our attention to points not far removed from A we can write $\delta=s^2(a+b)/2ab$, as can be easily shown by considering a and b as the longer sides of two right triangles similar to the small triangles which have the side s in common; δ is then equal to the sum of the short sides of the small triangles.

This gives us for the intensity

$$I=\left[\int \cos \pi\frac{(a+b)s^2}{ab\lambda}ds\right]^2+\left[\int \sin \pi\frac{(a+b)s^2}{ab\lambda}ds\right]^2.$$

Writing $\frac{\pi(a+b)s^2}{ab\lambda}=\frac{\pi}{2}v^2$, which gives us

$$s=v\sqrt{\frac{ab\lambda}{2(a+b)}} \text{ and } ds=\sqrt{\frac{ab\lambda}{2(a+b)}}\,dv,$$

and the expression for the intensity reduces to

$$I=\frac{ab\lambda}{2(a+b)}\left[\left(\int \cos\frac{\pi}{2}v^2\,dv\right)^2+\left(\int \sin\frac{\pi}{2}v^2\,dv\right)^2\right].$$

The two integrals occurring in this expression are known as the Fresnel integrals. Integrating them between certain values of v gives us the resultant of the secondary disturbances from a corresponding portion of the wave-front, v varying with s the distance of the wave-front elements from the pole of the wave, the latter taken with reference to the point at which we are determining the illumination. The values of these integrals between 0 and upper limits of various values have been evaluated by different methods by Fresnel, Knochenhauer, Cauchy and Gilbert, and the results given in tables. As we gradually increase the upper limit, the values of the integrals pass through maxima and minima, approaching $\frac{1}{2}$ as a limit, as we see from substitution in the formula

$$\int_0^\infty \sin mx^2\,dx=\int_0^\infty \cos mx^2\,dx=\sqrt{\frac{\pi}{8m}},$$

which gives $\int_0^\infty \cos\frac{\pi}{2}v^2\,dv=\int_0^\infty \sin\frac{\pi}{2}v^2\,dv=\sqrt{\frac{\pi}{4\pi}}=\frac{1}{2}.$

The illumination consists of the sum of the squares of two integrals. The two integrals, therefore, represent the components along two rectangular axes of the resultant amplitude. The illumination is thus represented by the square of a line joining the origin with a point, the coördinates of which are the two integrals. Taking ξ and η as the coördinates of the point for different values of v, we will investigate the curve along which the point moves as v varies. This geometrical

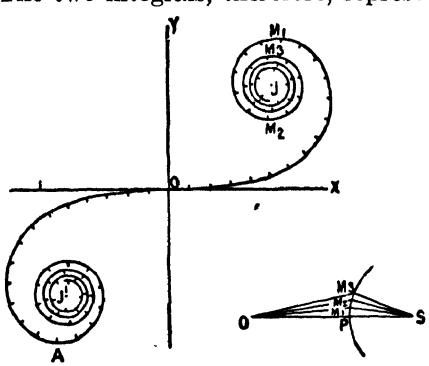


FIG. 147

discussion of the equation is due to Cornu, and the curve is known as Cornu's Spiral. (Fig. 147 and plate at end of book.) By its aid the classical problems of diffraction can be solved in a geometrical manner, the intensity curve of the diffraction pattern being plotted from measurements made on the spiral.

Cornu's Spiral. — Let

$$\xi = \int_0^v \cos \frac{\pi v^2}{2} dv, \quad \eta = \int_0^v \sin \frac{\pi v^2}{2} dv.$$

The curve passes through the origin, since for $v=0$, ξ and η also equal zero. Changing the sign of v does not change the values of ξ and η, but only their sign; the curve is therefore symmetrical about the origin.

The tangent to the curve makes an angle τ with the ξ axis given by

$$\tan \tau = \frac{d\eta}{d\xi} = \tan : \frac{\pi v^2}{2} \text{ or } \tau = \frac{\pi}{2} v^2.$$

At the origin, where $v=0$ the curve is parallel to the ξ axis. For $v=1$ or $s=1$ it is parallel to the η axis, for $s^2=2$ again parallel to the ξ axis, and for $s^2=3$ parallel to the η axis.

The radius of curvature is given by

$$\rho = \frac{ds}{d\tau} = \frac{1}{\pi v} = \frac{1}{\pi s}.$$

For $v=0$ the radius is infinite, and the curve has a point of inflection at this point: as v increases the radius decreases, the curve having the form of a double spiral, which winds about the asymptotic points P and P^1, which correspond to the values of the integrals when the upper limits are

$$+\infty \text{ and } -\infty.$$

The spiral curve given at the end of the book was plotted from a table of the Fresnel integrals, in which successive values of ξ and η are given. This table can be found in earlier editions of this book.

The use of the curve for the solution of diffraction will now be taken up.

The effect of each Fresnel zone is represented by a half turn of the spiral, and if we consider the action of the whole wave the spiral will make an infinite number of turns, finally subsiding to asymptotic circles at J and J' of sensibly zero radius. The spiral OJ represents one-half of the complete wave, and the spiral OJ' the other. The line JJ' joining the two asymptotic points repre-

sents the action of the complete wave. Any portion of the wave is represented by the corresponding chord of the spiral. The effect, for example, due to the second, third and fourth zone on the one side of the pole will be found by joining the ends of the second, third and fourth half turns of the spiral. It is possible by means of this spiral to plot graphically the distribution of light in the diffraction patterns formed when the wave is partially cut off by screens of various types. We will now consider a number of cases.

CASE 1. **Straight Edge.** — The elementary treatment has shown us that in this case we have a system of fringes of decreasing width, outside of the edge of the geometrical shadow, while within the edge the illumination falls off rapidly, without, however, passing through maxima and minima. Consider first the illumination outside of the edge as represented by the spiral.

Since our spiral represents *amplitudes*, we shall find our intensities by squaring the resultant amplitude lines. At the edge of the shadow the intensity will be represented by the square of the distance OJ, since one-half of the complete wave is operative at this point. As we pass out from the edge, the lower part of the spiral begins to operate, and on reaching a distance such that the whole of the first or central Fresnel zone is exposed, the intensity will be found by squaring the line joining the point A with J (Fig. 147); *i.e.* it will be considerably greater than the intensity due to the entire wave, which is represented by the square of JJ'. If we represent the intensities as ordinates, our abscissae must be taken proportional to the distances, measured along the spiral from O, at which the point which we join with J, is located. To facilitate measurements, equal distances have been marked off on the spiral. The first maximum occurs at abscissa 1.4, the distance from O to the bottom of the spiral.

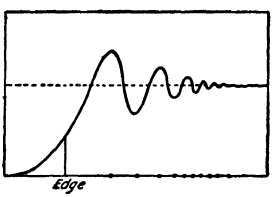


FIG. 148

The intensity as we proceed outward will be represented by the square of the line joining J with a point which travels around the lower half of the spiral. The intensity thus passes through maxima and minima, soon reaching a nearly constant value, owing to the small diameter of the convolutions.

Within the edge of the shadow we have the intensity represented by the square of the line which joins J with a point travelling from O towards J along the *upper* half of the spiral. This line rapidly shortens, without passing through maxima and minima, consequently the illumination drops rapidly to zero. A curve plotted in this way is shown in Fig. 148.

CASE 2. **Narrow Slit.** — In this case the amplitude of the vibration is measured by an arc of the spiral, the length of which is proportional to the width of the slit. The intensity will be represented by the square of a line joining the extremities of a constant length of the spiral. Suppose the width of the slit and its distance from the screen to be such that it subtends exactly one-half of the central zone. The length of the arc which we are to employ is then equal to that of the first half turn of the spiral, OM_1, and within the geometrical projection of the aperture, the arc will lie partly in the upper and partly in the lower branch of the spiral. At the centre it will be symmetrically placed; that is, with its centre at O. As we proceed from the centre, we push the arc of constant length along the spiral, squaring the line joining its extremities at regular intervals, plotting these values as ordinates, at abscissae corresponding to the distances advanced along the spiral as before.

It is at once apparent that the illumination at the centre of the fringe system may be either a maximum or a minimum, according to the width of the aperture. If the aperture just covers the entire central zone, the illumination will be a maximum, and will have a larger value than that due to the whole wave, while it will be a minimum if the aperture covers two zones.

CASE 3. **Narrow Wire.** — This case is a little more complicated, for the effect of the wire is to cut out a constant arc of the spiral just the reverse of the condition in Case 2. The amplitude is the resultant of the two remaining portions of the spiral, which must be compounded, paying attention to the directions as well as the lengths of the lines joining the extremities of the curves. The direction is always found by measuring from J' to J. This can be seen from the following consideration: The effect due to one-half of the wave is $J'O$, that due to the other half is OJ. The whole wave will produce an amplitude equal to the sum of these vectors. If we take their directions as measured from J' to O, and from O to J, the amplitudes will be added, and we shall have amplitude $J'J$. If, however, we measure from O to J' and from O to J, the vectors will be opposed, and cancel each other. If this is borne in mind, no difficulty will be found in remembering how to determine the direction in which the amplitude lines point.

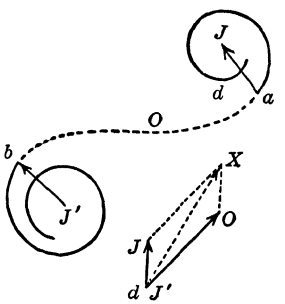


FIG. 149

Suppose now that the wire cuts off one-half of the central zone, *i.e.* one half turn of the spiral measured from O. At the centre the

first elementary distances on each side of O (see Fig. 149, dotted line) will be absent, and the resultant amplitude will be found by compounding $J'b$ with aJ, as shown in the figure. At the edge of the geometrical shadow of the wire, we compound the line $J'O$ with the short line which joins the lowest point, say "d," of the upper branch of the spiral with J (as shown in the lower part of the figure), the resultant amplitude being JX. Abscissae are laid off as before, proportional to the distances advanced along the spiral.

The intensity curve should be plotted in this manner for a wire covering say two zones. It will be found interesting to compare it with the curve obtained in the case of a straight edge. Regarding one side of the wire as a straight edge, we see that the exterior fringe system is complicated by the effects due to the exposed portion of the wave beyond the opposite edge.

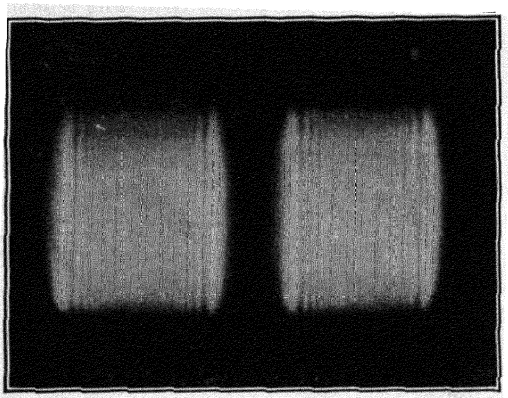


FIG. 150

In Fig. 150 we have a photograph of the fringes produced by a vertical thick wire stretched across a rectangular aperture. The straight-edge fringes appear outside of the edges of the wire, and bordering the left- and right-hand edges of the aperture. No fringes border the upper and lower edge, since the source of light was a vertical illuminated slit.

Within the shadow of the wire are seen the fringes produced by the overlapping of the radiations which penetrate within the shadow. They are equidistant, and we can regard them as interference-fringes produced by two similar sources of light bordering the two edges of the wire.

CASE 4. **Two Parallel Slits.** — The effects in this case are found by compounding the resultants of two arcs, of lengths proportional to the widths of the slits, separated by a distance proportional to the distance between them.

Fraunhofer Class of Diffraction Phenomena. — The mathematical as well as the experimental treatment of diffraction was much simplified by Fraunhofer who placed the diffracting screen between two lenses, with the source and receiving screen at their

foci. This arrangement ensures that parallel rays are incident on the diffracting screen, and the phase is uniform over its area, and that parallel diffracted rays are united at the focus of the second lens. These conditions obtain if we place the screen between the telescope and the collimator of a spectroscope, removing the prism and bringing the tubes in line.

The graphical method can be used in this case also, and we will begin with a very elementary treatment of the effects produced by a narrow slit.

Narrow Slit. — We mount the slit, as described above, between telescope and collimator, or we may simply hold the slit close to the pupil of the eye, placing the source of light at a distance, giving us nearly parallel rays. In this case the lens of the eye focusses the parallel diffracted rays on the retina.

Let DC (Fig. 151) be the aperture upon which parallel waves are incident in a normal direction. The phase of the vibration will then be the same across the aperture, or along any line parallel to

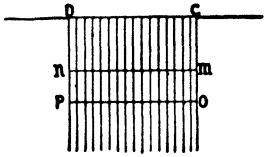


FIG. 151

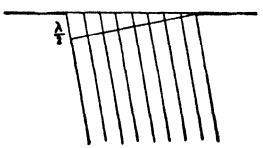


FIG. 152

it, such as p, o. We will consider the normally diffracted ray bundle as divided into 16 elementary ray bundles, which are united by the lens at a point situated at the centre of the diffraction pattern which we are to study. Let the amplitude produced at the focus by one of these elementary bundles be a, then the amplitude produced by all 16 will be $16a$, since they all arrive in the same phase, and the intensity of the illumination will be

$$I = 16^2a^2 = 256a^2.$$

Consider next a bundle of parallel diffracted rays which leave the aperture at such an angle that the path-difference between the extreme rays is exactly one-half wave-length, as shown in Fig. 152. These extreme rays will then arrive at the focus with a phase-difference of 180° and destroy each other. The other elementary bundles destroy one another to a certain extent, and we determine the resultant effect by the graphical method given above.

There will be 16 vectors (giving total amplitude A when in line) instead of the eight given in Fig. 152 the last pointing in a

direction opposite to that of the first (phase-differences between edge bundles 180°). The closing side will be a vertical line, and if we consider that we have an infinite number of diffracted rays, the broken line will become a semicircle, and the closing side its vertical diameter. The length of this line will be the required amplitude "a," while A, the semicircumference, is the amplitude resulting from the normal rays.

We thus have $\pi a = 2A$, or for the relative intensities

$$a^2 = \frac{4A^2}{\pi^2} = \frac{4}{9.86}A^2 = .4056A^2$$

or the intensity at the point at which the lens unites the parallel diffracted rays which leave at such an angle that the edge rays have a path-difference of $\lambda/2$, is about .4 of the intensity at the centre.

For diffracted rays leaving at a greater angle such that the path-difference for the edge rays is λ, the disturbances from the one edge and the centre ($\lambda/2$ path-difference) cancel each other, the same being true for all other pairs of points separated by the same distance. Or if we go back to our 16 original ray bundles, the first and ninth, second and tenth, etc., destroy each other and the illumination is zero, the first minimum bordering the central maximum. By the graphical method our line has now curled around into a closed circle. As the angle of diffraction increases the circle winds up on itself, the line joining the ends of A, our original line of vectors, reaching a maximum when the path-difference beween the edge rays is $3\lambda/2$. In this case we have $\pi a = \frac{2}{3}A$; $a^2 = 4A^2/9 \times 9.86 = A^2/22.2$ or $a^2 = .045A^2$, the intensity being a little over .04 of that at the centre (Fig. 153).

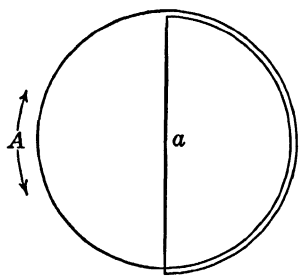


FIG. 153

This graphical method should be thoroughly understood as we shall employ it in an interesting manner in discussing the somewhat complicated fringe system that occurs in the case of diffraction by a number of parallel slits (diffraction grating). The second minimum falls in such a direction that the path-difference for the edge rays is 2λ, the line A having coiled up into two superposed circles.

The diffraction pattern thus consists of a central bright band which is very intense, bordered by alternate dark and bright bands, the intensity of the latter decreasing very rapidly. As the width of the slit is made less, the angle of diffraction necessary to give the

extreme rays a path-difference of $\lambda/2$ becomes greater and the minima retreat from the centre of the system, the fringes broadening.

This can be seen by holding before the eye a slit made of black paper, the width of which can be varied, and viewing a distant lamp through it. The fringes produced by diffraction through a single slit were termed by Fraunhofer "Spectra of the First Class."

Two Parallel Slits. — We will now take up the case of diffraction by two similar parallel slits. They produce spectra of the first class in the same position, *i.e. superposed,* but we shall find that the maxima are broken up by a new set of minima which run through them. These minima are especially noticeable in the bright central maximum, and are produced by destructive interference between diffracted rays from the first slit and corresponding rays from the second.

Let us assume the slits AB and CD to be so narrow that in the direction represented in Fig. 154 the path-difference between the extreme rays A and B, C and D is $\lambda/4$. Rays diffracted in this direction will then be the rays which bring about the illumination of that portion of the central maximum lying midway between the centre and the point where the illumination is .4 of its value at the centre. If but one slit were acting, the intensity would be about .7. If the slit CD is at such a distance from AB that the path-difference between corresponding rays from A and C is $\lambda/2$, the corresponding rays from the two slits will destroy each other and the illumination will be zero.

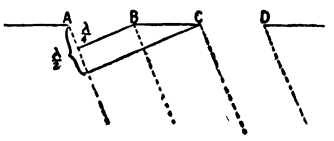


FIG. 154

It is in this way that the new minima are produced, and though they may be investigated in a manner similar to that employed in the case of a single slit, it will be best to postpone their more complete investigation until we come to the discussion of the grating. These maxima and minima were termed "Spectra of the Second Class" by Fraunhofer.

Transition from Fresnel to Fraunhofer Class. — The gain in the brilliancy of the diffraction effects when we use a lens to bring the parallel diffracted rays to a focus, can be shown by passing gradually from the Fresnel to the Fraunhofer class.

In Fig. 155 we have at the left the diffraction pattern produced by a single slit, a bright central maximum bordered by fainter maxima. These fringes belong to the Fresnel class and can be seen by examining the region behind a narrow slit with an eye-piece; a scratch made with the point of a knife on the film of an old photo-

graphic negative answers the purpose. If now we make two slits side by side, say 1 mm. apart, we have the pattern shown at the right-hand side of the figure. In the overlapping portions we have new maxima and minima, much closer together than the others due to the interference of disturbances coming from the two slits, as seen within the faint fringes between *A* and *B*. These maxima and minima do not appear in the bright central maxima *A* and *B*, since these are illuminated each by its own slit, in other words, there

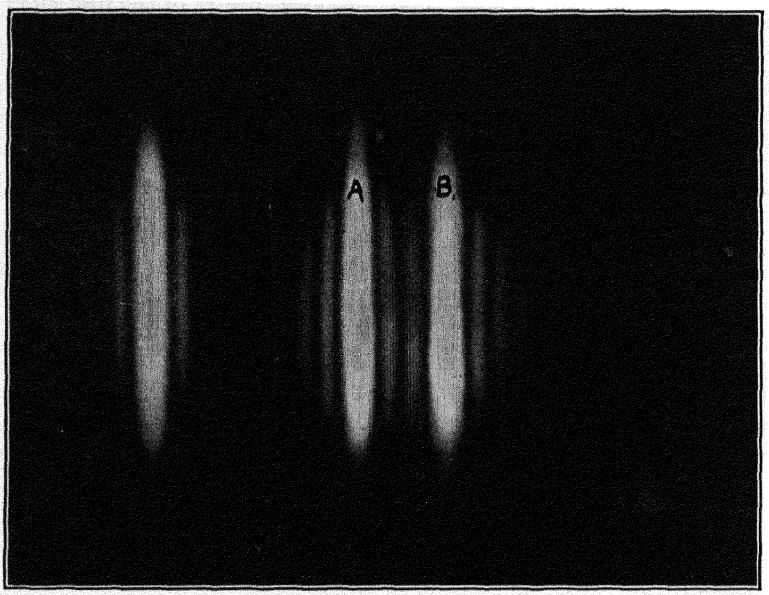


FIG. 155

is no overlapping here, or at least only that of the very faint third maxima, which are unable to produce visible interference in *A* and *B*. If now we put a lens behind the two slits, the two patterns *A* and *B* will be superposed at the focus, and we shall have the new maxima and minima furrowing the brightest part of the pattern. We can pass gradually from one condition to the other, by holding the eye-piece close to the lens and gradually moving it back to the focus. As we do this *A* and *B*, which are at first widely separated, gradually approach, fuse together, and immediately fill up with very bright and narrow maxima and minima. This experiment is extremely instructive, and gives a clear picture of the relation of the Fresnel to the Fraunhofer class. It is apparent that the conditions are the same in this experiment as in the one with the interference refractometer described in the previous chapter.

The following abridgment of the mathematical treatment of diffraction problems of the Fraunhofer class is given to familiarize the reader with the methods employed in developing expressions for the distribution of the illumination in diffraction patterns.

Mathematical Treatment of Fraunhofer Diffraction Phenomena. —If a converging lens is placed behind the aperture upon which plane-waves are falling, the lens transforms that portion of the plane-wave which gets through the aperture into a concave wave, which, if the laws of geometrical optics were followed, would collapse to a point at the focus.

We have then to determine the effect of a small portion of a spherical concave wave at a plane passing through the centre of curvature of the wave. To do this we first get a general expression

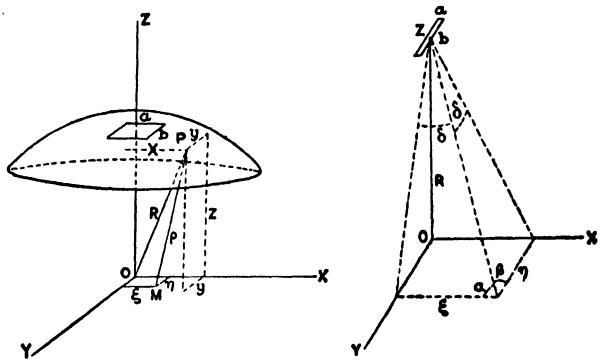


FIG. 156

in the form of a double integral for the effect of a complete hemispherical wave, and then integrate this expression over the aperture (Fig. 156, left-hand figure).

Let the centre of the concave wave be at the origin of three rectangular coördinates. We are to determine the effect of a disturbance starting from the point P with coördinates x, y, z at the point M with coördinates ξ, η, and then the collective effect of all the disturbances coming from the entire wave-front. Call the distance $MP=\rho$ and $dxdy$ the element of the wave at P.

The amplitude at M produced by the secondary disturbances from the area $dxdy$ will be $\kappa dxdy \sin 2\pi(t/T-\rho/\lambda)$, in which κ is a coefficient depending on the inclination of the surface element $dxdy$ to ρ and on the distance of P from M. Since we limit the area of the wave, by blocking off the greater part we can regard the inclination as the same for all portions considered, and κ therefore becomes a constant.

The collective effect of all points P at M, taking into account their mutual interference, is

$$a=\int\int \sin 2\pi\left(\frac{t}{T}-\frac{\rho}{\lambda}\right) dxdy$$

(in which a is the displacement).

From this expression the intensity due to the whole wave, at various points in the xy plane is found to be

$$I=\left(\int\int \cos 2\pi \frac{x\xi+y\eta}{R\lambda}\, dxdy\right)^2+\left(\int\int \sin 2\pi \frac{x\xi+y\eta}{R\lambda}\, dxdy\right)^2.$$

(The development of the above and following expressions will be found in earlier editions.)

Diffraction by a Rectangular Aperture. — Suppose we have a small rectangular aperture of length a and width b, so placed that its sides are parallel to x and y axes, and the z axis passes normally through its centre. To determine the intensity of the illumination the above expression is integrated between the limits $+a/2$ and $-a/2$, $+b/2$ and $-b/2$, and gives

$$I=a^2b^2\frac{\sin^2 \pi\frac{a\xi}{R\lambda}}{\frac{\pi^2a^2\xi^2}{R^2\lambda^2}}\cdot\frac{\sin^2 \pi\frac{b\eta}{R\lambda}}{\frac{\pi^2b^2\eta^2}{R^2\lambda^2}}.$$

In diffraction problems it is, however, more convenient to employ angular measure, as illustrated at the right of Fig. 156. If parallel diffracted rays, which are brought to a focus by the lens, make angles δ and δ' with the planes xz and yz we can substitute $\sin\delta$ and $\sin\delta'$ for ξ/R and η/R in our expression for the intensity due to the aperture a, b.

We thus see that the intensity at the point x, y (or for directions determined by δ and δ') is dependent on two variables of the form $\sin^2 u/u^2$, and that the illumination will be zero if either one equals zero, which will be the case if, m being a whole number, $\pi a\xi/R\lambda=m\pi$ or $\pi b\eta/R\lambda=m\pi$ or $\sin\delta=m\lambda/a$ or $\sin\delta'=m\lambda/b$.

There will therefore be two systems of dark regions of which the equations are $\xi=mR\lambda/a$ and $\eta=mR\lambda/b$, the former parallel to the y axis, the latter parallel to the x axis. If we bear in mind the fact that we oriented the aperture with its side a parallel to the x axis, and that ξ is the value in the direction of this same axis, we see that the distances between the parallel interference minima are inversely as the width of the aperture measured in a direction perpendicular to them, and directly as the distance R and the wavelength λ.

The minima form a reticulated meshwork. The central maximum is very bright and in shape similar to the aperture, but rotated through 90° with respect to it. The other maxima are arranged as shown in Fig. 157, in which A represents the aperture. Practically all of the light goes into two bands at right angles to each other, the faint maxima in the angles only being seen when the brighter portions are cut off by a screen. This diffraction pattern can be well seen by covering a lens of about 50 cms. focal length

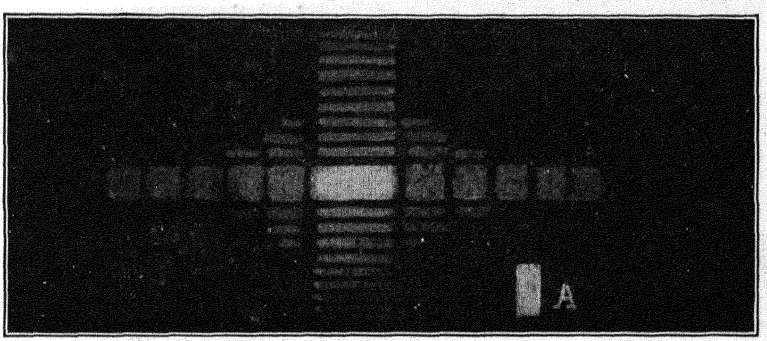


FIG. 157

with a screen perforated with an aperture measuring say 2×4 mms. and placing a pin-hole backed by a mercury arc and green filter and an eye-piece at the conjugate foci of the lens, *i.e.* at distances of a metre from the lens. The experiment should be made in a dark room.

Diffraction by a Single Slit. — If we let one dimension of our small rectangular aperture become large, we have the condition of a narrow slit, the diffraction by which we will now investigate as an introduction to the study of the diffraction grating. In this case we can substitute for the luminous point a luminous line, parallel to the slit, without altering the diffraction pattern, a device which enables us to use much more light, though we do not thereby increase the illumination at any given point to any great degree. With a point source of light the diffraction pattern is reduced to a series of maxima and minima, distributed along a line which is perpendicular to the slit. With a linear source of light the maxima are extended in a direction parallel to the slit, the minima appearing as dark bands.

We will express the intensity by the equation in which the position of the illuminated point is defined by the angle of diffraction δ, and since, if the slit be parallel to the y axis, $(\eta = 0)$, the diffraction will only occur in directions parallel to the x axis, we can put $\delta' = 0$.

This makes the second factor equal unity, and we can write the expression for the intensity

$$I = a^2 \frac{\sin^2 \pi \frac{a \sin \delta}{\lambda}}{\pi^2 \frac{a^2 \sin^2 \delta}{\lambda^2}} = a^2 \frac{\sin^2 u}{u^2}.$$

It will be seen that b^2 has been omitted. This is because varying b does not change the *distribution* light in the maxima. It, however, affects the intrinsic intensity.

The minima of this function, as we have seen, are given by $u = m\pi$ where m is a whole number (not for $m = 0$, however).

The intensity is zero for all directions δ for which $\sin \delta = m\lambda/a$, or if δ is small the directions for zero illumination

$$\delta = \frac{m\lambda}{a} \text{ MINIMA.}$$

To find the directions in which the minima lie we differentiate the expression $\sin^2 u/u^2$ with respect to u and equate to zero, the resulting equation falling into two,

$$\frac{\sin u}{u} = 0 \text{ and } \frac{u \cos u - \sin u}{u^2} = 0.$$

The first of these two equations gives the position of the minima, the second that of the maxima.

The second equation takes the form $u \cos u = \sin u$, $u = \tan u$.

This last equation can be solved graphically by plotting the curves $y = x$ and $y = \tan x$ as shown in earlier editions.

The maxima lie in directions given by $\sin \delta = u_n\lambda/\pi a$, in which u_n is one of the roots of the equation $u = \tan u$, or for small values of δ,

$$\delta = \frac{u_n\lambda}{\pi a} \text{ MAXIMA.}$$

The maxima are the spectra of the first class which we have already investigated in an elementary way.

If white light is employed, the central maximum is white, the other maxima colored, owing to the fact that their position is a function of the wave-length of the light, the red maxima being farther apart than the blue.

Diffraction by Two Parallel Slits. — This case, which we have already studied by elementary methods, is the next step which is taken in the development of the theory of the diffraction grating.

The width of the slits we will call a and the distance between them d.

The diffracted rays coming in a parallel direction from a slit at angle δ with the normal give a resultant intensity

$$A^2 = a^2 \frac{\sin^2 \pi \frac{a \sin \delta}{\lambda}}{\pi^2 \frac{a^2 \sin^2 \delta}{\lambda^2}},$$

in which expression A is the amplitude.

Parallel rays coming from corresponding parts of the two slits have a path-difference of $(a+d) \sin \delta$, and the vibrations when brought together by the lens will have a phase-difference

$$\frac{2\pi(a+d) \sin \delta}{\lambda}.$$

We have seen in the Chapter on Interference that the resultant intensity of two streams of light of amplitude A, with phase-difference e, is $I = 2A^2 + 2A^2 \cos e$.

The resultant intensity in this case is therefore

$$I = 2A^2\left(1 + \cos 2\pi \frac{(a+d) \sin \delta}{\lambda}\right) = 2A^2 2 \cos^2 \pi \frac{(a+d) \sin \delta}{\lambda},$$

and substituting A^2 from above,

$$I = 4a^2 \frac{\sin^2 \pi \frac{a \sin \delta}{\lambda}}{\pi^2 \frac{a^2 \sin^2 \delta}{\lambda^2}} \cdot \cos^2 \pi \frac{(a+d) \sin \delta}{\lambda}.$$

This expression contains two variable factors, one of form $\sin^2 u/u^2$, which we have already investigated, the other giving equidistant minima equal to zero, given by the equation

$$\sin \delta = \frac{(2n+1)\lambda}{2(a+d)} \quad \text{MINIMA},$$

and maxima given by $\sin \delta = \dfrac{n\lambda}{a+d}$ MAXIMA,

which expressions simply state that in the first case the rays coming from homologous parts of the two slits meet with a path-difference of an uneven number of half wave-lengths, and in the second case with an even number.

The intensity will be zero when either of the two variable factors is zero, *i.e.* when

$$\sin \delta = m\lambda,$$

$$\sin \delta = \frac{(2n+1)\lambda}{2(a+d)}.$$

The minima given by the first expression are the diffraction minima of a single slit which we have already studied, the second are interference minima resulting from the meeting of homologous rays from the two slits: they are chiefly noticeable in the central maximum of the first-class system, and the maxima which lie between them were called by Fraunhofer spectra of the second class. They are the spectra yielded by the diffraction grating. See also the interference refractometer in the Chapter on Interference.

If now we increase the number of slits we shall find that we have in addition spectra of a third class, which, however, practically disappear if the number of slits be very large.

Diffraction by Any Number of Parallel Equidistant Slits (Diffraction Grating). — The complete expression for a grating of n lines (or slits) of width a, separated by equal distance b, is

$$I = a^2 \frac{\sin^2 \pi \frac{a \sin \delta}{\lambda}}{\pi^2 \frac{a^2 \sin^2 \delta}{\lambda^2}} \cdot \frac{\sin^2 n\pi \frac{(a+b) \sin \delta}{\lambda}}{\sin^2 \pi \frac{(a+b) \sin \delta}{\lambda}}.$$

An elaborate formula, involving double integrals, the development of which requires several pages of mathematics, and is finally solved by graphical methods, shows that between the principal maxima produced by the grating there are present $(n-2)$ secondary maxima, where n is the number of lines of the grating.

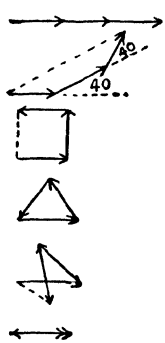


FIG. 158

The following graphical method will be found to account for the secondary maxima, and show as well their number, position and intensity, for a grating of any number of lines. We shall make use of the well-known method of compounding vibrations, which is employed in the elementary development of Cornu's spiral, and shall show that we have minima equal to zero whenever the amplitude lines form a closed symmetrical figure, or mutually annul each other in pairs. The closed figures are either triangles, squares, regular polygons or star-shaped figures, which can be plotted in a very simple manner.

Fraunhofer's treatment shows that a single slit produces maxima and minima, which recede from the centre and broaden as the slit width decreases. These he called spectra of the first class. In the case of the gratings used for optical purposes, the lines are so fine

that the central maximum of the first-class spectra occupies the entire field; *i.e.* there are no minima, a single line scattering light of decreasing intensity throughout the entire range between 0° and 90°. In the present treatment we shall consider lines of this degree of fineness. "Absent spectra," resulting from finite width of the line, and the consequent existence of first-class minima, can be separately dealt with.

We consider parallel rays incident normally upon the grating, the parallel diffracted rays being brought to a focus by a lens.

Each line of the grating acting alone, we will suppose to produce unit amplitude at the focus.

We find the resultant amplitude produced by a number of lines operating together by compounding vectors as in the graphical

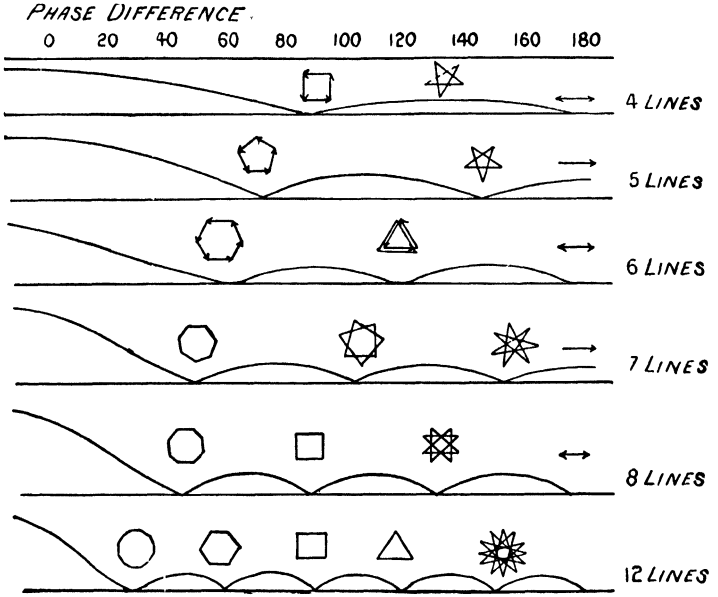


FIG. 159

treatment previously given, the resultant amplitude being the closing side of a polygon, the sides of which (vectors) represent the amplitudes and phases of the vibrations coming from the grating-lines. We can plot the intensity curve for a three-line grating, by considering phase-differences (P.D.) which increase by 20°. In the normal direction (P.D.0°) the intensity will be 3^2 or 9; in a direction such that we have a P.D. of 90° the intensity will be 1, while with a P.D. of 120° we have a triangle, there is no "closing side" and the intensity is zero. From now on it increases,

attaining the value 1 again with a P.D. of 180° when the three vectors are superposed; two of these vectors cancel each other, the illumination being that due to the outstanding one. The various stages are shown in Fig. 158 for different values of P.D. The first order spectrum comes in such a direction that the P.D. is 360° or the path-difference is λ; consequently the point for which P.D. = 180° is midway between the "central image" and the first spectrum, and the diffraction pattern is symmetrical about it. We thus see that there is a secondary maximum at this point.

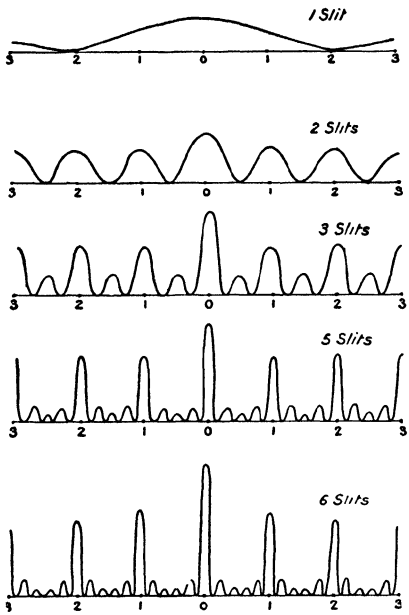


FIG. 160

From now on we shall only determine the positions of the minima when more than three lines operate. The complete curve can be calculated in the same manner.

In the case of a four-line grating we have intensity 16 at the centre, zero when the P.D. is 90° or 180°, and unity when the P.D. is 120°. This gives us two secondary maxima between the principal maxima, their intensity being about $\frac{1}{14}$ that of the latter; these maxima occur when the P.D. is 135° and 225°. In Fig. 159, we have the positions of the minima and the form of the closed amplitude figure at each, for gratings of 4, 5, 6, 7, 8 and 12 lines. The ordinates of the amplitude curves are not drawn to a scale, of course.

With a five-line grating we get our first minimum when the five amplitude lines form a pentagon, the phase-difference being 72°, and a second when they form a star, the phase-difference in this case being 144°. At the centre (180°) we have intensity one, as in the case of the three-line grating.

We thus see that, in the case of a grating of n lines, we have $(n-2)$ secondary maxima between the principal maxima, the intensity of which can be easily calculated from diagrams similar to those given.

The intensity curves for 1, 2, 3, 5 and 6 slits are shown in Fig. 160, from which we see that by increasing the number of lines of the

grating we throw more light into the principal maxima, the secondary maxima decreasing in intensity and becoming more crowded together. We also cause the curves of the principal maxima to become steeper, *i.e.* the light is concentrated more and more nearly into a line. When we have a very large number of lines the principal maxima are exceedingly bright and narrow when monochromatic light is employed, and the secondary maxima disappear entirely. The principal maxima then constitute the narrow spectrum lines seen with the grating.

The formation of spectra by gratings can be well illustrated by mercury ripples, as was shown by Vincent. If we fill a large shallow

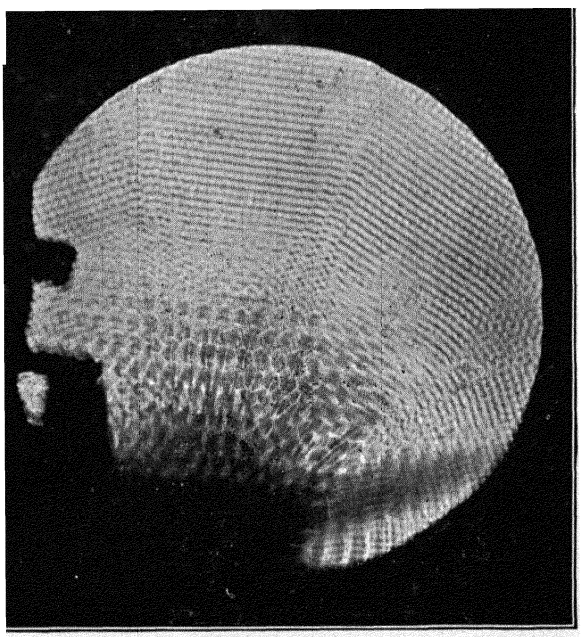


FIG. 161

basin with mercury, and cause a strip of cardboard attached to a vibrating tuning-fork to touch its surface, plane-waves will be emitted, parallel to the edge of the strip. If, however, we substitute a coarse comb for the continuous strip, and allow the teeth to touch the surface, we shall have a number of systems of plane-waves, corresponding to the central image and lateral spectra. The experiment has been repeated by Reese [1] and illustrated by a photograph similar to the one made by Vincent. It is reproduced

[1] *Astrophysical Journal*, xxiv. 48.

in Fig. 161. The author has found that a ring of glycerine run around the edge of the mercury surface completely prevents disturbing reflections from the sides, by damping the waves.

We will now consider the case of the optical grating with many thousand lines, and see what part the secondary maxima, which accompany the spectrum lines, play.

Our broken line now becomes essentially a smooth curve. We have our first minimum when it forms a complete circle, the phase-difference between disturbances from the first and last lines being 360°, or the path-difference λ. The first secondary maximum occurs when the line has wound up into a circle and a half. The ratio of the intensity of the secondary to that of the principal maximum is obviously the ratio of the square of the diameter of the circle of $1\frac{1}{2}$ turns to the square of the total length of the line. This we easily find by winding up a strip of paper of known length. It will be found to be about 1: 23.

This shows us that, no matter how many lines we have in the grating, our spectrum lines will always be accompanied by close companions, having at least $\frac{1}{23}$ of their brightness. For an eight-line grating the ratio is not very different, being about $\frac{1}{21}$. It occurs for a phase-difference approximately such that we have a regular pentagon, three sides of which are made of double lines. For a four-line grating it is about $\frac{1}{16}$ (closing side of star at 135°), and for a three-line grating $\frac{1}{9}$. This last is the maximum value of the ratio.

It seems quite surprising that even for optical gratings the secondary maxima have a brilliancy very nearly one-half of that which obtains in the case of a three-line grating. Their angular distance from the spectrum lines is such as to make the path-difference between disturbances coming from the first and last lines of the grating $3\lambda/2$ more than the path-difference at the spectrum line. This angle is obviously that subtended by one and one-half waves, at a distance equal to the width of the ruled surface. The distance between the secondary maxima is thus seen to depend upon the width of the grating, and not upon the number of lines. At first sight there may appear to be some difficulty about this, since there are $n-2$ secondary maxima, and we might very naturally expect an increase of n to push them nearer together. This is, however, only true when the "grating space" remains constant, *i.e.* when we add new lines of the same spacing.

Suppose we have a grating of given width with 20 lines which gives 18 secondary maxima. If we interpolate lines between the lines already present, we double the number of secondary maxima, to be sure; but the principal maxima (spectra) of odd order dis-

appear by interference; in other words, they are now twice as far apart as before, the spacing of the secondary maxima remaining the same.

The analogy between the secondary maxima and the fringes produced by a rectangular aperture of the same size as the ruled surface, can be studied to advantage by means of coarse gratings made by ruling four or five lines on a piece of smoked *plate* glass, and making the lower third of the grating clear by wiping out the lines. Sun or arc light filtered through red glass should be used with a small spectrometer, the grating and aperture being covered in succession or used simultaneously.

It is easy to see that, in a spectrum of order m, formed by a grating of n lines, the path-difference between disturbances coming from the first and last line is $mn\lambda$. The first secondary minimum to the right or left of the central maximum will lie in such a direction as to make the path-difference between the extreme rays change by one whole wave-length; that is, the path-difference in this case is $(mn \pm 1)\lambda$. For, as we have seen, this first secondary minimum occurs when our line of vectors turns around into a closed circle, the first and last vectors pointing again in the same direction. This means that the last vector has turned through 360°, which corresponds to a path increment of λ.

We shall use these expressions when we come to the resolving power of gratings; they are of fundamental importance.

Number of Spectra and Intensity Distribution: Absent Spectra. — In the case of gratings formed of opaque lines and transparent intervals the spectra of progressively increasing order lie in such directions that the path-difference between disturbances coming from the corresponding edges of two adjacent apertures is 1, 2, 3, 4, etc., wave-lengths. Referring back to Fig. 154, suppose the direction of the dotted lines to be the direction of the second order spectrum. The path-difference between disturbances coming from A and C is then 2λ. If the widths of the opaque bars are equal to the widths of the apertures, as in the figure, the path-difference in this direction, between the waves coming from A and B will be λ, and the total effect of the disturbances from the aperture will be zero, as we have seen in the elementary treatment of a single slit. If each aperture produces zero illumination, the total effect of all must be zero, or the second order spectrum will be absent. The spectra will in general be absent in the directions corresponding to the minima of the first class due to a single slit. It is easy to see that if $a = b$, as in the above case, the spectra of even order 2, 4, 6, etc., will disappear. If $a = b/2$, the spectra of order 3, 6, 9, etc., will be absent; if $a = b/3$, spectra 4, 8, 12, etc., will disappear.

Verification of Results. — The results which we have deduced for the diffraction grating can be easily verified by experiment. A piece of thin plate glass is smoked over a flame, and one edge moistened with alcohol. The alcohol spreads over the film, and on drying leaves it compact enough to enable us to rule lines through it with a sharp steel point by means of the dividing engine.

If no dividing engine is available it is not difficult to fit up an arrangement by which the plate may be advanced through equal distances by turning a screw provided with a roughly graduated head, while the ruling is done with a needle point mounted on a pivoted arm, arranged to slide back and forth along a fixed arc. The distance between the lines should be as nearly as possible equal to the width of the lines. A dozen or so lines will be found sufficient. The plate thus ruled should be provided with a cover, also of thin plate glass (ordinary window glass will not do), to prevent injury to the film, and a movable slide of thin black paper so arranged that the lines can be covered or exposed in succession. If we mount the plate on the table of a spectrometer and illuminate the slit with sunlight which has been passed through a sheet of dense ruby glass, we can verify in succession all of the results which we have deduced.

If all of the slits but one are covered we shall see the broad spectra of the first class, the central one being by far the brightest; on uncovering another slit, this central maximum, as well as the lateral maxima, appear furrowed by narrow dark bands, the bright bands between them being the spectra of the second class. If three lines be uncovered the faint third-class maxima appear, one between each pair of second-class maxima. As we proceed with the uncovering process we shall see the secondary maxima crowd in between the principal maxima, until, when all the lines are exposed, they are too faint and too close together to be detected.

If a piece of dense cobalt glass, which transmits the extreme red and blue only, be substituted for the ruby glass, a series of red maxima and blue maxima will be seen simultaneously, the latter being closer together than the former.

There is another cause which may operate in causing the disappearance of spectra of certain orders, which we will now investigate.

The Laminary Grating. — If we consider the opaque strips of the grating which we have just studied replaced by transparent strips of such a thickness that some one wave-length in the spectrum suffers a retardation of $\lambda/2$, we have a type of grating which was first studied by Quincke. These gratings have the peculiar property of failing to show the central image when light of the specified wave-length is used. If we employ white light, the central image, which is white in the case of the ordinary grating, appears colored

owing to the absence of the wave-lengths in the immediate vicinity of the one for which the retardation by the thin lamina is $\lambda/2$.

In the upper figure (Fig. 162) it is clear that the normal rays, which traverse the laminae interfere at the focus with the rays which have passed between the laminae. The light, absent in the central image, goes to the first order spectrum as shown by the lower figure, disturbances along AA' and CC' having a path-difference of λ. A disturbance along BB', having traversed the lamina now has also a path-difference of λ with respect to AA', since the half-wave retardation due to the lamina is added to the geometrical $\lambda/2$ retardation. Consequently all diffracted rays reenforce in the direction δ.

Gratings of this type for demonstration purposes may be made by ruling lines on a silver deposit on glass, and then converting the silver into the transparent iodide by covering the surface with

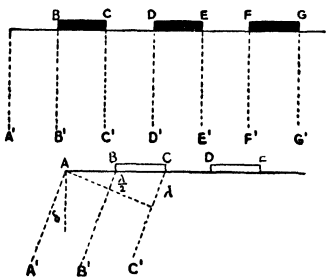


FIG. 162

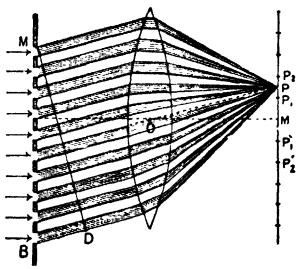


FIG. 163

iodine crystals. Or they may be ruled on a film of bitumen as described in the section treating zone-plates in the first chapter.

The Plane Grating. — Diffraction gratings for spectroscopic apparatus are usually ruled on a reflecting surface of speculum metal. If the surface is plane the case is analogous to that of the transparent grating. The formula for the grating we have already deduced in considering the diffraction fringes of the second class, due to two parallel slits. It is $n\lambda = (a+b)(\sin i \pm \sin \theta)$ in which θ is the angle of diffraction, i the angle of incidence, n the order of the spectrum and $(a+b)$ the grating constant. For normal incidence we have $\sin \theta = n\lambda/a+b$. This condition is represented in Fig. 163 from which it is apparent that the disturbances which originate at the grating elements, when a plane-wave is incident normally upon it, pass through the point P in succession. It is thus apparent that even if but a single wave or pulse struck the grating, we should have a periodic disturbance at P. The grating is thus able to manufacture, as it were, light of a definite wave-length or

color. The optical paths from what we may call the diffracted wave-front MD to the point P are equal, and since the path-difference between two adjacent diffracted streams for the first order spectrum is λ, the path-difference between the extreme rays is $m\lambda$, m being the number of lines in the grating. As we have seen, if we employ monochromatic light from a slit, made parallel by a lens, we shall have at P a sharp maximum accompanied by very faint maxima, which lose in intensity and crowd closer to the principal maximum as we increase the number of lines in the grating. The path-difference between the extreme rays for the position of the two minima immediately adjacent to the principal maximum is $(m \pm 1)\lambda$, an expression which we shall make use of in considering the resolving-power of the grating. If n is the order of the spectrum, the path-difference for the principal maxima and adjacent minima are $mn\lambda$ and $(mn \pm 1)\lambda$. The formula for the grating shows us that the position of the diffracted image depends upon the value of λ.

The dispersive power of the grating is represented by

$$\frac{d\theta}{d\lambda} = \frac{n}{(a+b)\cos\theta},$$

which shows us that the dispersion increases with the order of the spectrum, and that it is inversely proportional to $(a+b)$, the grating constant. For small values of θ, $\cos\theta = 1$ approximately, and the spectrum is normal, *i.e.* equal increments of wave-length correspond to equal increments of θ. In the higher orders of spectra, however, the dispersion increases with θ and therefore with λ, and the spectra are more drawn out at the red than at the violet end, exactly the reverse of what we have in the case of prismatic spectra. On this account the grating should be mounted on the table of the spectrometer so as to stand normal to the observing telescope, for in this position $\theta = 0$, for the centre of the spectrum and the other values of θ are small.

The general formula for a reflecting grating, for any angle of incidence i, is

$$(a+b)\ (\sin i \pm \sin\theta) = m\lambda.$$

Overlapping Spectra. — The formula for the grating shows us that the spectra of the different orders overlap, for by doubling n and halving λ we have the same value of θ. This overlapping gives no trouble in the visible region, if we limit ourselves to the first or even the second order, but in photographic work it must be taken into account, for the ultra-violet of the second order is superposed on the visible region of the first order, wave-length 25 of the second

coinciding with 50 of the first. This overlapping is often of use, as we shall see when we come to the consideration of the concave grating; if it is desired to eliminate it, color screens or prismatic analysis must be resorted to: glass cuts off practically everything below $\lambda = 32$, consequently a glass lens or plate is all that is necessary when working in the first order spectrum.

Resolving Power of Gratings. — Since the diffraction grating is used largely in place of a prism, for the formation of spectra and the examination of spectrum lines, it is of importance to determine upon what the resolving-power of the grating depends. Lord Rayleigh has shown that in order to separate two spectrum lines, the distances between the central maxima of their diffraction images must be at least as great as the distance of the first minima from the central maximum. If this condition is as represented in Fig. 164, the resultant illumination (dotted) on half scale being given by summing the ordinates of the two curves, the lines will not appear clearly separated, but the duplicity of the line can be recognized from the slight shading down the centre. Let AB (Fig. 165)

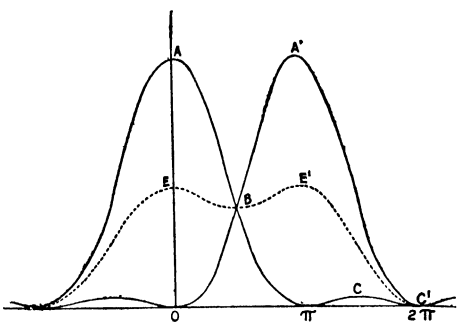


FIG. 164

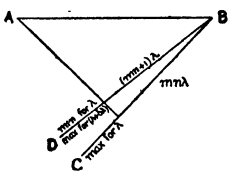


FIG. 165

be the grating and BC the direction of the diffracted rays, which form, for wave-length λ, the central maximum of the spectrum of the mth order. As we have seen, the central maxima in each spectrum are accompanied by secondary maxima which decrease in intensity, and become crowded together as we increase the number of lines of the grating. The first minimum at π on each side of a central maximum will obviously lie closer to the central maximum of the spectrum line, if we employ a large number of lines in the grating, or more exactly if the grating has a considerable width. In other words, the intensity curve becomes steeper, the light crowding together more nearly into a geometrical line. As an expression for the resolving-power we require the change in wave-length necessary to shift the central maximum into the position of the first minimum. For the central maximum for wave-length λ, in a spectrum of order m for a grating of n lines the retardation between the dis-

turbances coming from the edges of the grating is $mn\lambda$. The first minimum to the left for this wave-length is represented by the line *BD*, which will be the direction of the central maximum for some other wave-length slightly greater than λ, say $\lambda+\delta\lambda$. The retardation of the extreme rays of the first minimum is $(mn+1)\lambda$ or $mn(\lambda+\delta\lambda)$, if $\delta\lambda$ is the increment of wave-length which will shift the central maximum into the position previously occupied by the first minimum. (See treatment of diffraction grating.) Since these two are identical, we can write

$$mn(\lambda+\delta\lambda)=(mn+1)\lambda, \text{ or } \frac{\delta\lambda}{\lambda}=\frac{1}{mn}.$$

For the *D* lines of sodium $\delta\lambda/\lambda=1/1000$, so that in order to resolve them in the first order spectrum, we must utilize at least 1000 lines of the grating. They may be separated in the second order spectrum with 500 lines, etc. This treatment is due to Lord Rayleigh.

It is especially to be noticed that the resolving-power of the grating does not depend upon the closeness of the ruling, but merely upon the number of lines. Let us take for example a grating one inch in width, ruled with 1000 lines, which in the first order spectrum will barely resolve the sodium lines. Suppose now we interpolate an additional 1000 lines, making them bisect the original spacing. The spectra of odd order will disappear by interference, the energy being thrown into the spectra of even orders, which increase in brilliancy, the gain being fourfold since the amplitudes are double.

The resolving-power in each spectrum is exactly the same as it was before, since what is now the first order spectrum was previously the second order. If one-half the grating is cut away, leaving 1000 lines in half an inch, the dispersion will not be altered while the brightness and the resolving power are halved. The sodium lines are now just barely resolvable in the first order spectrum. If the grating had been cut in halves before the interpolation of the second ruling, the sodium lines would have been just barely separated in the second order spectrum, 500 lines only operating. This spectrum is identical in position with the first order spectrum in the second case. It is thus clear that with a grating of given size, the resolving-power *in a given direction* is quite independent of the number of lines in the grating. With few lines we have a high order spectrum at a given point, and as we increase the number of lines by ruling more in the *same space*, the spectra move out from the central image, and the order of spectrum at the given point becomes less.

If, however, we decrease the order at the given point by compress-

ng the grating, that is, decreasing the grating space without increasing the number of lines, then the resolving-power at the given point becomes less. The advantage of ruling the lines close together is twofold. In the first place, for a given aperture of telescope and collimator, we can bring more lines to bear, and consequently increase the resolving-power; in the second place, we concentrate the light into fewer spectra, and obtain increased brilliancy.

Focal Properties of Gratings: Zone-Plates. — If the distance between the lines of a plane grating, instead of being constant, varies according to the same law which governs the spacing of the Fresnel-Huygens zones, the grating will give focussed spectra without the aid of mirrors or lenses.

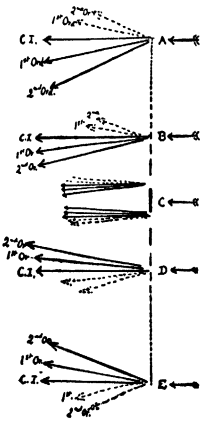


FIG. 166

It is obvious that the diffraction angle for a spectrum of given order must be, for each element of the grating, such as to cause the diffracted rays to meet at a point, instead of remaining parallel to one another.

We may regard the zone-plate, described in the Chapter on The Rectilinear Propagation of Light, as a circular diffraction grating of variable spacing.

In Fig. 166 let $ABCDE$ represent a section of the zone-plate through which parallel rays of light are passing. At the points A and E, where the elements are close together, we have the diffracted rays corresponding to spectra of different orders leaving at the angles designated by arrows. It is clear that the rays for the first order spectra will meet at a point at some distance from AE, and the second order rays at a point nearer to the plate.

At B and D, the distance between the diffracting lines being greater, the diffraction angles are less and the rays from these elements meet at the same points as those from A and E. We thus see that the different foci of the zone-plate merely represent spectra of different orders. The diffracted rays indicated by dotted arrows meet only if they are produced backwards behind the zone-plate. These foci are of course virtual, and the plate therefore acts as a concave, as well as a convex, lens.

Intensifying Glass Gratings. — Gratings ruled on glass can be much improved by a method described by the author in 1906.[1] They usually give rather weak spectra, the diamond cuts being

[1] R. W. Wood, "Intensification of Gratings and the Diffraction Process of Color Photography," *Phil. Mag.*, Dec., 1906.

deep and narrow, usually much narrower than the clear spaces between them. Dilute hydrofluoric acid widens the cut and increases the intensity of the spectra often as much as tenfold. The process can be controlled most perfectly, and there is no danger of injuring the grating. Coat the inside of a saucer with hot paraffine and allow it to cool. Fill the saucer with water and pour in two or three c.c. of hydrofluoric acid.

Stir well with a small glass rod and then apply a minute drop to one corner of the grating. After ten seconds, wash the grating and dry. It will be found that the spot covered by the drop shows enhanced spectra. Now fasten a handle to the back of the grating with soft wax and lower the face into the acid, taking care that none flows over the edge. After ten seconds wash and dry the grating. If the spot first etched shows up brighter than the rest of the ruled surface repeat the process until the whole grating shows the same intensity as the small spot. This is the point at which to stop. Further etching will produce no increase in brilliancy. It is most important to avoid getting even the smallest drop of this acid on the skin. If an accident happens apply wet bicarbonate of soda.

False Lines or "Ghosts" Due to Errors of Ruling. — No gratings are perfect since, up to the present time, no method has been found of advancing a plate by exactly equal increments. It could be done, of course, with a ruling engine on which the motion was controlled by interference-fringes, and small gratings of great perfection could undoubtedly be ruled in this way, employing either the eye or the photo-electric cell for observing the transit of the fringes across a fine wire.

The high degree of perfection attained with the dividing engines employed for the ruling of optical gratings depends first upon the perfection of the screw, the turning of which advances the plate. Rowland devised a method of making a practically perfect screw, and of very nearly compensating the slight periodic errors due to a small to-and-fro motion of the screw with respect to the thrust-block or end-bearing resulting from a lack of parallelism of the bearing surfaces and other factors, but even in the very best gratings that have been turned out, optical tests show the presence of periodic errors. These give rise to spurious lines which may lie close to the real lines, or be widely separated from them.

There are three quite distinct types of "Ghosts": Those which result from a periodic error with a period identical to the pitch of the screw were first studied by Quincke but are usually referred to as the Rowland ghosts, and appear to the right and left of the true lines, with an angular separation corresponding to spectra formed by a grating with a spacing equal to that of the threads of

the screw of the ruling engine. If not too intense they are not very troublesome, as they can usually be recognized at once from their symmetrical distribution about the main line which they accompany. They are somewhat useful at times for determining the true centre of a greatly overexposed line, which lies midway between each pair of ghosts. Their intensity with respect to the main line, varies as the square of the order of the spectrum in which they are observed. The best that has been done thus far (ruling 15,000 lines to the inch) is to reduce them to about $1/1000$, $1/250$ and $1/110$ of the intensity of the main line in the first, second and third orders respectively. With 30,000 lines to the inch $1/250$ in the first order is about the best to be expected.

The production of ghosts by periodic errors can be seen by the following construction, which accounts also for the circumstance that their intensity with respect to the main line increases with the square of the order of spectrum. In Fig. 167 the numbers represent the lines of a grating, the first order spectrum lying in the direction indicated. Here the path-difference between rays from lines 1 and 6 is 5λ. In the directions in which the ghosts are to appear we will suppose that, if the grating is perfect, we have complete darkness as a result of a path-difference of $5\lambda \pm \lambda/2$ between diffracted rays from lines 1 and 6, 2 and 7, 3 and 8, etc., across the grating. Now suppose lines 6 and 2 (dotted) and all other lines of this wider spacing very slightly misplaced, say by an amount making the original path-differences of $5\lambda \pm \lambda/2$, (which accounts for 0 illumination) increase or decrease by $\lambda/20$. The destructive coöperation of these lines with the others will now be incomplete and we shall have a slight illumination of say amplitude 1, in the specified direction. This will be the "ghost." In the second order spectrum the change of path-difference due to the displacement will be $\lambda/10$ and the destructive interference will be still less complete giving amplitude 2, or intensity 4; that is, the intensity has increased as the square of the order. The periodicity of the error is, of course, that of the dotted numbers.

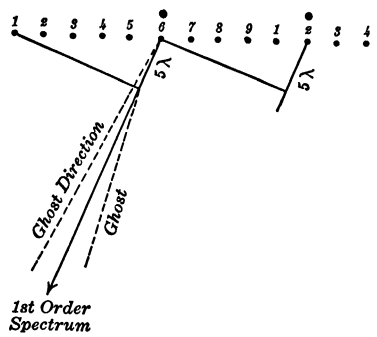


FIG. 167

If one line in m is omitted or an extra line inserted at the same intervals, ghosts are produced, but in this case their intensity with respect to the main line which they accompany is independent of

the order of the spectrum. This can be seen experimentally by superposing two glass gratings, one ruled with, say 100 lines to the millimetre and the other with ten, and viewing a slit backed by a mercury arc through the combination. Ghosts of similar type occur if one line in m is ruled heavier than the remainder.

A very clear and thorough treatment of the theory of imperfect gratings has been given by C. M. Sparrow [1] in which he points out that errors of displacement of lines from their true position affect the phase, while errors of the type last considered affect only the amplitude; the former increase in intensity with the square of the order of spectrum.

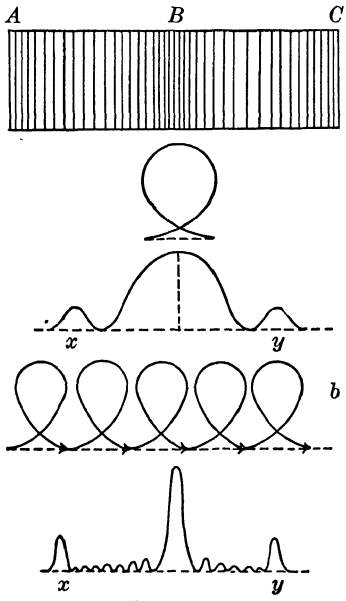


FIG. 168

The phase errors are the ones which occur commonly in the case of gratings, the spacing of the lines varying slightly in a periodic manner. As Sparrow shows this means that the line of vectors representing the resultant of all grating lines in a single element of the periodicity does not form a closed circle for the direction corresponding, say, to the first minimum which borders the main spectrum line, as they would do if the rotation of each vector with respect to its neighbor was the same.

Suppose that the grating consists of p groups of lines in which the spacing, instead of being uniform varies as shown on an exaggerated scale in Fig. 168. Considering now as a grating, only the group of lines AB the direction of the first minima bordering the principal maxima will be such as to make the path-difference of disturbances from A and B increase by λ from its value at the maximum, *i.e.* the disturbances arrive in similar phase. For a perfect grating the vectors representing the amplitudes due to the individual lines close around into a circle for this direction, giving zero illumination, since each rotates through an equal angle with respect to its neighbor, but in the above case the rotation is greater for the vectors representing lines midway between A and B since they are farther apart, and instead of a closed circle we have a curve of the type shown by a. The amplitude at x and y of the

[1] *Astrophysical Journal*, *49*, 65, 1919.

curve below will therefore be that represented by the dotted arrow of curve *a* instead of zero. If now we have *p* such groups as *AB*, the curve of vectors becomes a series of loops as shown by *b* in the figure instead of *p* superposed circles, as would be the case for equal spacing in *AB* and all subsequent groups. The "errors of closure" (dotted arrows) are all in phase, and their sum gives the amplitude at *x* and *y*. But the principal maximum has now been narrowed to $1/p$ of its former value (owing to the employment of *p* more lines than in the first case), and in the direction of the old minima (at *x* and *y*) we now have secondary maxima, or ghosts, as shown in the lower amplitude curve.

The periodic structure can usually be seen on the surface of the grating if a telescope is focussed on the ruled surface from the position of the green mercury line. If the ghost lines are prevented from entering the telescope objective by suitably placed screens the structure disappears, for reasons which will be clear when we come to the subject of diffraction in its relation to image formation.

A very rapid and simple method of measuring the intensity of these ghosts is the following: A thin metal disk 20 cms. in diameter is mounted on the shaft of a small electric motor, taking care to have it accurately "centred" so that its rim runs true. A small section of the rim is cut away, and over this aperture a radial slit of variable width can be formed of two small pieces of stiff black paper attached with soft wax, or better by an adjustable metal sector. The motor is mounted in such a position that the disk covers the main line, but passes the light of the first order ghost. When the disk is running, we observe with a low-power eye-piece the main line (seen intermittently through the slit) and the ghost (seen directly) and adjust the width of the slit until they have the same intensity. The ratio of intensities will then be that of the slit width to the circumference of the disk. If the slit width is under a millimetre, under the conditions specified, the grating is first class, so far as the Rowland ghosts are concerned. We can check our results by observing in the second and third orders. The green line of the mercury arc should be used.

A second type of ghost, first described by Lyman, is more insidious, as the faint spurious lines are far removed from the real lines. They are due to a periodic error of ruling of the order of magnitude of six or eight lines of the grating, and have been responsible for the supposed discovery of lines in the infra-red region, which were in fact merely ghosts of visible lines. Their cause and a method of eliminating them was found by the author. The Rowland engines were driven by a leather belt, and it was found that the periodicity involved was that of the belt. When thicker por-

tions (*e.g.* the splice) passed over the pulley, there was an abnormally strong pull on the spacing wheel which warped slightly the frame of the engine. A portion of the belt was intentionally thickened by winding it with string, and Lyman ghosts of great intensity were found, quite strong green lines appearing in the red and infrared regions of the mercury arc.

The driving mechanism of the engine was now altered employing a device previously designed for the construction of a rotating mercury mirror to eliminate the transmission of vibrations from the motor to the rotating dish of mercury. A pulley wheel of the same diameter as that of the large driving pulley of the engine was mounted on a separate support close to the latter, the two being attached by light steel springs close to the rim. The belt passed around the new pulley, which dragged around the driving wheel of the ruling engine by the tangentially applied forces of the springs.

This device successfully "laid the ghosts" which have never appeared in gratings ruled since its adoption.

A third, and more troublesome type of ghost has been extensively studied by the author, the spurious lines appearing very close to the real lines, and resembling the satellites of complex lines, for which they may easily be mistaken. They are due to non-periodic errors resulting from variable friction on the ruling engine. To detect their presence it is best to employ the yellow line of neon, which is very bright and has no companions. If visual observations or photographs show faint companion lines between the main line and the first-order Rowland ghosts, we can be sure that we have ghosts of this type.

If the grating is observed with a telescope by the light of a single line, and the main line is screened off by a knife edge, those portions of the ruled surface which contribute to the formation of the spurious line will appear luminous while the rest of the surface is dark.

The method of crossed gratings, which is analogous to Newton's method of crossed prisms, enables us to recognize a "ghost" in cases where it is not otherwise possible to do so. As this method has been used in the interpretation of the results obtained with the interferometer of Lummer and Gehrke, which will be described in the Chapter on Interference Spectroscopes, it may be well to take it up in this connection.

In Fig. 169 let O be the source of light (slit of spectroscope, for example) viewed through two superposed gratings, with their lines perpendicular to each other, one of which shows "ghosts." Let AB be a first-order spectrum produced by this grating, with bright

lines 1, 2, 3, and a very bright line 5 accompanied by two ghosts 4 and 6.

The other grating produces the spectrum *CD* of the source and the oblique spectrum *EF*, which we can regard as the spectrum produced by the good grating (lines horizontal) of the spectrum *AB*, produced by the bad grating (lines vertical) that is to say, the different elements of *AB* are deviated by different amounts proportional to their wave-lengths. The lines 1, 2 and 3, being true lines and having different wave-lengths, will lie in the positions 1, 2, 3, on *EF*. Lines 4, 5 and 6, however, have the same wave-length, consequently they will be deviated by equal amounts and occupy positions shown by 4, 5, 6, on *EF*. In other words, the ghosts are thrown out of the spectrum. The same method could be used with reflection gratings.

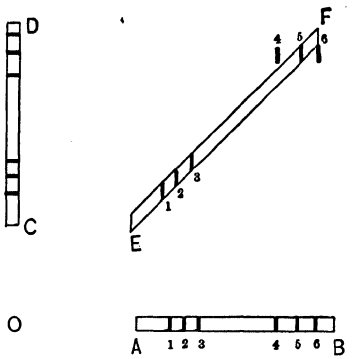


FIG. 169

In Fig. 170 will be found photographs of the spectra of a helium tube, made by the author with two crossed gratings. The left-hand

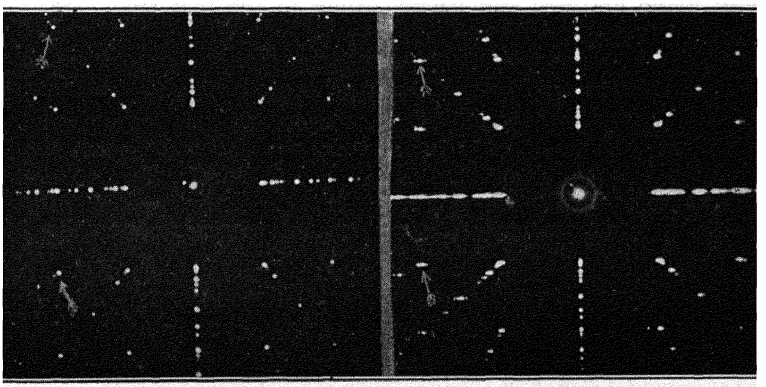

FIG. 170

figure was made with a pair of gratings free from ghosts. The source of light was the small brilliantly illuminated circle of an "end-on" tube. The spectra show the helium lines in the form of rows of dots, and the dots lie along single straight lines in the diagonal spectra. In the right-hand figure one of the gratings has been replaced by a very poor one with a bad periodic error; the

horizontal spectra produced by this grating are almost unrecognizable, owing to the numerous ghosts, which produce multiple dots overlying one another. The vertical spectra produced by the other grating are seen to be free from ghosts. In the diagonal spectra the horizontal arrangement of the ghosts at each spectral line (dot) is clearly seen. Compare points indicated by arrows.

Concave Gratings. — To obtain sharply focussed spectra by means of the plane diffraction grating, two lenses are required: one to render the light parallel before its incidence upon the grating, the other to unite the parallel diffracted rays in a focus. The brilliant discovery was made by Rowland that gratings ruled on concave spherical surfaces would of themselves furnish focussed spectra, excelling in sharpness those obtained by means of lenses. This discovery marked an epoch in the history of spectroscopy, for by dispensing with the lenses, and the absorbing action which they exerted on the ultra-violet, the region of short waves could be explored with an accuracy never before attainable, and the gratings ruled upon Professor Rowland's machine have become the standard instruments for spectroscopic work throughout the world. They combine the image-forming power of concave mirrors with the spectrum-producing power of gratings. A marked advantage of the concave grating lies in the fact that the superposed spectra of different orders are all in focus, which is not the case with plane gratings and lenses, owing to the fact that complete achromatization can never be obtained. It is thus possible to measure the relative wave-lengths with great accuracy. An ultra-violet line of wave-length 2950 of the second order spectrum will be photographed nearly in coincidence with the *D* lines of sodium, and its wave-length can be very accurately measured relatively to these lines. This method of coincidences was originated by Rowland. But the greatest advantage of all is the fact that, when properly mounted, the concave grating yields spectra which are truly normal, *i.e.* spectra in which the distances between the lines are proportional to the difference of their wave-lengths.

Various methods of mounting the concave grating have been devised. That due to Rowland is the following: The theory of the grating, which we shall take up presently, shows that if the grating and the illuminated slit are both situated on a circle, the diameter of which is equal to the radius of curvature of the grating, the spectra of different orders will all be in focus upon the same circle. The spectra are normal along that portion of the circle diametrically opposite the grating, consequently if a photographic plate is placed at this point and bent to the radius of curvature of the circle, the photographic image will be everywhere in focus, and the spectrum

will be normal. To pass from one part of the spectrum to another we have only to move the slit around on the circle, a method sometimes employed. With fixed sources of light, such as the image of the sun formed by a lens in combination with a heliostat, this is impossible, and Rowland devised the following extremely ingenious mechanical device, by which the camera and grating could be moved, with reference to a fixed slit, so as to comply with the required conditions.

Two tracks *AB*, *AC* (Fig. 171) are rigidly mounted on fixed beams, so as to meet accurately at a right angle. On these tracks roll a pair of carriages which support a trussed tube of iron, the length of which is equal to the diameter of the large circle, *i.e.* the radius of curvature of the grating. One of the carriages carries the camera, the other the grating *G*, while the slit is permanently mounted above the point where the rails meet. As the camera is moved away from the slit, the grating is drawn towards it, the three always remaining on the circumference of the circle, with the grating and camera always at opposite ends of a diameter. The grating is turned into such a position that its centre of curvature coincides with the centre of the photographic plate. This method of mounting the grating is now seldom employed.

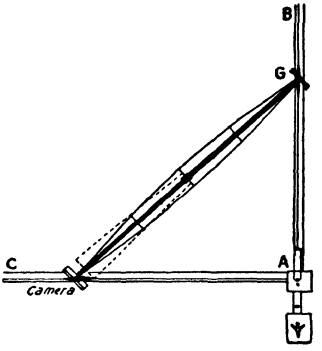


FIG. 171

Concave gratings show always strong astigmatism, which increases as we pass to higher orders. This means that a point on the slit is drawn out into a vertical line. If the slit is not adjusted so as to be exactly parallel to the lines of the grating the astigmatism will cause the lines to broaden, and the resolving power suffers. A simple way of securing this adjustment visually is to clamp the centre of a vertical rod to the slit, and fasten two strings to the end of the rod. One then observes some line showing fine structure such as the mercury line 5890 and, by alternately pulling the strings, set the slit for the sharpest definition. A more accurate setting can be made by taking a number of photographs changing the setting of the slit by a micrometer screw.

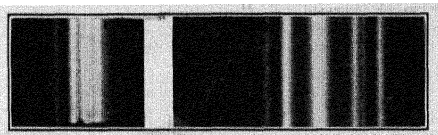
When the astigmatism is very pronounced, as in higher orders, and the slit not very long, if we place our eye at the focussed line and view the grating we see that the light which forms any given part of the line comes from a comparatively narrow horizontal

strip on the grating surface, this strip moving up or down on the grating according to the portion of the line at which the eye is located. On this account there is no advantage in making the lines of the grating very long, the custom being to rule a wide area with short lines, rather than to cover a large portion of the surface, as is done with plane gratings to secure more light. The latter do not show astigmatism, and the accurate setting of the slit is not necessary. The length of the slit must also be considered in connection with the intensity of the lines, a short slit giving small intensity when strong astigmatism is present. The mathematical theory of the concave grating, given in earlier editions of this book, has been omitted.

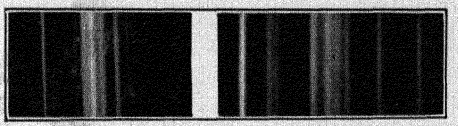
Comparison Spectrum with Concave Grating. — Owing to the astigmatism of the concave grating a comparison prism cannot be placed against the slit as with prism spectrographs. It was shown by Sirks, however, that if a small right-angle prism is placed at a certain distance in front of the slit, *i.e.* between the source and slit, and the light from the comparison source is reflected from this to the slit, the comparison spectrum will cut across the main spectrum with sharply defined edges. Rowland used in front of the plate a thin wooden or metal shutter with a narrow horizontal aperture, which could be rotated so as to alternately screen a narrow strip along the centre of the plate or two areas immediately above and below this strip, the two exposures being made in succession.

Testing of Resolving-Power and Focus. — In determining the focus of a concave grating the most convenient source of light is a small commercial glass Cooper-Hewitt mercury arc, with the paint scraped off from the bulb to permit end-on use. A current of hot air can be directed against the spot serving as a window, to prevent condensation of the mercury. The yellow mercury line 5790 is the best line to observe for visual tests. Its hyperfine structure is shown on Plate 5, on which are reproduced photographs of a number of the mercury lines made by the author with a 7-inch plane grating and a lens of 40-feet focus at the Loomis laboratory at Tuxedo. These will be found of use in determining the correct focus, as many gratings show a spurious structure of exceedingly fine lines even when considerably out of focus. (Lines 5460 and 4358 by J. A. Anderson.) The absorption spectrum of iodine is also a good test of resolving-power, but gives little or no evidence of the faint spurious lines between the main line and the Rowland ghosts, which are due to aperiodic errors, and are shown by many gratings.

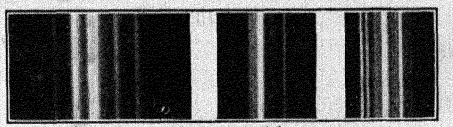
Paschen's Mounting of Concave Gratings. — Rowland's method of mounting the grating is seldom used. The most satisfactory



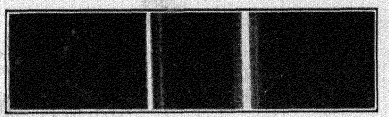
5769.6 5790.6



5460.7 4358.3

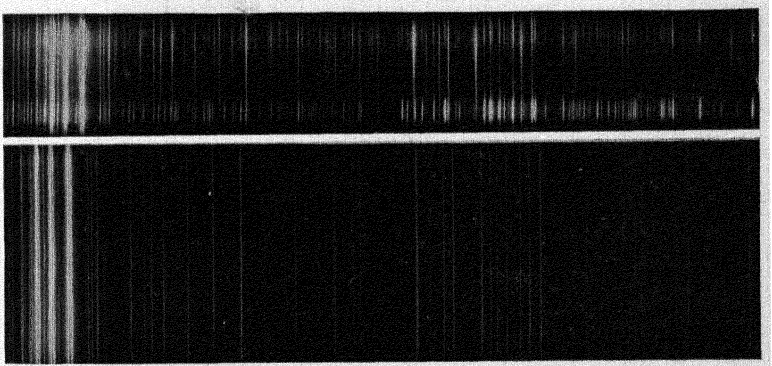


4046.5 3650.1 3654



3662.9 3663.3

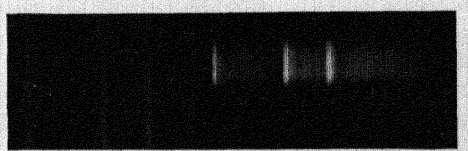
Hyperfine Structure of Mercury Lines with Plane Grating 40-Foot Spectrograph.



Manganese Spectrum near 2700. 21 Foot 30,000 Lines (to inch) Concave Grating, Paschen Mounting. Exp. Times Equal.

Upper Spectrum: Angle of Incidence 16°. Minimum Astigmatism. Lines Produced Only at Electrodes Easily Recognized.

Lower Spectrum: Angle of Incidence 38°. Astigmatism Strong. All Lines Extend Uniformly across Spectrum.



N_2 Bands. Left Incidence Angle 38°.
Right Incidence Angle 16°. Concave Grating.

PLATE 5. CONCAVE GRATING PHOTOGRAPHS ILLUSTRATING PAPER BY G. H. DIEKE ON ASTIGMATISM.
J. O. S. A., *23*, 274, 1933.

mounting for general work is that devised by Paschen. It has the advantage of extreme rigidity and the entire spectrum in practically all orders can be taken simultaneously if desired. No plate-holder is used, the plates being clamped to a circular steel track mounted on a cement pier of semicircular form as shown in Fig. 172. The grating is mounted in a fixed position on a separate pier, on a continuation of the Rowland circle, as is also the slit which is best mounted in a hole in the wall of the grating room, the source of light being in the adjacent room. The walls of the grating room are painted black, and the temperature should be controlled by an electrical heater and thermostat if long exposures are required. A horizontal ventilating fan mounted on the ceiling at the centre of the room keeps the air in circulation and prevents "striae" of warm and cool air from drifting about which seriously impair the definition. A basement or sub-basement room should always be used for this type of mounting. It is advantageous to arrange a mounting for a second slit which can be placed on the track at about the position of the second order red in the figure, the light entering the room through a tube passing through the wall. The astigmatism is much less in this case, and the spectra enormously brighter. This arrangement permits two independent experiments to run simultaneously without dismounting any apparatus, a distinct advantage when using complicated sources of light, as in the study of the Zeeman and Stark effects.

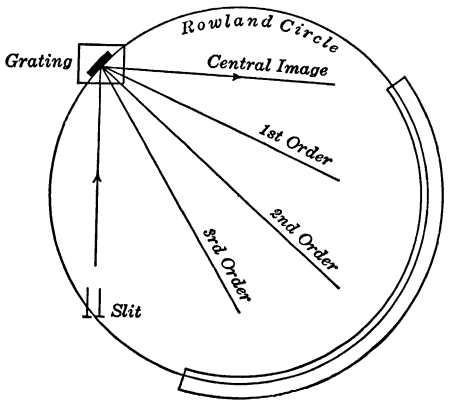


FIG. 172

Eagle's Mounting. — A very compact form of mounting for the concave grating was designed by Eagle. The plate-holder, capable of rotation on a vertical axis, is mounted at one end of a long light-tight box. The slit tube enters the side of the box at such a distance in front of the plate-holder as to cause its virtual image in the total reflection prism (which send the rays down the box to the grating) to be in coincidence with a point just below the centre of the photographic plate. The grating is capable of rotation on a vertical axis, and can be moved towards or away from the slit on a track according to its orientation. The relation of the Rowland

circle to this form of mounting is shown in Fig. 173. The circles, all of the same diameter, centre in points 1, 2, 3 and 4 and the corresponding positions of the grating and plate are indicated by the same numbers. As is clear from the diagram, slit, grating and plate lie always on the circle, which shifts its position as the grating is turned and moved towards or away from the plate.

Other advantages of the Eagle mounting are increased brightness of the spectra due to less astigmatism. In the third order of the Rowland mounting the astigmatism is 4.3 times the length of the lines on the grating: with the Eagle mounting it is only 0.47 the length. One is able also to reach higher orders than are possible with the Rowland mounting, and spectra on either side of the normal can be used.

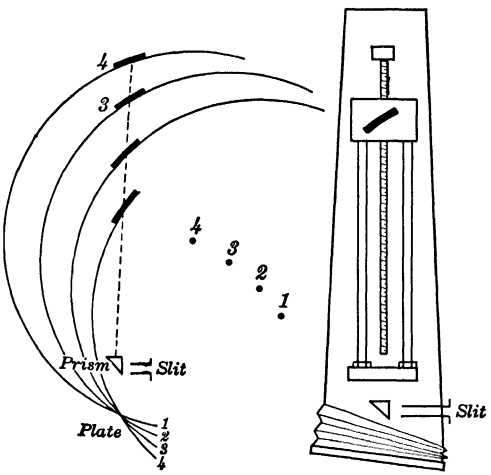


FIG. 173

Concentration of Light into a Single Spectrum. — If a grating is ruled with a diamond ground to, say, an angle of 90° so oriented as to rule grooves one side of which is very steep and the other inclined at a small angle to the surface of the grating, most of the light will go to the spectrum lying in the direction in which the light is reflected from the less inclined sides of the grooves. It is possible in this way to get as much as 75% of the total light into the first spectrum on one side, with 15,000 lines to the inch. A grating recently ruled by the author showed this efficiency. The central image was practically absent, for a bright sodium flame could not be seen in the ruled portion by direct reflection. If a grating is ruled on a copper "flat" with, say, 7000 lines to the inch and the diamond set so as to rule a groove throwing the light into the fourth order spectrum, a celluloid replica will concentrate the light in the first order, the transparent prisms forming the grating refracting the light in this direction. Extremely brilliant and highly efficient transmission gratings can be made in this way. Attention was first drawn to the possibility of making such gratings by the late Lord Rayleigh many years ago.

The "Echelette" Grating for the Infra-Red. — Gratings of this type were first ruled by the author [1] in 1910, with a natural crystal of carborundum on flat copper plates such as used for "half-tone" work. They were made with various spacings from 2000 to 3600 lines to the inch with a groove form similar to that just described.

With visible light they sent the greater part into a group of spectra on one side, say from the 12th to the 16th or from the 24th



FIG. 174

to the 30th order. They may thus be regarded as reflecting echelons of comparatively small retardation and were named "echelette" to distinguish them from the ordinary grating and the Michelson echelon which will be treated presently. They showed no image by reflection in the normal direction, *i.e.* they gave no central image, but when held at an angle of 20° it was possible to see a very good image of one's face. A photograph of an early echelette grating standing normal to its label, seen also reflected in the eighth order spectrum is reproduced in Fig. 174. With infra-red radiation say around 3 μ the greater part of the total energy goes into the first order spectrum, and they can be ruled to meet this condition for any desired wave-length even up to 50 or 100 μ.

With the imperfect ones of poor figure the author, in collabora-

[1] *Phil. Mag.*, Oct., 1910.

tion with A. Trowbridge [1] in 1910 resolved the CO_2 band into a doublet with maxima at 4.3 μ and 4.43 μ. This was two years before Bjerrum predicted the doublet structure of vibration-rotation bands (also foreseen by Lord Rayleigh as long ago as 1892).

The recent increase in the field of research in infra-red spectroscopy, resulting from the development of the theory of molecular band spectra, has made it appear worth while to improve this type of grating.

It was found that carborundum crystals were too fragile to be depended upon, but excellent results were obtained, with a diamond ground and polished to the proper angle, on optical flats of copper which were plated with chromium after the ruling was finished.

If a mercury arc is viewed in an echelette grating one sees most of the orders for the monochromatic radiations though some are missing since they fall in directions corresponding to the minima produced by a single element of the grating, and in this way the grating can be calibrated for infra-red work.[2] They have proved very efficient in infra-red investigations where high resolving-power is required. With visible light the central image and spectra of lower orders are practically absent with a strong concentration in the region of, say, the 21st to the 23rd order, as shown on Plate 6, the upper spectrum showing the appearance with white light and the next with a sodium flame. The blaze of white light is formed by the superposition of many orders, as can be seen by plotting on cross section paper the spectra of the 20th to the 25th order for wave-length 4000 to 6000, inclusive; we find that at the point where the 29th spectrum for violet falls we have colors from 9 spectra of lower orders superposed, which form the white blaze of light in the direction of reflection from the edges of the grooves.

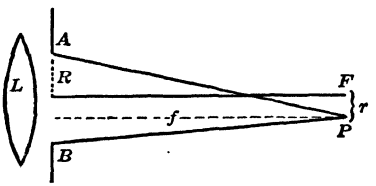


FIG. 175

Diffraction by a Circular Aperture. — A rigorous mathematical treatment of the formation of the central bright disk surrounded by rings was given by Airy and will be found in earlier editions of this book. The following elementary treatment gives an approximate solution and is easier to visualize. In Fig. 175 we have a lens L illuminated by parallel rays from a distant point source

[1] *Phil. Mag.*, Nov., 1910.

[2] These gratings with various spacings and concentrating in various regions of the infra-red are now being ruled on one of the Rowland engines, on disks of copper 5 inches in diameter and 1 inch thick, worked to optical flats and polished. The surface is plated with chromium after the ruling is finished. They can be supplied to laboratories desiring them.

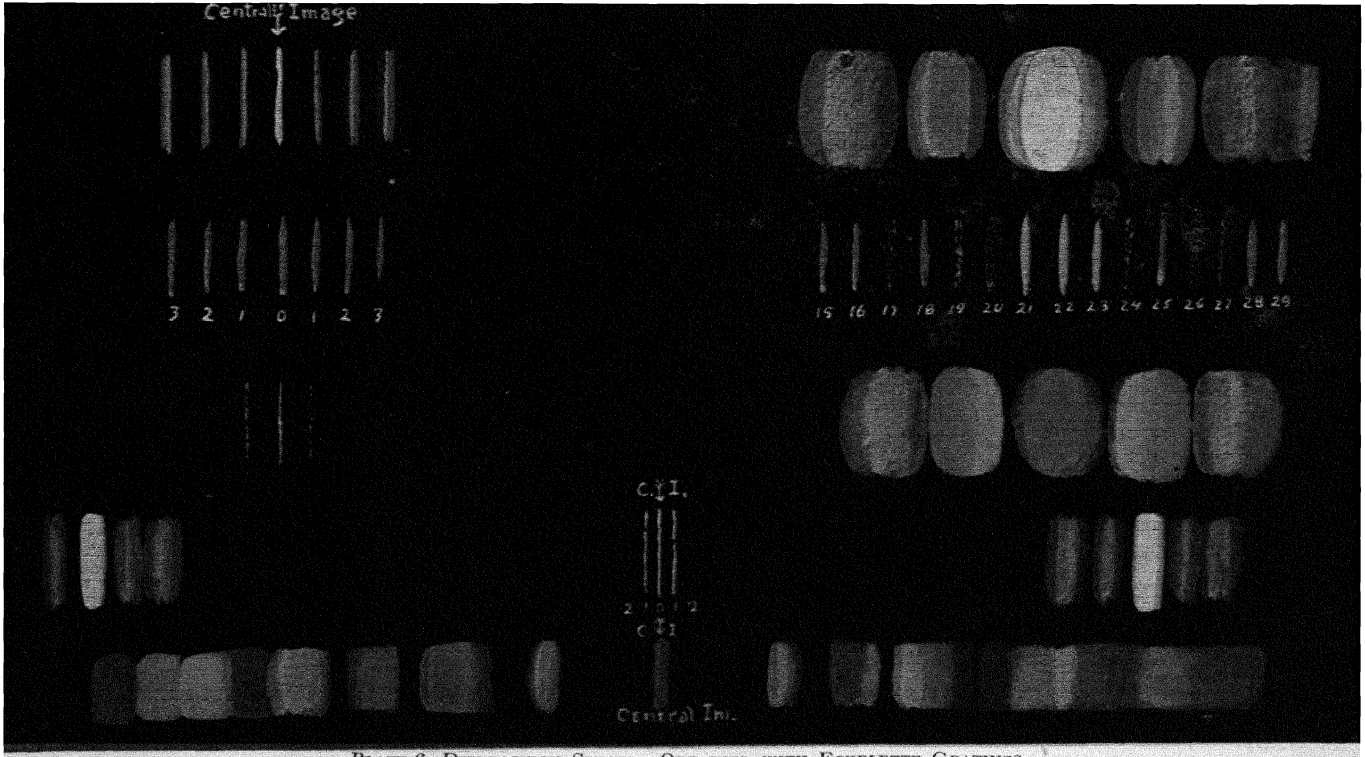


PLATE 6. DIFFRACTION SPECTRA OBTAINED WITH ECHELETTE GRATINGS.

which focus at F. We require the illumination at P situated at distance r from F. If R is the radius of the circular aperture behind the lens and f the focal length we have the optical paths

$$AP^2=f^2+(R+r)^2$$
$$BP^2=f^2+(R-r)^2$$
$$\therefore\ AP^2-BP^2=4Rr$$

and the path-difference $AP-BP=2Rr/f$ since r is small compared to f, and the illumination will be zero or a maximum according as

$$\frac{2Rr}{f}=n\lambda \text{ or } (n+\tfrac{1}{2})\lambda.$$

Rotation of the figure around the axis represents the illumination as a bright disk surrounded by bright and dark rings, which become smaller as the diameter of the aperture is increased. As a result the images of stars appear smaller in telescopes of large aperture than in smaller instruments. The intensities for the disk and first two rings are 1, .017 and .0041.

Resolving-Power of Telescope. — The images of two stars can be seen separated if the central spot of the diffraction pattern of one falls at or beyond the first minimum (*i.e.* dark ring) of the image of the other. Let R be the radius of the telescope's aperture. The diffraction angle θ for the first minimum is given by

$$\sin\theta=0.61\frac{\lambda}{R}.$$

The angular distance between two stars must therefore be greater than θ, as defined above, if they are to be seen separated, *i.e.* we must have the angular separation $\Phi>.61\lambda/R$ (writing Φ for $\sin\Phi$).

Calling $\lambda=.00056$ mm. and expressing Φ in minutes, we have $\Phi>1.17'/R$. A telescope of 200 mms. aperture will therefore resolve a double star with an angular separation of $.0117'=.7''$. The equation shows us that the angular separation of two stars which can be separated by a given lens is roughly equal to the angle subtended by the wave-length of light at a distance equal to the diameter of the lens.

In some cases it becomes necessary to reduce the aperture of the telescope, as when viewing a very brilliant object such as the sun or the brighter planets. This is usually accomplished by covering the lens with a screen perforated with a circular hole. The resolving-power is immediately reduced, and if the reduction is carried too far all definition disappears.

Lord Rayleigh has pointed out that a better method would be to cover the central portion of the lens with a circular opaque screen.

This would permit of the interference of the rays from the outer zones of the lens and the resolving-power would not be reduced. In fact calculations showed that an actual improvement resulted, the diameter of the first dark ring around the central bright spot being less than with the full aperture.

Babinet's Principle. — This principle is one which is applied to complementary diffraction screens, by which we mean a pair of screens in which the transparent portions of one are replaced by opaque portions in the other, and *vice versa*. An example would be a metal plate with a number of small circular apertures and a glass plate with metal disks of similar size and distribution. Babinet's principle states that the diffraction patterns are the same in each case. This we can see from the following considerations:

In the case above the illumination at a point M on the screen, where the parallel diffracted rays of diffraction angle δ from the collection of circular apertures come together, is represented by the sum of the squares of two integrals taken over the areas of the apertures. This we will call $A_1^2+B_1^2$. In the same way the illumination at the same point due to the collection of disks is $A_2^2+B_2^2$. If the two sets of disturbances act simultaneously, *i.e.* if the wave is disturbed by no screen, the illumination is zero, provided the point M is situated at some point not coincident with the point at which the wave comes to a focus; in other words, no diffraction effects are produced. This means that the resultant of one set of disturbances is able to exactly destroy the resultant of the other set, or

$$(A_1+A_2)^2+(B_1+B_2)^2=0, \text{ or that } A_1=-A_2 \text{ and } B_1=-B_2.$$

The illumination is, therefore, the same in the two cases, and the only result of changing the screens is to alter the resultant phase by 180°.

The principle of Babinet cannot be applied universally to all diffraction problems, for example the circular aperture and disk in the Fresnel class, one of which gives maxima and minima along the axis, the only other a maximum. Its application is restricted to points lying outside of the projection of the aperture, where the illumination due to the whole aperture is zero.

Diffraction by Two Small Apertures. — In the case of two small circular apertures close together we have interference between diffracted rays as in the case of two slits; parallel dark bands cross the ring system perpendicular to the line joining the two apertures. By Babinet's principle we may substitute two circular disks for the apertures, though the phenomenon is difficult to see in this case. The author has observed it in viewing a small and very brilliant

lamp filament in a horizontal mirror of plate glass, lightly dusted with lycopodium, at a fairly large angle of incidence. Each particle and its reflected image form a vertical pair, and the dark fringes cross the rainbow-colored rings horizontally.

Halos. — The halos which are sometimes seen surrounding the sun or moon are due to diffraction by small drops of water, which by Babinet's principle will produce the same effects as small circular apertures of the same size. The smaller the drops the larger the halos, but we distinguish between the diffraction halos, which are always close to the sun, and the large rings due to ice spicules floating in the air. These halos can be imitated by viewing a candle flame or other source of light, through a glass plate, on the surface of which lycopodium has been dusted, or better, by viewing the light through a large glass flask, wet on the inside and connected with an air pump. On partially exhausting the flask with one or two strokes of the pump a cloud forms in the flask, and the light is seen to be surrounded by brilliantly colored rings.

When the halo is produced at a great distance, as is the case in the atmospheric phenomenon, instead of by particles immediately in front of the lens of the eye or telescope, the complete ring system as seen is of course not produced by each individual particle.

The production of the colored ring is illustrated in Fig. 176. The broad arrow indicates the direction in which the sun is seen by an eye at X. A, B, C, D, etc., are small globules of water. The dotted arrows represent the directions of the diffracted rays, giving the first maximum to the left of the central maximum for the blue rays, the long solid arrows the directions of the diffracted rays for the green, and the short arrows for the red. It is obvious from the diagram that the particle D will send blue light to the eye, the particle E green light and the particle F red light. The phenomenon in space will be represented by rotating the diagram on AX as an axis. Each particle of water thus forms an infinitesimal element of the halo. If the particles vary in size in different parts of the sky, the angles of diffraction will vary also, and we may thus have a halo which is not a perfect circle.

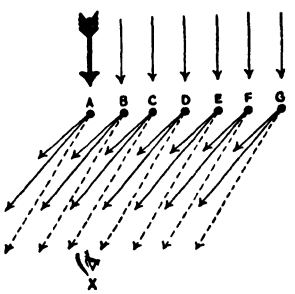


FIG. 176

In the same way a cobweb in the sunshine sends approximately monochromatic light to the eye, the color depending on its angular position, and a plane diffraction grating at a distance of eight or

ten feet from the eye appears illuminated in light of a uniform color.

Young's Eriometer. — The dependence of the diameter of the halo on the size of the diffracting particles was utilized in an ingenious piece of apparatus devised by Young for measuring the diameters of fibres, or small particles of any sort. It consists of a metal plate with a small hole 5 mms. in diameter, surrounded by a circle of smaller holes about 1 cm. in radius. The plate is placed in front of a lamp flame, and viewed through the particles or fibres to be measured, which are best spread out on a glass plate. The halo surrounding the central aperture can be brought into coincidence with the circle of small holes by varying the distance of the screen, which can be done by sliding the plate carrying the particles along a graduated rod, on the end of which the diffracting screen is mounted, the distance varying inversely as the diameter of the halo, which in turn varies inversely as the diameter of the particles. The constant of the instrument is determined by making an observation with particles of known size. If d is the distance between the screen and the particles of known radius r, when the halo is in coincidence with the ring, and d' is the distance for particles of unknown radius r', we have

$$\frac{r'}{d'}=\frac{r}{d} \text{ or } r'=r\frac{d'}{d}.$$

Diffraction by Small Particles on the Surface of a Mirror. — The so-called diffusion rings observed when a small source of light is viewed in a silvered glass mirror, the front surface of which is slightly dimmed with a deposit of dust, such as lycopodium, are in reality diffraction phenomena. They are sometimes erroneously attributed to the interference of diffused light, and Stokes was the first to treat them as diffraction effects.

We have here a case of the interference of the secondary disturbances from a particle interfering with the reflected secondary disturbances from the same particle, the path-difference depending on the thickness of the glass plate and its refractive index. By a suitable arrangement of the apparatus employed for viewing them, they may be brought under the Fraunhofer class of diffraction phenomena. The method is due to Lommel. A full treatment is given in earlier editions.

Diffraction in Relation to the Formation of Optical Images. — Diffraction plays a very important rôle in the functioning of all optical instruments and we will commence our study of the subject by considering the simplest possible case — the formation of an optical image by a simple aperture without the use of either lenses

or mirrors. The only practical use that has been made of this method of image formation is in pin-hole photography in which the lens of the camera is replaced by a pin-hole in a piece of thin metal foil. The very clear and illuminating treatment of this case we owe to Lord Rayleigh, who showed that in certain cases an aperture acts quite as well as a lens, expressing the relation as follows:

"The function of a lens in forming an image is to compensate by its variable thickness the differences in phase which would otherwise exist between secondary waves arriving at the focal point from the aperture. If we suppose the diameter of the aperture or lens ($2r$) to be given, and its focal length f to increase gradually, these differences of phase at the image of an infinitely distant luminous point diminish without limit. When f attains a certain value, say f', the extreme error of phase to be compensated falls to $\lambda/4$. Now, as I have shown on a previous occasion, an extreme error of phase amounting to $\lambda/4$, or less, produces no appreciable deterioration in the definition; so that from this point onwards the lens is useless, as only improving an image already sensibly as perfect as the aperture admits of. Throughout the operation of increasing the focal length, the resolving-power of the instrument, which depends only upon the aperture, remains unchanged; and we thus arrive at the rather startling conclusion that a telescope of any degree of resolving-power might be constructed without an object-glass, if only there were no limit to the admissible focal length. This last proviso, however, as we shall see, takes away almost all practical importance from the proposition.

"To get an idea of the magnitudes of the quantities involved, let us take the case of an aperture of $\frac{1}{5}$ inch (inch = 2.54 cms.), about that of the pupil of the eye. The distance f_1, which the actual focal length must exceed, is given by

$$\sqrt{\{f_1^2+r^2\}}-f_1=\tfrac{1}{4}\lambda,$$

so that

$$f_1=\frac{2r^2}{\lambda}.$$

Thus, if $\lambda=\frac{1}{40000}$, $r=\frac{1}{10}$, $f_1=800$.

"The image of the sun thrown on a screen at a distance exceeding 66 feet, through a hole $\frac{1}{5}$ inch in diameter, is therefore at least as well defined as that seen direct. In practice it would be better defined, as the direct image is far from perfect. If the image on the screen be regarded from a distance f_1, it will appear of its natural angular magnitude. Seen from a distance less than f it will appear magnified. Inasmuch as the arrangement affords a

view of the sun with full definition and with an increased apparent magnitude, the name of a telescope can hardly be denied to it.

"As the minimum focal length increases with the square of the aperture, a quite impracticable distance would be required to rival the resolving-power of a modern telescope. Even for an aperture of four inches f_1 would be five miles."

Returning to the subject in a later paper, Lord Rayleigh discusses its application to the so-called pin-hole photography, in which the lens of a camera is simply replaced by a small aperture. "If this aperture is too small, the image loses in definition owing to the spreading out of the waves, and on the other hand it is clear that no image can be formed, when the aperture is large. There must therefore be one particular size of the opening which gives the best result. The best result in general is found, when the aperture as seen from the image includes about nine-tenths of the first Fresnel zone, so that if a is the distance of the object, b that of the image from the screen and r the radius of the opening,"

$$r^2 \frac{a+b}{ab} = .9\lambda.$$

An interesting way of showing the formation of an image by a simple aperture, on a rather large scale, using an aperture 15 mms. in diameter, is as follows: Silver an interferometer plate or other piece of optically flat glass and cover it with a piece of paper perforated with a hole 15 mms. or so in diameter.

Mount the mirror on some sort of adjustable support (a heliostat, if one is available), and reflect sunlight from its surface into a moderately dark room through an open window or door from a distance of say 200 feet, receiving the image of the sun thus formed on a white wall or large sheet of paper. The best diameter for the mirror, calculated from Lord Rayleigh's formula for pin-hole photography would be about 6 mms., and if the room can be sufficiently darkened, this dimension can be tried. The aperture is best made elliptical so that, as seen from the direction of the reflected ray, it appears circular. One obtains in this way a surprisingly sharp image of the sun nearly 2 feet in diameter, showing the spots surrounded by their penumbrae and other details.

If a long corridor (which can be darkened) with a window at one end is available, the experiment is still more striking. A familiar example of image formation by small apertures is the appearance of solar images on the ground under trees, the apertures being chinks between the leaves. During an eclipse crescents are of course seen.

Image of Single and Double Apertures: The Edge-Wave. — The diffraction of light plays an important rôle in the formation of images of objects which have fine structure, especially in the case of the microscope as was first shown by Ernst Abbé over half a century ago, and the problem is usually treated from this point of view, the case chosen being that of a diffraction grating and the changes produced in the appearance of its image when certain ray systems in the transmitted light are blocked off by screens. We shall, however, gain a much clearer conception of the physical processes involved, and how they operate if we consider a simpler case, namely the formation of an image of a rectangular aperture (or two adjacent narrow apertures) by a lens. We shall see that the formation of a true image by a lens occurs only when the lens transmits the diffracted light as well as the direct.

As our object we shall take a vertical rectangular aperture measuring 9×80 mms., the upper half divided into two equal parts by a wire 3 mms. in diameter. The upper part is thus a diffraction grating of two elements in which spectra of even order are absent. To secure the proper ratio of widths it is best to form the aperture from three short lengths of steel rod 3 mms. in diameter, fastened across a narrow aperture in a card with sealing wax, using two extra bits of the rod to secure the proper spacing, which must be quite exact if the even order spectra are to disappear.

We must now illuminate, and observe (or photograph) our object under conditions which permit of our screening any portion or all of the diffracted light. The experiments made by the author for the purpose of illustrating this subject were made on rather a large scale with two lenses of about 20-foot focus. It can be made on a smaller scale employing two large spectacle lenses of as long focus as possible (or a single lens of about 6-foot focus), with the object reduced in its vertical dimension.

A mercury arc is placed behind a narrow slit covered with a green filter, and the emergent light, rendered parallel by one of the lenses, is brought to a focus by the other, the card with the wire aperture being placed between the lenses (or in front of the single lens). Diffraction maxima and minima will be seen at the focus. If the double aperture is covered we have the Fraunhofer fringes of the first class formed by a single aperture, while if the single aperture is covered we have the second class of fringes or grating spectra. With both apertures uncovered we have the two systems superposed.

We will consider now the case of the single aperture, placing a telescope directed at the aperture in such a position that the focussed fringes fall on the object-glass. We now focus the telescope

on the aperture, which appears uniformly illuminated with sharply defined edges in spite of the fact that most of the light is concentrated along a very narrow line on the object-glass, and if such a restriction of illumination were imposed under ordinary conditions by placing a screen perforated by an equally narrow slit in front of the telescope, the resolving-power would be impaired to such an extent that only a very fuzzy image of the aperture would be seen. This is the "Optical paradox" proposed by Lord Rayleigh [1] who showed that the sharpness of the image resulted from the passage into the telescope of the light of the faint maxima bordering the very bright central maximum. If these are screened off with a pair of cards so that only the light of the central maximum enters the telescope we see only a single broad image with ill-defined edges, and if we uncover the double aperture it cannot be distinguished from the single one, except that it is fainter. This shows clearly that the diffracted light must enter the lens if a true image is to be formed.

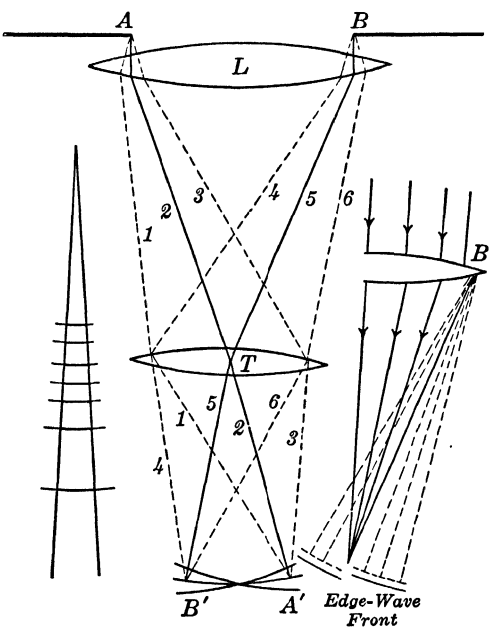


FIG. 177

Since the conditions of image formation are a little peculiar in the present case, it will be well to consider exactly what is happening. In Fig. 177 we have a diagram of our arrangement. AB is the aperture illuminated by parallel rays or plane-waves. These rays are brought to a focus by the lens L on the object-glass of the telescope T, as indicated by the solid lines 2, 5. We thus have an image of our distant narrow slit on the object-glass and the rays pass through the object-glass with practically no deviation as a diverging cone reaching the plane $B'A'$, which is the focal plane of the telescope *when focussed on* AB. With the reduced dimensions of the diagram, the image at $B'A'$ will be half the size of AB. We

[1] *Phil. Mag.*, ix, 779, 1905.

must now see how the image is formed. The diffracted rays, represented by dotted lines, are focussed at the edges of the telescope lens, rays 3, 6 corresponding to parallel rays leaving A, B in such a direction that their path-difference is $3\lambda/2$, this being the condition for the first lateral maximum. Rays 1, 4 form the maximum to the left. If we screen off these lateral maxima, we have only a blurred image at $B'A'$ for in this case the telescope lens *plays no part* in its formation. The convergent cone 2, 5 passes through it without deviation, the rays diverging from the focus point on the lens surface at the same angle as that through which they arrived. This can be shown by removing the lens T which does not alter the appearance of the image of the aperture. The same is true of the diffracted cones, except that their direction is changed by the lens, so that the rays 3, 6, converging on the lens instead of passing through the lens in their original direction, are bent to the left to position 3, 6 below the lens, the blurred image which they form falling also on $B'A'$. The small element of the lens through which the focussed cone passes can be regarded as acting as a prism. No such deviation is produced at the centre of the lens.

The image at $B'A'$ produced by any given diverging cone of rays is wider than would be predicted by geometrical optics since the diverging wave-front spreads out by diffraction as shown by the small diagram at the left of Fig. 177 the diffracted portion forming the widened and blurred edges of the image. When, however, the three images are superposed interference takes place and we get a fairly sharp image of the aperture. Rays 1, 2, 3, proceeding from A, are brought to a focus at A' by the lens, while rays 4, 5, 6 are focussed at B'. This is the ray method of accounting for the image. To account for it by wave-fronts and interference we superpose the three diverging waves coming from the three illuminated points on the lens as shown in Fig. 177 above $B'A'$ and find that, owing to their inclination with respect to each other (greatly exaggerated in diagram) they tend to reënforce each other between B' and A', while in the regions to the right and left a path-difference is established which results in destructive interference.

To obtain a complete picture of what happens between the lens and its focal plane, a series of photographs were made on an inclined plate 3×25 cms. enclosed in a narrow box, a cylindrical lens being used to prevent convergence in the vertical plane. In this way images of the intersecting ray cones and their interferences were recorded over their entire path from lens to the focus point. We must now determine the part played by the diffraction maxima (on each side of the central maximum) in the formation of the image.

When all of them operate (with central one excluded) we see in the telescope only the edges of the aperture illuminated, each edge appearing as a *double* line however. If only the first diffraction maximum operates we have a broad ill-defined image. If the first and second are passed this image becomes double, a broad dark minimum bisecting it. At first sight this seems difficult to account for since the two sources (the two lateral maxima) are supposedly in phase, and we should therefore rather expect the centre of the image to be enhanced in intensity and the edges diminished. If, however, we consider the diffracted wave-fronts which produce these maxima we find that they are half a wave-length apart *at their centres.*

The path-difference for the first lateral maximum, Fig. 178 for the diffracted rays (wave-front $A'B$) from A and B the edges of the aperture is $3\lambda/2$, and for the second maximum $5\lambda/2$ (wave-front $A''B$), therefore the distance $A'A''=\lambda$ and the distance between the *centres* of the diffracted fronts is $\lambda/2$. This relation holds after the two fronts have come to a focus on the second lens and are diverging on the other side, consequently when they are brought into coincidence at the focal plane of the lens destructive interference occurs at the centre, and reënforcement at the edges. The image of the aperture therefore has a dark centre with bright edges. A very simple explanation of the somewhat surprising circumstance that each luminous edge of the aperture appears as a *double* line is as follows:

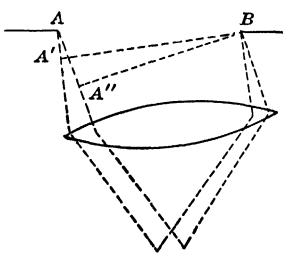


FIG. 178

The rigorous mathematical treatment of diffraction by Sommerfeld shows that there is a phase-difference of π between that portion of the cylindrical wave which diverges from the edge into the shadow region and the portion spreading out into the illuminated region as was pointed out by S. Banerji [1] who observed and explained the formation of the double line of light about to be described. We can thus regard the "edge-wave" (rand-welle) as having a dislocation of $\lambda/2$ along the line of the geometrical edge of the shadow, in other words analogous to a plane-wave, one-half of which has passed through a plate giving a half-wave retardation. Such a wave, when brought to a focus by a lens gives a double image as will be explained in the section dealing with Talbot's fringes. This can be demonstrated by viewing a distant point or vertical line of light through an aperture 0.5 mm. in width

[1] *Phil. Mag.*, *37*, 112, 1919.

one-half of which is covered by a transparent thin lamina of suitable thickness. Referring back to Fig. 177 we see that rays 4 and 6 belong to the edge-wave from B which is brought to a focus at B', the portion 4 having been diffracted into the region of illumination (cone 2, 5) and the portion 6 into the region of shadow. The $\lambda/2$ path-difference between them accounts for the fact that the image of the luminous edge (shining by diffracted light) and formed at B', is *double*.

The relation of the "edge-wave" to the ray diagram of Fig. 177 is shown at the right of the figure, in which the incident parallel rays are shown brought to a focus on the narrow screen which prevents their reaching the lens T. The rays of the edge-wave diverging from B and practically uninfluenced by the lens L are represented by dotted lines, and the wave-front, with its $\lambda/2$ dislocation at the edge of the shadow (*i.e.* the edge of the converging cone) is indicated. This may assist the reader to visualize the somewhat complicated effects with which we are dealing. If the lens T passes the light of all the maxima on one side of the central image (which is screened off) the edge of the aperture appears as a narrow line of light. If only the maxima on the other side are passed, the line of light appears, but displaced slightly from its previous position. If both sets of maxima are passed, two lines of light are seen at each edge of the aperture, that is the true edge is dark but bordered by two narrow luminous lines. If only the first maxima to the right and left are passed, the double line is seen but the components are broader and fainter. With a single aperture (narrower than the one employed) we thus find that, under certain restricted conditions of illumination, we have an image consisting of four bright lines.

If we place a lens at some distance behind a straight edge diffracting screen with an eye-piece (or photographic plate) at such a distance from the lens as to give a focussed image of the diffracting edge, we see the edge as self-luminous (by the edge-waves) which seems to show that they have a real existence, or rather that the lens interprets the interference wave-field as if made up of plane-waves and cylindrical waves centred on the edge.

The effects seen with the two slits are best studied by a repetition of the experiment.

Abbe's Diffraction Theory of Microscopic Vision. — The diffraction theory of microscopic vision was proposed over half a century ago by Ernst Abbe. It may be briefly stated as follows: For the production of a truthful image of an illuminated structure by a lens, it is necessary that the aperture be wide enough to transmit the whole of the diffraction pattern produced by the structure.

If but a portion of the diffraction pattern is transmitted, the image will differ from the object, and will correspond to an object the entire diffraction pattern of which is identical with the portion passed by the lens. If the structure is so fine, or the lens aperture so narrow, that no part of the diffraction pattern is transmitted, the structure will be invisible, no matter what magnification is used. Abbe devised a number of interesting experiments to illustrate the theory. By means of suitably perforated screens placed within or above the objective of a microscope, one or more of the diffraction spectra produced by a glass grating (the object viewed through the microscope) were cut off, and the appearance of the grating shown to be modified in a most remarkable manner. If all of the spectra were screened off, all trace of the lines vanished. It is sometimes assumed that there is an essential difference between microscopic and ordinary vision, and that the phenomena of diffraction play no part in the latter. Whatever difference there may be between the two cases, arises, however, from the relative size of the objects involved, and from the special methods of illumination employed with the microscope, as has been shown by A. B. Porter, who devised the following ingenious experiment for demonstrating that the images of periodic structures formed by the naked eye itself are due to diffracted light. Light from an arc lamp or the sun passes through a pin-hole in a screen and is focussed by means of a photographic lens on a cardboard screen, about 30 cms. from the lens — immediately in front of which a piece of wire gauze having about thirty wires to the centimetre is placed. The diffraction pattern produced on the screen by the wire gauze consists of a central image with a large number of radiating spectra surrounding it. There are two sets of spectra at right angles to each other, formed by the two sets of wires, with two intermediate sets also at right angles, but rotated through forty-five degrees with respect to the others. These latter may be regarded as the spectra of spectra, and are always seen when a source of light is viewed through a pair of crossed diffraction gratings. (See Fig. 170.) By cutting small holes in the screen we may transmit any portion of the diffraction pattern and allow it to enter the eye. If the screen is pierced by a hole only large enough to transmit the central image, the wire gauze is quite invisible. If a narrow slit is used which transmits only the central image and the horizontal line of spectra, the vertical wires alone are seen; if the slit is turned vertically so as to transmit the vertical line of spectra, the horizontal wires alone are visible. If the slit is turned at an angle of forty-five degrees so that the diagonal set of spectra are transmitted, neither the vertical nor horizontal wires are seen, but

a very real-looking set of wires appears running diagonally in a direction perpendicular to the slit. Such a set of wires would, if acting alone, give rise to the spectra transmitted. If the card is pierced with three pin-holes which transmit the central image and two second order spectra, a set of vertical wires is seen, twice as close together as the actual wires. By cutting two slits in the screen at right angles to each other, and arranging them so as to transmit the diagonal spectra, we see the gauze as if rotated through forty-five degrees.

In these experiments the object of placing a lens between the gauze screen and the eye is to enable us to remove any desired group of parallel diffracted rays. This can only be accomplished by bringing all of the parallel sets each one to its own focus, and intercepting them at this point. A white card should be mounted at the point where the spectra appear sharpest, and the size and shape of the aperture desired for the transmission of any portion of the diffraction pattern traced upon it with a sharply pointed pencil.

These experiments are similar to the ones devised by Abbé, and furnish a very easy and convenient means of illustrating his theory.

The Colors of Mixed Plates. — Interference colors of this type were discovered by Young, and described in the *Philosophical Transactions* for 1802, and were subsequently studied by Sir David Brewster. Verdet and other writers on Optics have classified them with Newton's thin-film colors, and have given treatments which are not very rigorous, and fail to show where the energy goes to.

The colors are very easily obtained by pressing a little white of egg between two pieces of plate glass, separating the plates and squeezing them together a number of times so as to form a froth. The plates are to be pressed firmly together with a rotary sliding motion just before the froth becomes sticky, enclosing a film made up of air and albumen in the form of a mosaic. The colors are best seen by holding the plate towards a distant window or other bright source of light on a dark field. Certain wave-lengths will be found to be absent in the directly transmitted light.

In Fig. 179 let AA represent the glass plates with the albumen and air elements between them. If the albumen elements give a half-wave retardation for green light, the rays passing through them, when brought to a focus by B, the lens of the eye will interfere with the rays passed by the air elements and we shall have an absence of green light at E. The source will therefore appear purple. At F, in an oblique direction, where the illumination would be zero in the absence of the albumen mosaic, the phases agree for green light, and we have a green halo or ring, the diagram being rotated about the axis of the lens. The distribution of the light in

the halo depends on the form of the elements of the mosaic. By pressing the plates firmly together and sliding one over the other, the circular air-bubbles can be deformed into ellipses. The light in the ring will be more or less concentrated on opposite sides of the halo. If the ellipses were drawn out indefinitely, we should pass over to the grating, and the points of concentration would become first order spectra, the rest of the halo disappearing.

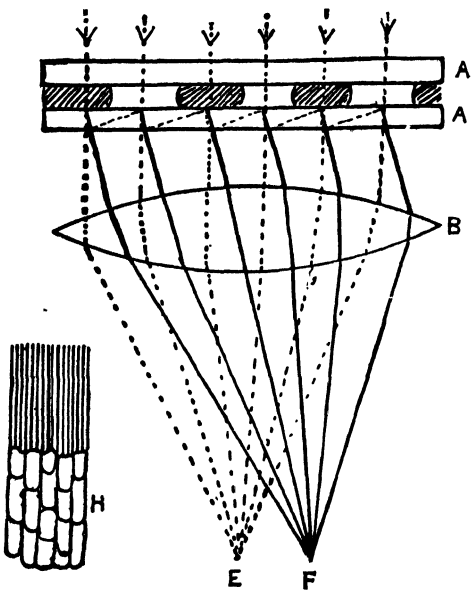


FIG. 179

A very curious and interesting example of this concentration of light in a halo was observed by the author when copying some diffraction gratings on bichromatized albumen. The original grating was ruled on glass, 14,400 lines to the inch, a spacing so fine that copies were only obtained with considerable difficulty.

Some of the films were found to have frilled in the process of washing, the buckling of the film following the grooves of the grating to a certain extent. The albumen surface was seen by the microscope to have frilled into oval patches of varying length, but of fairly constant width, the width being equal to three lines of the original grating. In Fig. 179, *H*, we have a diagram illustrating this condition. This plate when held before the eye showed a ring of wide aperture surrounding a brilliant source of light, with four distinct concentrations, two very bright, and two quite faint. The photograph, Fig. 180, was made with this plate. The appearance reminded one most forcibly of a solar halo with parhelia or mock suns.

Talbot's Bands. — A curious type of interference bands was described by Talbot [1] in 1837. If, when viewing a continuous spectrum in a spectroscope, we place a thin piece of glass or mica in front of one-half of the pupil of the eye, with its edge towards the red end of the spectrum, a set of extremely black bands appears,

[1] Gilbert, *Ann.*, 1837.

crossing the spectrum parallel to the Fraunhofer lines. If, however, the thin plate be turned with its edge pointing towards the violet, no trace of the bands is to be seen. A thin cover-glass, such as is used for microscope slides, answers the purpose. If the glass is placed in front of the telescope lens it must be introduced from the side towards the blue end of the spectrum.

Talbot gave an imperfect explanation of the bands on the elementary principles of interference. The thin plate retards certain colors an odd number of half wave-lengths. These waves arrive at the retina in a condition to interfere destructively with the waves which enter the uncovered portion of the pupil, the lens of the eye bringing them to the same point, consequently these colors are absent in the spectrum. The distinctness of the bands depends on the thickness of the plate, and Talbot's explanation neither accounts for this nor for the apparently remarkable circumstance that the bands do not appear at all when the plate is turned with its edge the other way, covering the other half of the pupil, a change which should produce no effect on the interference theory; neither is it quite clear what becomes of the light. The correct explanation was first given by Airy.[1] His treatment involves so much mathematics that it is difficult to follow the physical significance of the different steps.

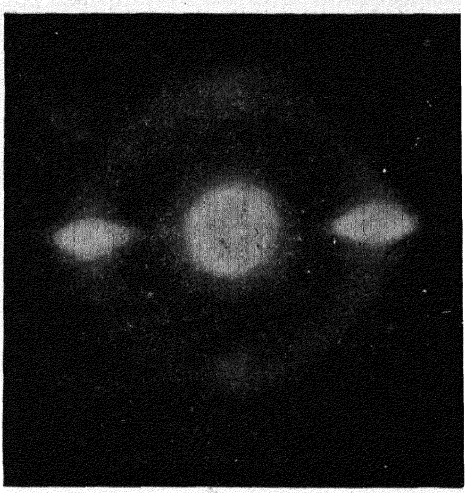


FIG. 180

The following elementary treatment will make clear just how the bands are formed, and why they appear only when the plate is introduced on one side. Compare this case with the laminary grating.

Incidentally we shall find that the echelon grating, which will be described presently, is a special case of Talbot's plate, the aperture partly covered by the thin plate forming an echelon grating of two elements. To see the bands at their best, the object-glass of the spectroscope should be covered with a screen perforated with an

[1] Airy, *Pogg. Ann.*, liii–lviii; *Philos. Trans.*, 1840, vol. ii; 1842, vol. i.

aperture about 1 cm. in height by 5 mms. in width. The glass or mica plate can be mounted so as to cover the right- or left-hand half of this aperture, or simply held in front of the eye.

Let AB, Fig. 181, be the aperture, one-half of which may be covered with the retarding plate.

In the absence of the plate, the lens will form an image of the slit (which for the present we will consider illuminated with monochromatic light), bordered by faint maxima and minima. For the present we consider the prism removed, and the telescope directed in line with the collimator. The intensity curve, which represents the image of the slit diagrammatically, is shown at I.

The first diffraction minima to the right and left of the central maximum will lie in directions such that the path-difference between the disturbances coming from the edges of the aperture is a whole wave-length. If the right-hand half of the aperture is covered

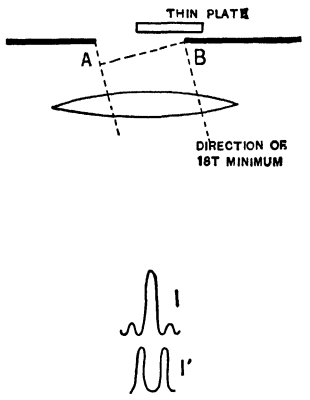


FIG. 181

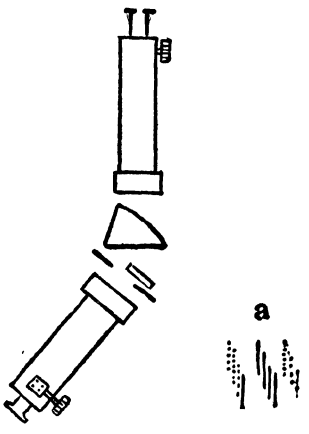


FIG. 182

with a transparent plate which retards the wave-train one-half wave-length, we shall have zero illumination at the centre, and two bright maxima of equal intensity in the position previously occupied by the first minima, as shown at I'. It is clear that the path-difference for the left-hand maximum is one wave-length, and for the right-hand one zero. We may therefore call the former the spectrum of the first order, and the latter the spectrum of zero order, considering the aperture as a grating of two elements. If we consider the change as taking place progressively, starting with a transparent lamina of infinite thinness which gradually thickens, it is clear that the central maximum moves to the right, decreasing in intensity, while its neighboring maximum to the left

increases in brightness moving in the same direction. When the thickness corresponds to a half-wave retardation these two maxima are of equal intensity, and occupy the positions previously occupied by the minima to the right and left of the central maximum. If we increase further the thickness of the plate the maxima drift further to the right, the left-hand one increasing in brilliancy, the right-hand one decreasing. When the retardation becomes one whole wave-length, we again have a single bright maximum at the centre, as with the uncovered aperture. In this case, however, it is the spectrum of the first order which is at the centre. As we go on increasing the thickness of the plate the march of the fringes continues, new ones coming into view on the left and disappearing on the right, and the order of spectrum increasing. The condition when we have a pair of bright maxima will be recognized as the condition which obtains when an echelon grating is set for the position of double order, the single order position corresponding to the case when we have a single bright maximum at the centre.

Suppose now that we have a thicker retarding plate and gradually decrease the wave-length of the light which illuminates the slit from which the plane-waves are coming. We shall have the same march of the fringes as before, the gradual increase in retardation resulting from the decrease in wave-length. If our source emits two wave-trains of different wave-length, and the retardation for one of the trains is an odd, and for the other an even number of half wave-lengths, it is clear that the single and double maxima exist simultaneously, and all trace of the maxima and minima will disappear owing to the form of the intensity curve; in other words, the lines are not resolved.

Consider now the slit of a spectrometer illuminated with monochromatic light, the prism, aperture, and retarding plate arranged as shown in Fig. 182. If the plate retards the stream an even number of half wave-lengths, we shall see in the instrument a single bright line with accompanying faint maxima and minima. If we now decrease the wave-length a trifle the line will move a little to the right, if we disregard the action of the prism. We can verify this experimentally by removing the prism and viewing the slit directly with the telescope. The slit can be illuminated with a monochromatic illuminator (spectroscope with narrow slit in place of the eye-piece), the slit of the instrument being substituted for the slit of the first spectroscope.

As we gradually decrease the wave-length by turning the prisms of the illuminator, we shall see the line become double and single in succession, the doubling being accomplished by the march of the fringes to the right and their fluctuating intensity in the manner

already described. It is a little difficult to put into words the changes which accompany the alteration of wave-length. The appearance is much like that presented by a picket fence in motion viewed through a narrow vertical aperture, *i.e.* the pickets are only visible as they pass across a narrow vertical region. If we imagine the visibility of the pickets to be a maximum as they cross the centre of the aperture and least when at the edges, we shall have a fairly accurate conception of the appearance of the moving maxima. Suppose that a change of 100 Ångström units in the wave-length is sufficient to change the single line to the double line. The changes which take place in the focal plane of the telescope (prism removed) as we decrease the wave-length over this range are indicated in Fig. 182*a*. To avoid confusion each successive appearance is represented a little lower down in the figure.

Now consider the prism in place, turning the telescope so as to view the deviated image. As we decrease the wave-length there will be a shift to the *right* as before, resulting from the retardation of the plate, and a shift to the *left*, due to the increasing deviation produced by the prism. If the compensation is perfect, the prismatic dispersion will shut up the picket-fence arrangement of lines in the diagram into single lines; in other words, if we illuminate the slit simultaneously with all of the wave-lengths in the given range of 100 Ångström units, we shall not see a short continuous spectrum, but single bright lines, to the intensity of which all the waves contribute. These lines are of course much narrower than the corresponding range of the continuous spectrum which would be formed in the absence of the plate. We thus see that the retardation has the effect of compressing a narrow region of the spectrum into a much narrower one which constitutes one of the bright Talbot bands.

The Echelon Grating. — We have seen that the resolving-power of a grating is represented by the product of the number of lines and the order of spectrum. The brilliant idea occurred to Michelson to secure high resolution by increasing the order of spectrum instead of the number of lines. The order of spectrum is the path-difference measured in wave-lengths between two adjacent elements of the grating, and this can be increased by the introduction of a glass plate. The echelon consists of a number of rectangular blocks of glass all cut from a single plane-parallel plate of optical glass, and arranged in the form of a flight of steps, each step forming a line of the grating, as shown in Fig. 183 in which the grating has 11 lines, one of them being a clear space at the edge of the widest plate. It is clear that the retardation of disturbances from the elements increases by equal amounts as we pass across

the grating, say 10,000 waves at each step, giving us a spectrum of the 10,000 order and a resolving-power of 110,000.

A very clear idea of the operation of the instrument can be obtained by following the transition from a diffraction grating made of parallel slits or wires to an echelon, and thence to the prism. Consider an aperture 1 cm. in width divided into ten elements by wires and covered by a lens which forms focussed spectra in the directions indicated by the lines, in which the angular divergence has been enormously exaggerated of course (Fig. 184). In

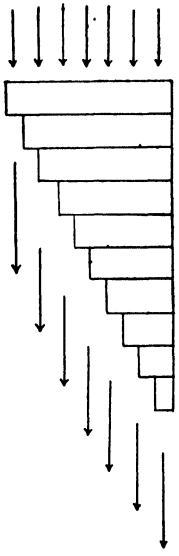


FIG. 183

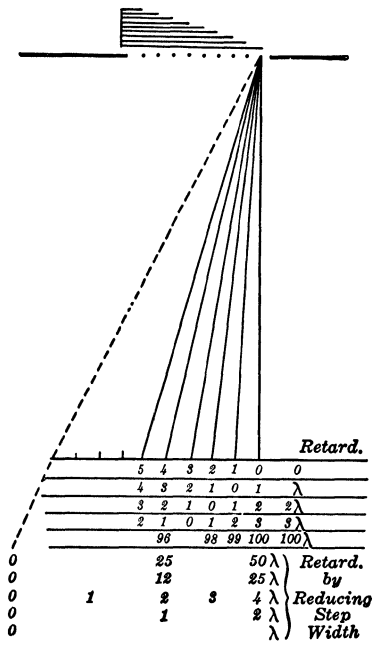


FIG. 184

reality, with a very coarse grating such as is here assumed there will be diffraction in directions inclined at only a very small angle to the right and left of the directly transmitted rays.

In the direction of the first order spectrum the path-difference between the edge rays will be 10λ. Now consider ten transparent laminae, each giving a retardation of λ, placed one above the other over the aperture forming a flight of steps, or an echelon as shown in the figure, the wires having been removed. The zero order spectrum now appears in the position previously occupied by the first order, since the thick end of the echelon gives a retardation of 10λ, which just compensates the former path-difference of

10λ in this direction. The old zero order direction is now that of the first order. If, instead of suddenly introducing the plates, we consider them as increasing from zero to a thickness giving a retardation of λ, the spectra will shift *gradually* to the left as shown in the treatment of the Talbot fringes.

We now consider the thickness of each plate doubled. The second order spectrum now appears in the normal direction, and the zero order moves one step more to the left, as shown in the figure. As we continue to increase the thickness of the plates, the zero order goes off to larger angles. Owing to the dispersion of the plates the retardation is greater for short λ's, and the shift to the left of the zero order spectrum (and all of the others) will be greater for blue than for red. This effect will be identified with the spectrum formed by the prism, into which we shall presently transform the echelon by making the width of the steps vanishingly small.

With plates retarding 100λ, we have the 100th order spectrum in the normal direction, and the 99th on its left and the 101st on its right. The resolving-power will now be 1000, just sufficient to resolve the *D* lines of sodium. A grating [1] of this type was made by the author out of a carefully selected sheet of mica, shortly after Michelson's announcement which proved of some value for demonstration purposes. With the exaggerated angles shown in the figure it is impossible to represent correctly the position of the zero order spectrum. As shown, we must consider that there are 99 spectra between it and the 100th order spectrum which is on the normal, the 50th order coming halfway between.

In point of fact the zero order spectrum lies in the direction of the ray refracted by a prism of 1 mm. base and 1 cm. width, into which we shall presently transform the echelon.

We have assumed that we have been operating with monochromatic light thus far, but now suppose that we have two wave-lengths differing by a very small quantity such that the *difference* of their retardations by one of the plates amounts to only half a wave-length. The spectra formed by this second radiation will fall halfway between those formed by the first, and we shall see twice as many lines as before, in other words the two radiations will be "resolved."

If the difference in wave-length of the two radiations were such as to give a difference of 3λ in their retardations the 97th order of one would coincide with the 100th order of the other, and we should be unable to differentiate them. This is one of the disadvantages of the echelon, but means of overcoming it will be described in the further treatment of the instrument in the Chapter on Interfer-

[1] *Phil. Mag.*, June, 1901.

ence Spectroscopes. In the case which has been given, the plates will each be 0.1 mm. thick (since retardation $=[(n-1)d]/\lambda=.5\times.1/.0005=100$) and the total thickness at the left 1 mm.

Transformation of Echelon into Prism. — We shall now gradually transform the echelon into a prism with a width of 1 cm. and a base of 1 mm., and see how its resolving-power compares with that of the echelon. We double the number of plates, halving the thickness so that the slope of the flight of steps remains the same and have therefore in the normal direction the 50th order spectrum, but the same resolving-power as before, as we now have 20 elements in the grating. The zero order spectrum remains fixed in position and the spectrum halfway to the normal drops to the 25th order spectrum, as shown in the figure. The angular separation of the spectra becomes twice as great as before, as this depends solely on the increment of path-difference due to obliquity. But owing to the decrease of one-half in the order of spectrum any spectrum of given order on the left, except the zero order has swung back towards the normal direction. We continue this process until the thickness of each plate reaches its original value (retardation of λ). The first order spectrum now appears in the normal direction, and we have 1000 steps, giving the same resolving-power, which is now due solely to the large number of lines. The zero order spectrum will lie to the left at an angular distance such as to make the path-difference between the edge rays 1000λ.

With a further reduction of step width the spectrum of the first order swings away from the zero order spectrum, moving to the right until, when we have reduced the step width to λ, the first order spectrum is at 90° to the normal, that is to say it vanishes and the stepped surface has become optically flat in so far as the wave-length with which we are operating is concerned. We now have 20,000 steps and only the spectrum of zero order, which is in reality the image formed by the prism into which the echelon has been transformed. As we have seen the position of the zero order is a function of λ, which accounts for the prismatic spectrum.

The base of the prism being 1 mm. thick it will resolve only lines separated by a distance of 10 times the distance between the *D* lines showing that the echelon has 10 times the resolving-power of a prism of equal base thickness. Echelons used in practice are usually about 40 cms. thick (*i.e.* 40 plates of 1 cm. each) and have, therefore, the resolving-power of a train of prisms with a total base thickness of 4 metres. This is 400 times the base thickness required for the *D* lines, which require a resolving-power of 1000, hence the resolving-power of the 40 plate echelons works out as 400,000 about the right order of magnitude, or more than that of a 6-inch, 21-foot

concave grating ruled with 30,000 lines to the inch in the second order spectrum.

Anomalous Propagation of Waves in the Vicinity of Foci. — The curious discovery was made by Gouy in 1890, that in its passage through a focus a wave gained a half wave-length, as if its velocity of propagation was accelerated in the vicinity of the focus.[1] He employed a pair of Fresnel mirrors, one of which was concave, the light falling upon them at nearly normal incidence.

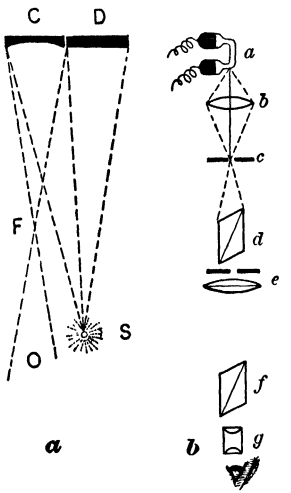


FIG. 185

The arrangement of the apparatus is shown in Fig. 185*a*. An illuminated pin-hole at *S* sends light to the two mirrors at *C* and *D*. *C* is a concave silvered mirror made by cutting a small square from a spectacle lens, *D* is a plane silvered mirror of the same size. The distance of the source *S* from the mirrors is 1.5 metres and the mirror *C* brings the rays to a focus at *F*, about 35 or 40 cms. from *C*. The fringes are observed at *O* with an eye-piece, or better with a microscope of low power (30–40 diameters). If there were no advance of the wave in its passage through the focus, we should have a fringe system with a white centre bordered on each side by colored fringes, as with the ordinary mirrors. What we see is a system with a black centre, with lateral fringes of colors complementary to those usually observed.

The fringes are circular since the two sources, one at *F*, the other at distance *SD* behind the plane-mirror, are practically in line. Gouy's experiments were repeated by Fabry [2] and later by Zeeman, who used a lens made of Iceland spar, placed between two Nicol prisms. This lens was doubly refracting and gave two real images of the source, at different distances. A similar arrangement subsequently used by Sagnac and by Reiche is shown in Fig. 185*b*. The light of a mercury lamp *a* is focussed upon a small aperture *c*, traversing the Nicol *d* and the doubly refracting lens *e*. The latter has focal lengths of 304 and 412 cms., and the distance between *c* and *e* was about 7 metres. The Nicols are crossed and the lens turned until its principal section makes an angle of 45° with the planes of the Nicols. The circular fringes are

[1] Gouy, *Ann. de Chim. et de Phys.*, 6e série, 24, pp. 145–213.
[2] Fabry, *Journ. de Phys.*, 3d series, 2, p. 22 (1892).

observed with a powerful eye-piece. With this arrangement we can see the fringes, not only between the foci, but in front of and behind them. If the diaphragm in front of the lens *e* has a diameter of 3 cms., we observe a system with a white centre about .5 m. in front of the first focus, *i.e.* between the focus and the lens; on moving the Nicol eye-piece back the fringes become indistinct until we arrive at a point about .5 m. behind the focus, when they again appear, this time with a dark centre, owing to the change of phase produced by the passage of the waves through the first focus. Behind the second focus we again find rings with a white centre. If the diaphragm is contracted to a diameter of 1 cm., it is found that as the focus is approached periodic changes of the centre of the system occur, and the fringe system can be followed through the foci, where *no sudden change occurs*. Sagnac explains the apparent anomalous propagation as a result of diffraction, and shows that we may have a somewhat similar phenomenon, when plane-waves pass through a circular aperture, without passing through a focus.[1]

We remember from our study of Huygens's principle that the resultant of the disturbances from a circular zone in the aperture at a point on the axis is retarded in phase with respect to a disturbance travelling from the centre of the aperture along the axis. The phase-difference between the two becomes less as we move the point away from the aperture. It is as if the phase of the resultant vibration from the circular zone was propagated along the axis with a velocity greater than that of a plane-wave, or greater than that of a disturbance from the centre of the aperture, for if it lags less and less upon the phase of the axial disturbance it must be travelling faster. This does not mean, however, that we are actually dealing with an abnormal velocity of light, for the direction of the axis does not coincide with the direction of propagation of the elementary waves from the circular zone. If we have a spherical wave starting from the circumference of the aperture we can regard the velocity with which its point of intersection with the axis travels, as the velocity of the resultant phase. It will be greater than the velocity of light in the vicinity of the aperture and equal to it at a great distance. Now the phase of the resultant at any given instant on the axis, of all of the disturbances coming from the aperture, depends upon the size of the aperture. The same is true of the illumination, which is zero for points so situated that the aperture contains 2, 4, 6, etc., Fresnel zones. If now we define the velocity of light as the velocity with which a given phase of vibration travels along the axis, it is clear that if we restrict the

[1] *Journ. de Phys.*, *2*, 721, 1903.

wave to the portion passed by the aperture, the velocity *as defined above* may be modified by diffraction.

We will now investigate exactly what happens in the case of a small aperture, and in the case of a wave passing through a focus.

Take first the case of a circular aperture, through which a wave-train has passed. If we are at a great distance from the aperture the resultant phase due to all of the disturbances will agree with that due to the one coming from the centre, since the path-differences are practically zero. This resultant phase will be accelerated 90° on the phase which we should have if the wave had not been restricted by the screen with its small aperture, as we saw in the Chapter on Huygens's Principle, for we must consider the secondary wavelets as starting with a phase one-quarter period in advance of the phase on the wave-front. If the phase which we should have at the point if the wave were unrestricted be represented by an arrow pointing vertically, the phase which we have with the small circular aperture can be represented by an arrow pointing to the right, retardations of phase being represented by counterclockwise rotations of the vector.

If now we approach the aperture along the axis the illumination will increase, reaching its maximum value when we reach a point such that the aperture contains one Fresnel zone. The phase will drift back 90°, however, as a result of the $\lambda/2$ path-difference between the disturbances coming from the edge and centre elements, so that it agrees with the phase given by the unrestricted wave (vertical arrow). Moving still nearer to the aperture, the illumination decreases, reaching the value zero when the aperture contains two zones: the phase suffers a further retardation of 90°, as we can easily see by compounding the resultant by the graphical method, the elementary vector arrows forming a closed circle at this point. This first arrow, representing the effect from the central element, points to the right. Succeeding ones, with increasing phase retardations, represent resultant effects from circular zones surrounding the central element. Just before we reach the point of zero illumination, the vector which gives us the resultant phase and amplitude is a short arrow pointing to the left. Passing the zero point, the phase springs back suddenly to the value which it had at the most distant point, that is, it is accelerated 180°. This same thing occurs over and over again as we approach the aperture. We find, however, that as we approach the aperture the minima are no longer equal to zero, that is the vectors do not form a closed circle, but a portion of a spiral. The phase at these points is the same as at the maxima, that is it is represented by the short vertical arrow joining the beginning and end of the turn of the spiral.

The maxima and minima gradually die out and we end up with a phase represented by a vertical arrow, that is retarded 90° on that of the distant point.

We thus see that as we pass from the aperture to the distant point we gain 90° in phase, the phase, however, oscillating back and forth through 180° a number of times, as was found in the experiment with the lens of Iceland spar.

This same thing takes place in the case of a wave passing through a focus. At the focus the path-difference between disturbances coming from the centre and edge of the wave becomes zero; that is, the focus corresponds to the distant point in the previous case, and we have the phase represented by an arrow pointing to the right. As we pass through the focus the phase arrow begins to turn down and we have similar oscillations, ending up with an arrow pointing downwards when we get so far from the focus that the maxima and minima have disappeared. The passage of the wave through a focus is thus seen to result in an acceleration of phase of 180°; we can regard the decrease in area of the wave-front as we approach the focus as analogous to restricting it by a small circular aperture.

CHAPTER VIII

INTERFERENCE SPECTROSCOPES

In the Chapter on Diffraction we have discussed the action of the diffraction grating, and we will now take up the subject of the more recently devised spectroscopes, which should have been treated in the Chapter on Interference, but which has been postponed for the reason that certain points cannot be well understood without previously considering the theory of the grating. We will begin with the Michelson interferometer, which is, perhaps, the best-known type.

The Michelson Interferometer. — Interference fringes were employed by Fizeau and others for the study of the composition of light in a crude way (for example the periodic disappearance of the fringes of sodium light in the case of Newton's rings) but A. A. Michelson was the first to construct an instrument of precision based upon interference, and to study systematically the fine structure of spectrum lines.

By his brilliant invention of the interferometer he opened up a wide field of research and furnished us with an instrument capable of showing in a most beautiful manner the interference phenomena of thin and thick plates, wedge-shaped or plane-parallel, which we have studied with simpler apparatus.

The essential parts of this instrument are four plates of glass arranged as shown in Fig. 186. Plates A and B are cut from the same piece of glass accurately plane-parallel. Both may be transparent or A may be half silvered on the surface opposed to B. Plates C and D are heavily silvered on their front surfaces. Plate D is mounted on a carriage arranged so that it can be moved along parallel ways by means of a screw. The action of the apparatus is as follows: Light from a source S made slightly convergent by a lens falls upon the plate A, the beam being divided into two portions by the half-silvered surface. If the source has a considerable area, such as a flame, no lens is required. One portion is reflected to the mirror D, the other transmitted through B to the mirror at C, which is fixed in position. The mirror D returns the light to A, a portion of it escaping through the half-silvered film and entering the observer's eye, which is located at O, at which point the convergent beam should come to a focus. The light reflected back

from C is in part reflected from the silver film and enters the eye over the same path. If the path-difference is an odd number of half wave-lengths, these two streams will interfere destructively and we shall have darkness. The path-difference between the two rays can be altered by moving the mirror D by means of the screw, consequently the point in question upon the half-silvered surface will appear alternately bright and dark as the carriage is moved along the ways. The plate B is not essential, and its object will be explained presently. We can get a better idea, perhaps, of the action of the instrument in the following way: The mirror C is seen by reflection in the half-silvered film in coincidence with the mirror D

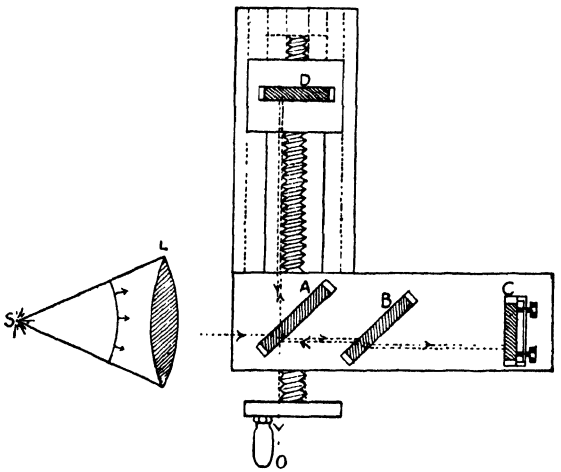


FIG. 186

if the optical paths are the same. The instrument is thus the equivalent of two parallel reflecting surfaces, one real the other virtual, the distance between which can be varied. The phenomena presented by the interferometer are thus similar to those shown by thin films, the difference lying in the fact that in the present case we may make the distance between the reflecting surfaces as great or as small as we please.

The fact that the two beams which interfere are widely separated (at right angles) makes it possible to introduce objects of large area (*e.g.* a tube filled with a gas, the dispersion of which is to be investigated) is of great importance. One of the reflecting planes being a virtual image, results can be accomplished which are impossible with two material reflectors, for one can be brought into exact coincidence with the other, or made to pass through it. If the virtual reflector makes a small angle with the real one, and the

distance between them is small, we observe the fringes shown by a thin wedge-shaped film (such as are seen in the case of Newton's rings). These are located in the plane of the reflector. By moving the mirror forward or back we can cause the reflecting planes to intersect along a line, and along this line we have the black "central fringe" seen with white light which corresponds to the Newton black spot observed on very thin soap-films. The change of phase occurs at the oblique mirror, one reflection being from a glass-silver surface, the other from silver-air. If the two reflecting planes (the real and the virtual one) are exactly parallel, we have circular fringes (located at infinity) of the type first described by Haidinger. These are best seen with a telescope or with the eye focussed for distance. This type of interference we have already studied in the case of thick plates. We can, with this instrument watch the increase in the diameter of the circular Haidinger fringes as the reflecting planes approach each other, until, when they are practically in contact, the whole field becomes uniformly bright or dark, according to the exact setting of the mirror. This means merely that the central circular spot of the fringe system has become so large that it fills the entire field of view.

The plate B is called the compensator, and is introduced to make the two optical paths symmetrical. In its absence it is obvious from the diagram that one of the interfering beams which enters the eye has traversed the plate A three times, while the other has passed through it but once; the double transit of the latter ray through the compensator makes the two paths optically equivalent. The compensator has also another use, for by turning it slightly we can increase or diminish the optical path, thus compensating for and measuring a change produced in the other path, as, for example, by the introduction of a thin film, the refractive index of which we wish to determine.

Adjustment of Interferometer. — The following directions for the use of the instrument are taken from Mann's *Manual of Advanced Optics*. "Measure roughly the distance from the silver half-film upon the rear of the plate A to the front of the mirror C. Set the mirror D, by turning the worm wheel, so that its distance from the rear of A is the same as that of C from A. This need not be done accurately. It is suggested because it is easier to find the fringes when the distance between the mirror D and the virtual image of the mirror C is small. This distance will hereafter be called the distance between the mirrors.

"Now place a sodium burner, or some other source of monochromatic light, at S, in the principal focus of a lens of short focus. It is not necessary that the incident beam be strictly parallel. Hold

some small object, such as a pin or the point of a pencil, between L and A."

A pin-hole in a card is preferred by the author, as the vertical and horizontal adjustments can be made with greater precision. It must be removed when looking for fringes of course.

"On looking into the instrument from O, three images of the small object will be seen. One image is formed by reflection at the front surfaces of A and D; the second is formed by the reflection at the rear surface of A and the front surface of D; the third is formed by reflection from the front surface of C and the rear surface of A. Interference fringes in the monochromatic light are found by bringing this third image into coincidence with either of the other two by means of the adjusting screws upon which the mirror C rests. If, however, it is desired to find the images in white light, the second and third of these images should be brought into coincidence, because then the two paths of the light in the instrument are symmetrical, *i.e.* each is made up of a given distance in air and a given thickness of glass. When the paths are symmetrical, the fringes are always approximately arcs of circles as described above. If, however, the first and third images are made to coincide, then the two optical paths are unsymmetrical, *i.e.* the path from A to C has more glass in it than from A to D, and in this case the fringes may be ellipses or equilateral hyperbolae, because of the astigmatism which is introduced by the two plates A and B. It is quite probable that the fringes will not appear when the two images of the small objects seem to have been brought in to coincidence. This is simply due to the fact that the eye cannot judge with sufficient accuracy for this purpose when the two are really superposed. To find the fringes, then, it is only necessary to move the adjusting screws slightly back and forth. As the instrument has here been described, the second image lies to the right of the first.

"Having found the fringes the student should practise adjustment until he can produce at will the various forms of fringes. Thus the circles appear when the distance between the mirrors is not zero, and when the mirror D is strictly parallel to the virtual image of C. The accuracy of this adjustment may be tested by moving the eye sideways and up and down while looking at the circles. If the adjustment is correct, any given circle will not change its diameter, as the eye is thus moved. To be sure, the circles appear to move across the plates because their centre is at the foot of the perpendicular dropped from the eye to the mirror D, but their apparent diameters are independent of the lateral motion of the eye. For this reason it is advisable to use the circular fringes whenever possible.

"To find the fringes in white light, adjust so that the monochromatic fringes are arcs of circles. Move the carriage rapidly by intervals of a quarter turn or so of the worm wheel. When the region of the white-light fringes has been passed, the curvature of the fringes will have changed sign, *i.e.* if the fringes were convex toward the right, they will now be convex toward the left. Having thus located within rather narrow limits the position of the mirror *D*, which corresponds to zero difference of path, it is only necessary to replace the sodium light by a source of white light, and move the mirror *D* by means of the worm slowly through this region until the fringes appear."

A better control of the motion can be obtained by placing a small white gas flame behind the sodium flame. This gives us a white spot in the centre of the field, on which the colored fringes appear when we reach the centre of the system.

"These white-light fringes are strongly colored with the colors of Newton's rings. The central fringe — the one which indicates exactly the position of zero difference of path — is, as in the case of Newton's rings, black. This black fringe will be entirely free from color, *i.e.* perfectly achromatic, if the plates *A* and *B* are of the same piece of glass, are equally thick, and are strictly parallel. If they are matched plates, *i.e.* if they are made of the same piece of glass and have the same thickness, their parallelism should be adjusted, until the central fringe of the system is perfectly achromatic. When this is correctly done, the colors of the bands on either side of the central one will be symmetrically arranged with respect to the central black fringe."

If the instrument is illuminated with sodium light it will be found that the fringes become invisible periodically as the mirror is moved, for reasons which have been given in the Chapter on Interference. It will be found instructive to illuminate the instrument with a lithium flame containing a little sodium, and note the shortness of the periods of indistinctness. In using the instrument to measure the refractive index or dispersion of a gas, the tube containing the gas can be closed with plates of thin plate glass, which, if of good quality, do not much affect the appearance of the fringes. The tube is highly exhausted and the gas then slowly admitted, the shift in the fringe system being determined by counting the number of bands which cross the hair in the telescope used to view them.

The interesting investigation by Johonnott [1] on the thickness of the "black spot" on soap-films, is an example of the many interesting applications of the interferometer. If we know the thickness of a transparent plate we can measure its refractive index by in-

[1] *Phil. Mag.*, *47*, 501. 1899.

serting it in one of the optical paths of the instrument and measuring the fringe displacement. The white system must be used of course in conjunction with the sodium or other monochromatic system, as the central fringe is the only one that can be identified. The abnormal displacement of the central band referred to in the Chapter on Interference must also be remembered.

It is evident now that if the refractive index of a film is known the thickness can be determined. Johonnott found that, by employing a battery of 54 soap-films mounted on frames, it was possible to get a measurable shift of the fringes even when the films were so thin that they refused to reflect light, *i.e.* showed the Newton black.

The thickness was found to vary between .00006 mm. and .0004 mm.

Effects of Surface Films on the Mirrors. — In general the central fringe, with white light, is black: its locus being the intersection of the real mirror with the virtual image of the other. The path-difference being zero in this case, it is obvious that the destructive interference must result from a phase-change due to reflection of the two rays under different conditions, as is the case with the "black spot" of Newton's rings, as we have seen. With the type of interferometer used by Michelson in repeating Fizeau's experiment on the velocity of light in a moving current of water, the central fringe is white instead of black, as no phase-change occurs in this case. One ray is reflected twice from silver (once from the air side and once from the glass side), while the other ray is twice transmitted through the silver. As Zeeman [1] has shown, however, if the silver is modified by the action of chemical vapors, the central fringe may become black. A copper mirror obtained by cathodic sputtering gave a black fringe unless protected by a thin film of celluloid, applied immediately after the sputtering. Gold and silver mirrors give a white central fringe, but curiously enough an alloy of 10% Au–90% Ag gave a black one. In all cases where a very high degree of accuracy is required, it is well to guard against possible deterioration of the metal film with time. Zeeman found that a freshly sputtered silver film gave a white central fringe, midway between two dark ones, but in half an hour it had moved towards one of the dark fringes by one-sixth of the fringe width.

Twyman and Green's Application of the Interferometer to the Correction of Imperfections in Prisms and Lenses. — A very great advance in the method of testing and correcting imperfections in prisms and lenses was made at the firm of Adam Hilger by Twyman

[1] Zeeman, *Zeit. f. Phys. Chemie*, Cohen-Festband, 1927.

and Green. The interferometer is illuminated by plane-waves furnished by a point source at the focus of an extremely well-corrected lens. A screen perforated with a hole .5 mm. in diameter is placed in front of the source, which may be a mazda lamp with opal globe. The first objective is placed in the position L shown in Fig. 187 and a second objective placed at O which focusses the

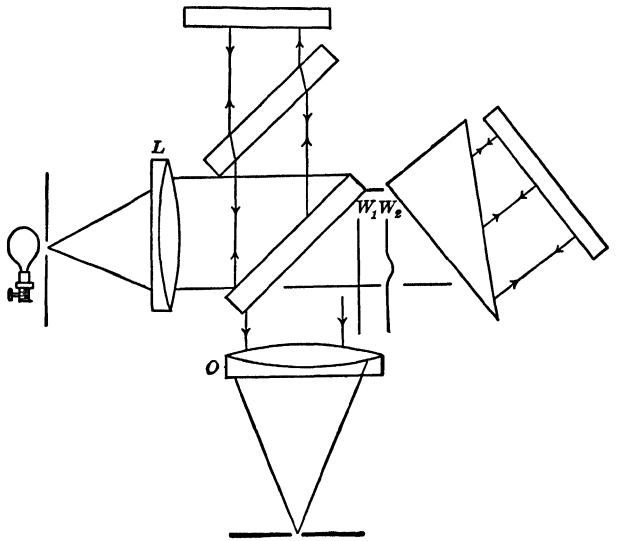


FIG. 187

waves upon a small aperture in a second screen, through which the observations are made. The improvement over the old method of using the interferometer lies in the use of truly plane-waves, and the second objective. If the two uniting wave-fronts are both truly plane, the whole field of view will appear of uniform intensity (or color). If now a prism P is introduced into one path as shown in the figure, any distortions of the plane-wave-front resulting from two transmissions through the prism cause fringes to appear on the uniformly illuminated field. If, for example, there is a globular region in the prism of slightly higher refractive index than that of the rest of the glass, a system of concentric rings will appear, and the area in which they appear is marked on the face of the prism. As Twyman shows [1] this fringe system may be regarded as contour lines of the deformed wave-front, which localize the regions of equal elevation (or depression) like the contour lines on a map.

The fringes would present the same appearance if the wave-front were reversed, the elevations becoming depressions, and we require,

[1] *Phil. Mag.*, *35*, 49, 1918.

therefore, some method of distinguishing between the two as we must know whether the bad spot in the prism has a higher or lower refractive index than that of the rest of the glass. This can be done by pressing against the support of the mirror behind the prism, which retards slightly the reflected wave-front and the fringes expand if the deformation is a "valley" in wave-front W_2 as shown in the figure, while if the deformation is a hill, the rings will close in. In the former case the region causing the deformation has a higher refractive index than that of the surrounding glass, and the correction is made by marking the area on the face of the prism and then rubbing away some of the glass from the surface with a very small polishing tool until the field becomes uniform again. In other words the thickness of the prism has been locally reduced to compensate for the local high index of refraction. The prism obviously has been corrected in this case for two transmissions, and would have to be used with the original plane (or uncorrected surface) silvered, the rays being reflected back to the single lens which serves as collimator and telescope (Littrow spectrograph).

For a prism which is to be used in the usual way with a single transmission, the correction would be one-half of that indicated by the interferometer.

The method has been adapted also to the correction of lenses, and practically perfect lenses, prisms and plates can now be made from glass which is not absolutely homogeneous, — and even the best optical glass is not free from defects.

The Twyman and Green interferometer can be adjusted by the following method, which can be used as an alternative to the one given for the Michelson instrument (which is merely the T. and G. one without the lenses).

Two spots of light are seen on the back of the perforated screen in front of the lamp, caused by the rays which pass through the half-silvered plate after reflection from the two back mirrors, and two similar spots on the screen at the focus of the second objective. The two pairs of spots should be brought simultaneously into coincidence with the two apertures, by tilting the half-silvered plate and one of the back mirrors. The opal lamp is then replaced by a sodium flame or other source of monochromatic radiations and on looking through the eye aperture, the fringes are seen, which can be widened by tilting one of the back mirrors until one of them fills the entire field. If the screens are painted white, it is easier to see the small spots.

There is an essential difference between this method of testing plates and that employed by Fizeau, which was described in the

Chapter on Interference. We are dealing here with two transmissions through, say, a plate of thickness t. The retardation on the other ray is therefore $2\mu t - 2t$ or $2(\mu-1)t$, which is the quantity tested. In the Fizeau instrument, in which the light is reflected from the upper and lower surface of a plate, the retardation is simply that of the plate, or $2\mu t$. Plates for echelon gratings, which will be described presently, require the constancy of $(\mu-1)t$, as they are used with a single transmission.

To correct a large plate which is to be cut up into echelon plates, the Fizeau fringes and the fringes obtained with the Twyman and Green interferometer are plotted, by which variations in the refractive index can be found to the sixth decimal place.

Determination of Refractive Index and Dispersion with the Interferometer. — The refractive index of a transparent plate and its dispersion can be obtained by means of white light in combination with monochromatic light of a single wave-length. The determination of the dispersion is based upon the shift between the *true* and the *apparent* position of the centre of the system of fringes formed by white light, which we have just studied. The plate should be sensibly plane-parallel, and should be cut in two, the two portions placed in the paths of the interfering beams in such positions that they cover the same portion of the field. We may illuminate the instrument with a sodium flame backed by a candle flame. The two pieces of the plate should be so arranged that they can be rotated about vertical axes, one of them very slowly and uniformly, the angles of rotation being measured with a mirror and scale. We can set them normal to the rays, by turning them to the point where the direction of motion of the fringes resulting from the increase of path with increasing incidence angle reverses.

Adjust the instrument so that both the white-light fringes and the sodium fringes appear in the field. Then turn one plate through a convenient angle, which is read from the scale. Turn the other plate very slowly, counting the sodium fringes as they pass over the cross hair of the observing telescope until the white fringes again appear and occupy their former position. Let the angle through which the plates have been turned be i, the fringe count $2N$, the thickness of the glass t, its refractive index μ, and the wave-length of the sodium light λ, it can be shown that

$$\mu = \frac{(t - N\lambda)(1-\cos i) + \dfrac{N^2\lambda^2}{2t}}{t(1-\cos i) - N\lambda},$$

in which the term $N^2\lambda^2/2t$ is negligible. We now restore one plate to its original position, and move the interferometer mirror until

the white-light fringes appear in their former position, counting the sodium fringes as they cross the hair. The number will be greater than $2N$, the difference, which we will call $2N'$, being due to the dispersion.

The Cauchy dispersion formula can be assumed, $\mu = A + B/\lambda^2$, and we have $N' = 2Bt'/\lambda^3$, in which t' is the thickness of the glass introduced by the rotation,[1] as was shown in the Chapter on Interference.

Light-Waves as Standards of Length. — Probably the most important use to which the interferometer has been put was the determination of the length of the standard metre in wave-lengths of the monochromatic radiations from cadmium. The invariableness of the wave-length of the radiation sent out from the atoms of a metal, brought to a state of luminescence by electrical discharges in a high vacuum, suggests their adoption as a standard of length. This proposition was first made by Lamont in 1823, and subsequently by Dr. Gould about fifty years ago. At that time the interferometer in its present form was unknown, and the method proposed involved the use of the diffraction grating, the measurement of its width, and the determination of angles, all of which measurements would have entailed no very inconsiderable errors. Michelson suggested the use of his interferometer, and through the efforts of Dr. Gould, who represented the United States in the International Committee of Weights and Measures, was asked to carry out the experiments at the International Bureau at Sèvres in collaboration with Benoit. A very complete description of the method will be found in Professor Michelson's book, *Light-Waves and Their Uses* (Chicago University Press, 1903).

The general principle of the method can be briefly outlined as follows: —

The problem is to measure the distance between the two marks on the standard metre bar in terms of the wave-length of light, or, in other words, find out how many light-waves there are in a beam a metre long.

A bronze bar 10 cms. in length, of the form shown in Fig. 188, was prepared, on the ends of which two silvered-glass mirrors were mounted which could be made accurately parallel by observing the interference fringes, formed in the manner to be described presently. The principle consisted in finding the number of light-waves in a beam whose length was equal to the distance between the planes of the two mirrors, and then to find how many times this distance was contained in the metre. In a length of 10 cms., there are, however, roughly 300,000 light-waves, and the direct

[1] Mann's *Manual of Optics*.

determination of this number by actual count would have involved too much labor and too great a risk of accidental mistakes. Nine other standards similar to the above were therefore prepared, each half as long as its predecessor, *i.e.* of lengths 10, 5, 2.5, 1.25, etc., cms.; the smallest unit had mirrors with reflecting planes only .39 mm. apart. The number of light-waves in this distance was first determined for the red, green and blue radiations from a vacuum tube containing cadmium vapor. This was accomplished by putting the bar with its two mirrors in the place of one of the mirrors of the interferometer; the other mirror was then brought into such a position that the central fringe (white light) appeared in the field of, we will say, the lower mirror. By moving the mirror back the centre of the system could be made to appear in the upper mirror, and by counting the number of fringes which passed during this operation the number of wave-lengths in the distance through which the mirror moved could be determined.

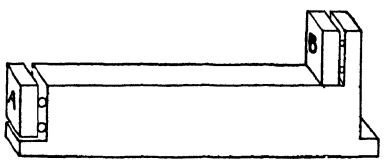


FIG. 188

This first "*etalon*," as it was called, was next compared with the second by mounting the two side by side, in place of the movable mirror of the interferometer. The field of view now consisted of four square areas corresponding to the four mirrors of the *etalons*. The longer of the two (No. II) was fixed in position, while the shorter (No. I) could be moved by turning the screw of the instrument. The reference plane (image of the interferometer mirror seen in the plate) was then brought into coincidence with the front surface of the lower mirrors of the two etalons (the plane of the lower dotted line in Fig. 189), by moving the interferometer mirror until the colored fringes appeared. This mirror, which is usually fixed, in the present type of instrument could be moved along parallel ways.

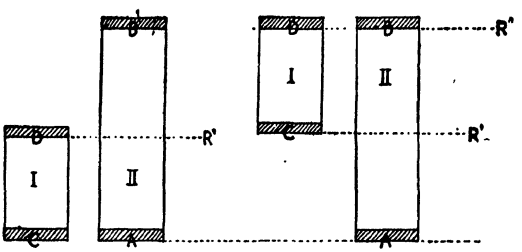


FIG. 189

It was then moved back until the reference plane coincided with the upper mirror *D* of etalon *I*, the plane *R'*. The fringes passing during this motion of the mirror were counted, the number of course corresponding with the number previously determined.

Etalon I was now moved back until C came into coincidence with the reference plane R' (Fig. 189). The reference plane was now moved to R'', until it coincided with D' in its new position, and was within a few wave-lengths of the plane of B, the number being found by turning the compensating plate. The second etalon was then compared with the third, and so on, until finally the number of wave-lengths in the 10 cm. etalon had been determined. A mark on this etalon was then brought into coincidence with one of the end marks on the metre bar under the microscope, and the etalon was then progressively advanced, its front mirror being brought into coincidence with the plane previously occupied by the rear mirror, the reference plane then moved back and the process repeated. In this way the total number of waves in a length equal to the standard metre was determined. The final results were as follows, for 15° C. and 760 mms. pressure.

Red line 1 m. = 1553163.5λ, *i.e.* $\bar{\lambda} = 6438.4722AE$,
Green line 1 m. = 1900249.7λ, *i.e.* $\lambda = 5085.8240AE$,
Blue line 1 m. = 2083372.1λ, *i.e.* $\lambda = 4799.9107AE$.

The values given by Rowland for these same lines are

6438.680, 5086.001, and 4800.097.

An idea of the accuracy of the work can be obtained by comparing three independent observations, the first two by Michelson, the third by Benoit:

1553162.7, 1553164.3, 1553163.6.

In addition to recording the length of the standard metre in terms of an invariable unit, this remarkable piece of work has given absolute determinations of three standard lines, which will doubtless stand for a long time, if not forever, as the standards from which all other lines will be measured.

It may be well to point out here that it has been recently shown by Michelson, and proven experimentally by Kayser, that Rowland's coincidence method is not accurate. As a result of small errors of ruling, the second order ultra-violet line of wave-length 2 may not fall exactly upon a first order line of wave-length 4. The use of the grating is thus restricted to obtaining the wave-lengths of lines between fixed standard lines, by interpolation, at least if the greatest accuracy is required. The standard wave-lengths in use at the present time have all been measured by interferometer methods.

The Visibility Curves. — As we saw in the Chapter on Interference, the fringe system formed with Newton's combination of a lens and flat plate, illuminated with sodium light, is not continuous. There are periodic regions of invisibility as we proceed outward

from the centre, due to the fact that when the maxima of D_2 coincide with the minima of D_1, uniform illumination results. If now D_1 and D_2 were infinitely narrow lines and single, the fringes would be equally distinct when "in-step," regardless of the path-difference. If, however, this is not the case, the visibility will vary at the different points of maximum distinctness. Suppose, for example, that each line is a close double; with a sufficiently large path-difference, the two components of D_1 will get out-of-step, and we shall have uniform illumination and invisibility entirely independent of the light from D_2. Fizeau and Foucault, who may be regarded as the founders of interference spectroscopy, only recorded the successive recurrences of the fringes as the path-difference increased. Michelson went a step further, and measured the distinctness of the fringes at each reappearance. From these observations he was able to compute the nature of the lines, *i.e.* whether they were single or double, broad or narrow, etc. If J_1 denotes the maximum brightness of a fringe, and J_2 the intensity of the dark region between, Michelson calls

$$\frac{J_1-J_2}{J_1+J_2}=V,$$

the "Visibility" a quantity which represents the distinctness with which the fringes appear to the eye.

If we know the nature of the distribution of the light in the source, *i.e.* whether the lines are single or double, accompanied or not by fainter companions, etc., it is possible to construct a visibility curve in which the values of V are plotted as ordinates and the path-differences as abscissae.

Michelson commenced by calculating the visibility curves which would result from various types of single, double and multiple lines. Examples of such curves are shown in Fig. 190, the intensity curves of the spectrum lines being shown to the left of each. The curves shown are resultant curves formed by the superposition of wave-trains such as would emanate from sources having a distribution of intensity as figured. The visibility curves are obviously the envelopes of the above curves. Michelson next took up the subject of the construction of an intensity curve from a visibility curve, a much more difficult problem. His work along this line was much aided by the invention of his harmonic analyzer, a machine which separates out of a complex curve the simple harmonic curves of which it is formed; in other words, makes a Fourier analysis of it.

As Lord Rayleigh [1] has shown, the rigorous solution of the prob-

[1] *Phil. Mag.*, *34*, 407, 1892.

lem is not possible, for, except in cases where there is symmetry in the group of lines, we may have a large number of different distributions of intensity, all of which give the same visibility curve. It is impossible, moreover, to decide from the visibility curve on which side of the principal line a fainter component lies. Michelson's predictions regarding the structure of many lines have been subsequently verified, however, and he is to be regarded as the pioneer in the field of investigations devoted to the minute study of spectrum lines.

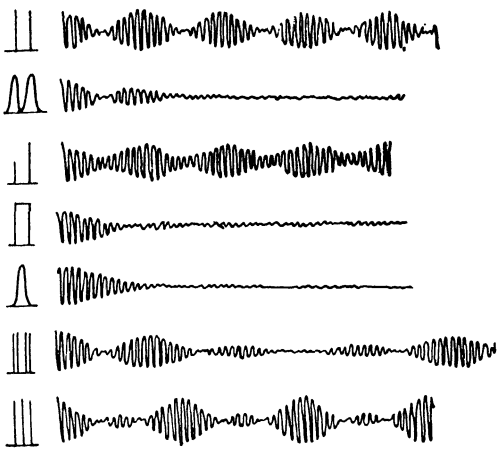


FIG. 190

The method has not been used to any great extent by other observers, partly from the great difficulty of estimating "visibilities" of the fringes, and partly from the difficulty in interpreting the results. Michelson's results were due to his great skill in this respect, which resulted from long experience and familiarity with his instrument. The more modern interferometers show objectively what before could only be guessed at, that is, they actually separate the line into its components just as the prism and grating separate the originally composite light into a spectrum of lines.

It should be noted, however, that the study of line structure by this method gives us practically unlimited resolving-power, and, as pointed out by W. E. Williams in his recent book on *Applications of Interferometry* might still be employed to advantage for examining the hyper-fine structure of a satellite line that could be isolated by a powerful auxiliary instrument. Very large path-differences would be required, and the difficulty of the small size of the rings could be avoided by employing the method of Twyman and Green (small point source at focus of a lens, the light eventually focussed on the pupil of the eye) which gives localized fringes of any size desired, no change resulting as the path-difference is increased.

Michelson's genius gave us the next instrument in the series which we are considering, and we will now take up the subject of

one of the most original and interesting optical instruments ever devised, the echelon grating.

The Echelon Grating. — A remarkable type of grating was constructed by Michelson.[1]

As we have seen in the Chapter on Diffraction the resolving-power of a grating is represented by mn, the product of the order of the spectrum and the number of lines. High resolving-power had been secured previously by ruling a very large number of lines; Michelson attacked the problem in a new direction and constructed a grating for which m instead of n had a large value. Michelson accomplished this by building up a flight of steps of glass plates all of exactly the same thickness and plane-parallel to within $\frac{1}{20}$ of a wave-length of sodium light. The plates were cut from a single disk 2 cms. in thickness which was figured with the greatest care by Mr. Petitdidier. The manner in which the echelon operates has been described fully in the Chapter on Diffraction. The mathematical treatment, given in earlier editions, has been omitted. The retardation by a plate 2 cms. in thickness is considerably over 20,000 wave-lengths; consequently we are dealing with a spectrum of the 20,000th order, if the plates have this thickness. The number of the plates cannot be increased above 30 to advantage, owing to the loss of light by absorption and reflection from the surfaces. Our resolving-power is thus about $30 \times 20{,}000$ or roughly 600,000, or the grating should separate lines only $\frac{1}{600}$ of the distance between the D lines apart. The Adam Hilger Co. made a great improvement in the construction of echelons by making the plates so flat that they could be brought into optical contact, thus doing away with the loss by reflection.

The echelon throws all its light into one, or at most two, spectra; consequently it is well adapted for the minute structure of faint spectrum lines. Its great disadvantage is the difficulty of interpreting the results obtained with it, and the impossibility of seeing more than a single line at a time unless an auxiliary spectroscope is employed. Even if sodium light is used nothing can be seen which can be interpreted. With certain thickness of plates the D_1 and D_2 spectrum lines may coincide, one being seen in, say, the 2000th order, and the other in the 2030th, owing to the difference of retardation. With plates of a different thickness the D_1 spectra may fall midway between those due to D_2. As the spectra of succeeding orders are very close together, it is obvious that, except when employing extremely homogeneous radiation, we shall have a confused jumble of lines.

Only three different orders can be seen at one time, but by turn-

[1] *Astrophysical Journal*, *8*, 36, 1898.

ing the echelon slightly others may be brought into view. We can set the echelon so as to have two adjacent orders of equal intensity, as in the first diagram of Fig. 191, or so as to have one bright line bordered by two faint ones. The latter condition is usually preferable.

The light must undergo previous prismatic analysis before it enters the collimator slit of the echelon spectroscope, or we may illuminate the slit with the heterogeneous light and place a prism between the echelon and the telescope. The instrument is especially adapted for the exhibition of the Zeeman effect, as it is compact, and extremely saving of light, and requires practically no adjustment if the plates are properly mounted in a metal case. The writer has had no difficulty in showing the Zeeman effect with an improvised echelon made by standing four interferometer plates on the table of a small spectroscope. A five-element grating is secured in this way, since a stream of unretarded light can be passed by the edge of the first plate. The width of the steps should not exceed one or two mms. and a cardboard screen should be so arranged as to cut off all the light except that which comes through the steps, a clear space of equal width to one side of the first plate, and a strip of the same width at the edge of the last plate, *i.e.* the top step. In other words, when looking at the echelon from the direction of the telescope the screen should hide everything except five vertical elements of equal width, four of them glass and one air. A direct vision prism can be put between the plates and the telescope to separate the echelon spectra of the different lines in the spectrum under investigation. A mercury vacuum tube between the conical poles of a powerful electromagnet is a suitable source of light to work with, the green line splitting up as soon as the current is turned on.

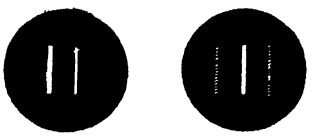


FIG. 191

The Reflecting Echelon. — Michelson's original suggestion that the echelon might also be used as a reflection instrument has only very recently been carried out. With the older method of construction it would have been impossible to make an efficient instrument of this type owing to the variation in the thickness of the air film between the plates, but with the Hilger Company's method of sliding and pressing the plates into optical contact, the height of the reflecting steps becomes constant. W. E. Williams has perfected the technique of employing such an instrument, and to his energy is due the construction of remarkably perfect reflection echelons of fused quartz by the Hilger Company. Quoting from his recent book on *Interferometry* we have the following method

of employing the instrument, which now, for the first time, can be used for the short ultra-violet waves.

"In a transmission instrument the variation of path difference needed to change the position from single to double order can be obtained by slightly tilting the Echelon, and this is done without any material displacement or deviation of the spectra.

"The same method cannot be used with the Reflection type as the necessary tilting would cause the reflected beam to go outside the field of view.

"Fig. 192 shows the method finally adopted for using a Reflection Echelon in conjunction with a spectrograph. The light to be analysed is focussed on a primary slit S_1 which may either be vertical or horizontal. It is reflected by a small quartz prism to the objective O and through the quartz or fluorite window W on the Echelon. The reflected beam is focussed by O on to the slit S_2 of the spectrograph. The latter makes a coarse analysis of the spectrum and separates the lines of widely different wavelengths so that these do not overlap as they do at the slit S_2. When S_1 is vertical the edges of the Echelon plates must also be vertical as in the figure, while the slit S_2 must be sufficiently wide to accommodate at least two orders of any line. By also taking a plate with the Echelon turned through two right angles, it is possible to determine whether a satellite lies on the long or short wavelength side of a main line. In the one position of the Echelon, its dispersion assists that of the spectrograph, while in the other, it opposes it. With a very rich spectrum such as that of the iron arc, the slit S_1 and the Echelon plate edges are arranged horizontally so that its dispersion is along the slit S_2 which can now be narrowed down as desired. In order to change a line from single to double order position as required, the air pressure in the chamber C is varied by connecting it to a suitable pump, and provided there is no leakage of air to or from the chamber during an exposure, the pattern remains steady in spite of variations of temperature. This is because the density and consequently the refractive index of the gas in the chamber remains constant, while the expansion of the silica Echelon plates with ordinary temperature changes is negligibly small."

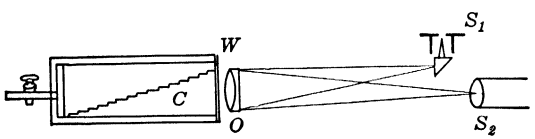


FIG. 192

The Interferometer of Fabry and Perot. — This instrument enables us to observe the Haidinger fringes formed by two parallel

reflecting films of silver, the distance between which can be varied continuously. In the Michelson instrument the fringes, due to a double line, disappear when the distance between the mirrors is such as to bring the maxima due to one line over the minima due to the other. As was shown in the Chapter on Interference the multiple reflections, which occur with two opposed thin films of silver, render the bright rings seen by transmitted light very narrow in comparison to the dark regions between them, so that, with increasing distance between the plates, the ring systems due to two or more different wave-lengths can be seen simultaneously.

The theory of the Fabry and Perot instrument was given in the treatment of the Haidinger fringes and the influence of multiple reflections in the Chapter on Interference. Multiple reflections result in a narrowing of the interference maxima, in much the same way as in the case of the diffraction grating, but it is to be especially noted that in this case we obtain an exceedingly narrow spectrum line (ring) without employing a slit as the source.

One of the plates is moved on a carriage by the rotation of a screw, as in the Michelson apparatus. With this instrument we can gradually increase the distance between the silver reflecting films and follow the division of a ring into its components in the case of light in which two or more different wave-lengths are present. In their later work a simpler form of mounting was used, the instrument being called an étalon. In this case the plates were separated by a metal ring with three studs (preferably of invar) as shown in Fig. 193. The plates are held against these studs by springs, accurate parallelism of the silvered surfaces being secured by varying the tension on the springs. The preliminary adjustment is made by lightly filing the studs.



FIG. 193

The glass plates should be optically flat and it is better to have the two surfaces of each plate not quite parallel, otherwise a second set of rings may be formed. The test for parallelism is the same as for the two surfaces of a single plate. If only arcs of circles are seen, the centre being "off stage" we can tell by squeezing one side or the other which stud requires filing to improve matters. When we have the centre of the system in the field of view, we have only to move the eye up and down and to the right and left to make the final adjustment by pressure of the springs. If the rings expand, new ones coming out from the centre, as we move the eye to the right, it means that we have passed to a region of greater thickness, and the plates must be pressed on the right-hand side into closer contact with the distance piece or the pressure released

slightly on the left. When the reflecting films are exactly parallel, the diameter of the central ring remains unaltered as the eye is moved about.

The paths of the rays through the half-silvered films are shown in Fig. 194, in which the source of the light is located at S and the eye at O. The opposed surfaces of the glass plates are half-silvered, and the distance between them can be varied from zero to 30 or 40 cms. Starting with the plates close together and accurately parallel, we shall observe in the case of sodium light a system of circular fringes similar in appearance to those seen with Newton's glasses, except that the maxima are extremely narrow circles of light with broad dark regions between them. On increasing the distance between the mirrors the rings due to the two components of the sodium light will gradually get out-of-step; but instead of disappearing, as they do in the Michelson instrument, they merely become double in the present case. In other words, we can observe the separation of the D_1 and D_2 fringes. On further increasing the path the fringes will again coincide.

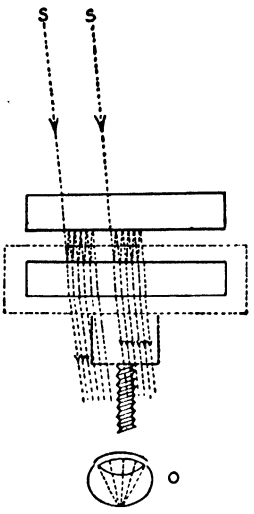


FIG. 194

It is clear that in the present case we can observe *directly* that which we were obliged to infer from the visibility curves in the case of the Michelson interferometer. Faint components lying close to a bright line can be observed directly with this instrument by making the distance between the plates sufficiently great. Fabry and Perot studied the minute structure of a large number of spectrum lines, and have obtained results which the Michelson instrument is incapable of yielding. Photographs of the rings given by the instrument illuminated by the green light of the mercury arc are shown in Fig. 197, page 317.

As the resolving-power of the Fabry and Perot instrument depends upon the number of reflections it is important to have the thin silver films highly polished, a condition which occurs automatically in the case of films put down by cathodic sputtering or evaporation in vacuo from an electrically heated tungsten wire on which a deposit of silver has been made electrolytically. If facilities for this technique are not available the silvering can be done by the formaldehyde process (given in earlier editions). With films of the correct thickness, reflecting about 80% as many as forty

images can be seen, when a tungsten lamp is viewed through the films, superposed at a small angle. They are white at first, but speedily become reddish owing to the lower reflecting power of silver for the blue. If the row of images is viewed through a prism, the falling off of the blue end of the spectrum by repeated reflection is easily seen. A photograph of the phenomenon is reproduced in Fig. 195. The blue has practically disappeared at the 8th image and the green at the 10th. On this account the resolving-power is greatest in the red and falls off rapidly with decreasing wave-length.

Someone has recently found that exposure of a silver film to the vapor of hydrogen peroxide increases the reflecting power, a very remarkable action which is worthy of further study.

If the formaldehyde process is used the films are best polished

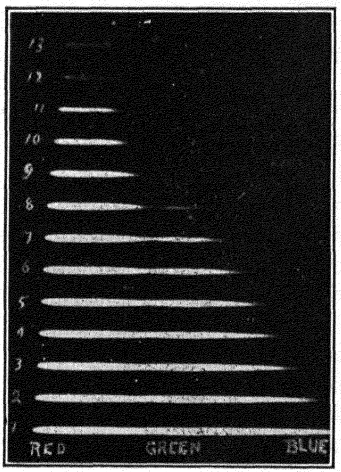


FIG. 195

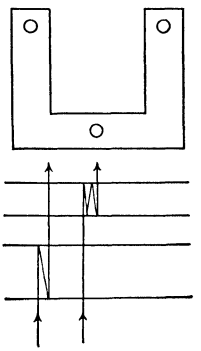


FIG. 196

by the method recommended by Pfund, a light brushing with an eiderdown powder puff, charged with rouge.

Determination of Wave-Lengths with the Fabry and Perot Interferometer. — A comparison of the standard metre with the wave-length of the red cadmium line was made by Benoit, Fabry and Perot by means of etalons of thicknesses 100, 50, 25, 12.5 and 6.25 cms. These etalons were made of invar of the form shown in section by Fig. 196. Three polished hemispherical studs of the same metal served as supports against which the half-silvered plates were pressed by springs, the tension of which could be adjusted. The method depends upon the principle that if white light is passed through two etalons, one of which has twice the thickness of the other, interference takes place between a ray which has passed through the thicker and suffered *two* to-and-fro

reflections in the thinner, and one which has been once reflected to-and-fro in the thicker and passed through the thinner, as shown in Fig. 196, since the total distance traversed is the same in each case.

It is, of course, impossible to secure the 2 : 1 ratio exactly, and the correction to be applied in each case is determined by the introduction into the light path of two lightly silvered plates which make an acute angle with each other.

The process consists in comparing the metre with the 100-cm. etalon, which is then compared with the one of 50 cms., and so on down to the one of 6.25 cms., the thickness of which in red cadmium wave-lengths is then determined by a method which will be described presently. The compensating wedge shows in each case the number of wave-lengths to be added or subtracted to make the 2 : 1 ratio exact. The central white fringe is observed with a lens focussed on the wedge at a point where the thickness is sufficient to effect the compensation by a to-and-fro reflection between its silver films. A very complete description of all of the details of the determinations will be found in Schuster's *Optics*.

Standard wave-lengths, which were first made with the diffraction grating are now made exclusively with the Fabry and Perot interferometer, which yields much more accurate results. The late Lord Rayleigh improved the method by employing fixed etalons, dispensing entirely with the sliding plate instrument which had been used by Fabry and Perot.

His method may be described briefly as follows:

Determination of Wave-Lengths. — If we know the absolute wave-length of the red cadmium line, the wave-lengths of any other lines which are known approximately, can be determined absolutely by comparing the ring systems seen in an etalon.

The cadmium red light, of wave-length λ, gives a system of rings. Call P the ordinal number of the first ring, counting from the centre. By this we mean the number of wave-lengths of path-difference involved in the formation of this ring. This integer may be roughly determined by mechanical measure of the thickness, and we shall see presently how it may be exactly determined. For the present we will assume this integer P known.

The order of interference at the centre of the system is slightly greater, since, as we have seen, the path-difference between the rays which interfere is greatest for the rays traversing the plate in the normal direction. Call the order of interference at the centre

$$p = P + \epsilon.$$

We have to determine ϵ, a fraction lying between 0 and 1. It is

obvious that ϵ increases with the diameter of the ring, so that by measuring the latter we can determine the former.

If the thickness of the etalon is e, the order of interference at the centre is $p=2e/\lambda$, or the number of waves path-difference. Suppose the ring to be formed by rays having an obliquity i, the order of interference is $p \cos i$, in which $\cos i$ is a fraction very nearly equal to 1, making the order of interference at the ring slightly less than at the centre. The cos term was explained in the section on Interference by thick plates, for the case of reflection. If x is the angular diameter of the ring P we have

$$P=p \cos \frac{x}{2}, \text{ or, since } x \text{ is small}$$

$$p=P\left(1+\frac{x^2}{8}\right) \text{ by McLaurin's Theorem.}$$

If, at the same time, we have another ring system formed by another wave-length λ' (known approximately), we write

$$p'=P'\left(1+\frac{x'^2}{8}\right)$$

in which x' is the diameter of the first ring, of wave-length λ'.

Substituting for p and p' their equivalents $2e/\lambda$ and $2e/\lambda'$ gives

$$\frac{2e}{\lambda}=P\left(1+\frac{x^2}{8}\right)$$

$$\frac{2e}{\lambda'}=P'\left(1+\frac{x'^2}{8}\right)$$

$$2e=P\lambda\left(1+\frac{x^2}{8}\right)=P'\lambda'\left(1+\frac{x'^2}{8}\right)$$

or

$$\frac{\lambda'}{\lambda}=\frac{P}{P'}\left(1+\frac{x^2}{8}-\frac{x'^2}{8}\right).$$

The required ratio of wave-lengths λ'/λ is thus given as a function of the angular diameters x and x' and the integers P and P'.

Fabry and Perot employed the sliding plate interferometer for determining P, the order of interference for the standard cadmium line.

Lord Rayleigh dispensed with this by the following method. Writing the above equation in the form

$$\frac{P'}{P}=\frac{\lambda}{\lambda'}\left(1+\frac{x^2}{8}-\frac{x'^2}{8}\right),$$

expressing P'/P as a function of λ/λ', regarded as known, and of the diameters of the rings. Quoting from his paper:

"To test a proposed integral value of P, we calculate P' from (5). If the result deviates from an integer by more than a small amount (depending upon the accuracy of the observations), the proposed value of P is to be rejected. In this way, by a process of exclusion the true value is ultimately arrived at.

"The details of the best course will depend somewhat upon circumstances. It will usually be convenient to take first a ratio of wave-lengths not differing much from unity. Thus, in my actual operations the mechanical measure of the distance between the plates was 4.766 mm., and the first optical observations calculated related to the two yellow lines of mercury. The ratio of wave-lengths, according to the measurements of Fabry and Perot, is 1.003650; giving after correction for the measured diameters 1.003641 as the ratio P'/P. From the mechanical measure we find as a rough value of P, $P=16460$. Calculating from this, we get $P'=16519.92$, not sufficiently close to an integer. Adding 22 to P we find as corresponding values

$$P=16482 \qquad P=16542.00,$$

giving P as closely as it can be found from these observations. This makes the value of P for the cadmium-red ring observed at the same time about 14824, and this should not be in error by more than ± 30.

"Having obtained an approximate value of P for the cadmium red, we may now conveniently form a table, of which the first column contains all the so far admissible (say 60) integral values of P. The other columns contain the results by calculation from (5) of comparisons between other radiations and the cadmium red. The second and third columns, for example, may relate to cadmium green and cadmium blue. These also suffice to fix the value of P, but any lingering doubt will be removed by additional columns relating to mercury green and mercury yellow (more refrangible)."

Lord Rayleigh's method of measuring the ring diameters was an improvement on the method used by Fabry and Perot, who employed a fixed interference gauge and telescope provided with an eye-piece micrometer.

He mounted the etalon on a horizontal turn-table so that by rotating it on a vertical axis (midway between the silver films), the right and left edges of a ring could be made tangential to a fixed vertical wire in the eye-piece of the telescope. The turn-table carried a thick strip of plate glass upon which was scratched a

radial line, a point on which was observed with a travelling microscope which read to .001 inch. If the reading of the ring diameter with the microscope was .18 inch, the angular diameter (x) was .018, since the point observed by the microscope was 10 inches from the axis of rotation. The separation of the colored ring systems was effected by means of a direct-vision prism held between the eye-piece and the eye.

In an apparatus modified and enlarged by the author for the determination of the thickness of steel gauges in wave-lengths of light,[1] the etalon is rotated on a horizontal axis by means of a long lever operated by the observer at the eye-piece of the telescope. The ring system is projected upon the rather wide slit of a large long focus-prism spectroscope, and the central vertical sections of the rings are seen in the spectrum lines, which appear as narrow vertical rectangles outlined by the rings which cut across them.

Absent Orders: the Compound Interferometer. — The device of two etalons in tandem employed by Fabry and Perot in the determination of the metre in terms of red cadmium waves, has been adapted by Houston [2] to the study of the fine structure of lines. In the paper referred to below he shows that if the light is passed in succession through two Fabry and Perot interferometers, the first will transmit light to the second only in certain directions, and that this light will be transmitted by the second only if these directions coincide with the inclination directions of the rings of the second instrument.

Let the thickness of the first be d, and of the second D, and suppose these values to be such that $2d \cos \theta = n\lambda$, and $2D \cos \theta = m\lambda$, in which m and n are integers. Both instruments will then transmit for rays inclined at an angle θ, and give a bright ring. The next maxima in general do not coincide and hence will be greatly weakened. (They will coincide, of course, if $d = D$.)

The coinciding maxima will be determined by the ratio of D to d. If $d = D$ the pattern transmitted by the two will be the same as for one, but as the author shows the resolving-power will be slightly increased, as the fringes become narrower.

If $D = 2d$, the separation of orders is that of the first interferometer (thickness d, and large rings) while the resolving-power is that of the separation D, *i.e.* only every other order of the second instrument appears. If $D = 3d$, only every third order is transmitted. It is important to note that a trace of the absent rings appears, and these must not be mistaken for companion lines.

[1] United States Patent 1, 455, 825.
[2] Wm. V. Houston, *Phys. Rev.*, *29*, 478, 1927.

The intensity of these will correspond to the intensity of the dark minima of a single interferometer. Photographs should be made with two different values for the ratio of thickness, the true companion lines or "satellites" can then be distinguished by the fact that their positions will be the same on both plates. Houston gives the following description of the instrument and directions for its adjustment:

"An attachment was made for the interferometer previously built, so that the preliminary interferometer can be clamped to the frame of the other. This clamp carries a steel plate which can be rotated about two axes to set the optical axes of the two instruments together. This plate then carries one mirror which slides back and forth and can be fastened by set screws, and another which can be made parallel to the first. All the adjustments are made by screws working at the ends of levers which are held tightly against the screws by springs.

"It has been found possible to make the necessary adjustments as follows. With the preliminary interferometer removed and the other set at about the separation to be used, the latter is adjusted until the plates are parallel. The other interferometer is then attached and its fringes are viewed from the side by means of a totally reflecting prism. In this way its plates can be made parallel. When the prism is removed the transmitted light shows the ring system of each interferometer as well as regions of brightness where the two systems coincide. If the axis of the preliminary interferometer is then adjusted until these regions of brightness are circles concentric with the other ring systems, the instrument is in adjustment. To make one separation an integral multiple of the other the movable interferometer is opened or closed to make the circles of bright rings move toward the center. As the desired separation is approached these regions become wider until they cover the whole field. A white light source is then put in and the adjustment continued until the colored fringes appear. When the white central fringe appears the desired point has been reached."

A more recent development of the compound interferometer is described in a paper by Lau and Ritter.[1] They employed two thick plane-parallel glass plates, each silvered on both sides, thus having only two instead of four plates to adjust. Moreover a glass etalon has a dispersion n times as great as one of air, or if compared for equal angles, n^2 times as great. The ring system for a quartz plate 3 mms. thick and for a 5-mm. air plate are shown in Fig. 197*a*.

[1] Lau and Ritter, *Zeit. für Phys.*, *76*, 190, 1932.

They employed a thickness ratio of 1 : 1.7. The light passed first through plate *a* of 2.94-mm. thickness and then through *b* of 5-mm. thickness. The combination gave a range of dispersion five times as great as that of the thicker plate, *i.e.* four intermediate orders were suppressed. The ring systems of the two plates acting in tandem are shown in Fig. 197*b* (center) both figures taken from their paper. They give a table showing the maximum number of lines that can be resolved between two adjacent orders for

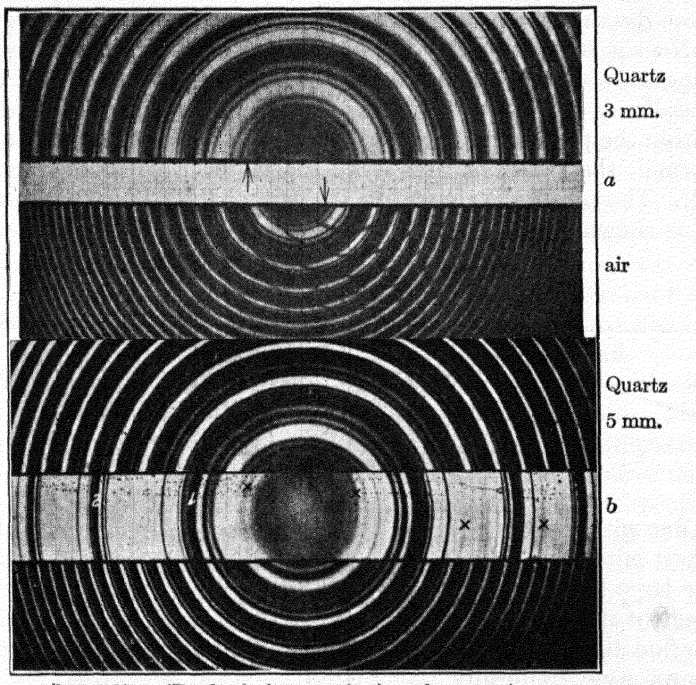


FIG. 197. (Dark circles are the interference rings)

the various types of interference spectroscopes, 30 for the echelon, 17 for the Lummer-Gehrcke plate, 40 for the Fabry and Perot and 200 for the tandem or multiplex as they term it.

The Interferometer of Lummer and Gehrcke. — As we have seen, one of the chief disadvantages of the Fabry and Perot interferometer, is the great loss of light due to the absorption by the two silver films. If we could find some substance to replace the silver which had a very high reflecting power and no absorption, the efficiency of the instrument would be enormously increased. Lummer and Gehrcke, making use of the high internal reflecting power of transparent substances in the neighborhood of the crit-

ical angle, perfected an interferometer superior in many respects to any other instrument. It is merely a long slab of very perfect optical glass or quartz, accurately plane-parallel, with a small prism cemented to one of the surfaces at the end of the plate. In the recent instruments of crystalline quartz made by the Adam Hilger Company of London, the prism is brought into optical contact with the surface, to which it adheres very tenaciously without the use of any cement. This avoids absorption of the ultra-violet, which would occur with any cement, and makes a more durable union than the older device of employing glycerine.

No slit is necessary with the instrument, for the interference fringes are of the same nature as the Haidinger bands of the Fabry and Perot, the only difference being that the angle of incidence within the plate is just short of the critical angle, instead of nearly normal: the emergent rays leave the plate at an angle of nearly 90°. The prism attachment in the case of the Hilger instruments is as shown in Fig. 198. This enables one to work with a source

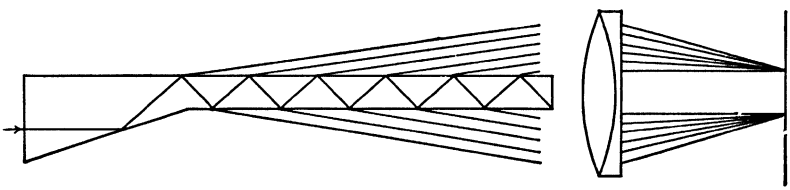


FIG. 198

in line with the plate, which is a convenience. The function of the prism is to pass light into the plate at the proper angle, without the large loss that would occur by reflection, if we employed rays incident at nearly 90° on the upper surface. In this latter case we find that the reflected light, when examined with a telescope, shows wide uniformly illuminated maxima, separated by very narrow dark minima, as with the Fabry and Perot instrument.

The transmitted light exhibits narrow sharp maxima with broad minima between. With the light entering through the prism we have narrow interference maxima formed by transmitted rays both above and below the plate, the two sets being separated by a dark zone. The fringes are most widely separated close to the plane of the plate. If the plate is slightly wedge-shaped, it has focal properties, for the internally reflected rays change their angle of incidence by each reflection, the angle becoming less and less if the radiation is travelling from the thick to the thin edge of the wedge, and the emerging rays make greater and greater angles with the surface and consequently converge slightly. Absolute

plane-parallelism of the plate is thus not essential, the only effect of the wedge form being to alter the position of the focal plane in which the fringes are formed. In the treatment of the diffraction grating it was shown that a plane grating may have focal properties as a result of unequal spacing of the lines.

It is generally stated that the light should be polarized to start with, since, when the electric vector is parallel to the surface, the reflection is stronger, and we obtain a greater number of multiply reflected rays which give increased resolving-power. This may be accomplished by placing a Nicol prism with its short diagonal parallel to the surface of the plate, immediately in front of the small prism. The gain in resolution, at least with a quartz plate, is not very great, however. Quartz plates are preferable to ones of glass, since if cut from a perfect crystal, they are absolutely homogeneous, while even the best optical glass is never perfect in this respect. They can also be used for investigations in the ultra-violet. Since quartz is doubly refracting two superposed sets of fringes are obtained, due to the two polarized components of the light. Either set can be suppressed by polarizing the light to start with, but if we are to work in the ultra-violet, a Rochon or Wollaston prism of quartz must be used in place of a Nicol.

The author has found it more convenient to dispense with the prism and eliminate one of the images by total reflection. This can be done by placing a horizontal slit, two or three millimetres in width with a collimating lens, in front of the light source and tilting the plate by raising or lowering the end next to the telescope, until one of the two sets of fringes disappears. The one to be retained is the one which can be seen through a Nicol with its short diagonal horizontal. Or we may use a horizontal mercury arc or vacuum-tube placed at a distance of a metre or more. To understand fully what happens when the plate is tilted we must examine a little more fully the path of the polarized rays of different colors in the quartz plate. The reader is supposed to have a knowledge of the elementary principles of polarization and double refraction, if not the following treatment had better be omitted until the chapters dealing with these subjects have been read. In the quartz Lummer plate, at least in both specimens in the author's possession, the optic axis is parallel to the long dimension of the plate. The incident light after entering the terminal prism and suffering total reflection from its under surfaces traverses the plate at an angle with the optic axis, and therefore suffers double refraction, the two oppositely polarized rays emerging from the upper surface at different angles. To determine these angles in their relation to the direction of the vibration (electric vector) we may proceed as

follows making use of reversed rays. Suppose two parallel rays from A and A' (wave-front B) incident on the plate at the angle shown in Fig. 199. When the lower end of the wave-front strikes the plate at C two secondary wavelets will spread out from the point of incidence, a sphere, and a spheroid as shown. In the case of quartz, the spheroid is within the sphere. Draw tangent lines from the point C' where the ray A intersects the plate to the sections of the sphere and spheroid, and join the point C with the two points of tangency, thus obtaining the two refracted rays inside of the plate. Now it is known that the electric vector (which is parallel to the short diagonal of a Nicol in the case of light transmitted by the Nicol) is, in the case of the spheroidal wave, in a plane containing the optic axis, and tangent to the wave-front. The direction of the electric vector is shown by the arrow drawn across the refracted ray. If we consider a ray from a source of light entering the plate, through the auxiliary prism in a direction coincident with this ray, its polarized component with electric vector in the plane of the paper will be refracted out into the air

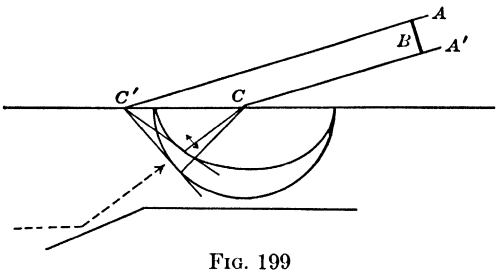


FIG. 199

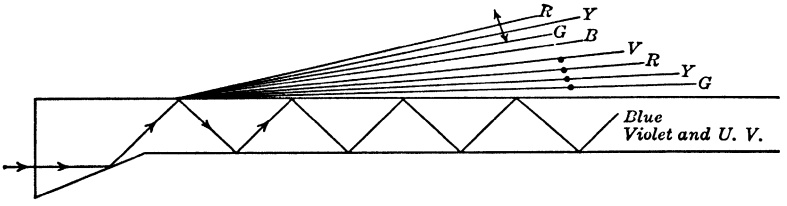


FIG. 200

along CA'. The other component, with its electric vector perpendicular to the plane of the paper will be bent down still more towards the plate, as is clear if we consider the other refracted ray reversed in direction after having been turned up into coincidence with the ray just considered. The result of this is that when light from a narrow horizontal source, collimated by a lens enters the auxiliary prism, two sets of polarized rays of different colors are reflected out into the air as shown in Fig. 200 with electric vectors as shown by the arrows, and by the black dots, which mean electric vectors

perpendicular to the paper. In the case as figured the violet rays of the lower set are totally reflected back and forth within the plate, and do not escape until they reach the rough ground end, where they are diffused. With a wider source of light the two sets overlap. The lower set of rays is the one to be used for reasons previously specified. If we raise the source of light, or what amounts to the same thing, raise the end of the plate nearest the telescope, the colored rays of the lower set will, in turn, enter the plate, and we shall see the interference fringes formed by the upper set. The student should familiarize himself with these changes both by placing the eye a trifle above the prolongation of the upper surface of the plate, using a Nicol prism, and raising and lowering the end of the plate. The multiple images of the source are seen in this way if the eye is focussed on the plate. A telescope should then be substituted for the eye. A mercury arc is the most convenient source of light.

With the plate adjusted for visible light, the ultra-violet is totally reflected to and fro within the plate and none escapes, at least with a source of restricted height (as with a horizontal slit). To bring the ultra-violet fringes into existence it is necessary to raise the end of the plate next to the telescope. This will be clear from Fig. 201 in which the *dispersion* of an incident white ray on

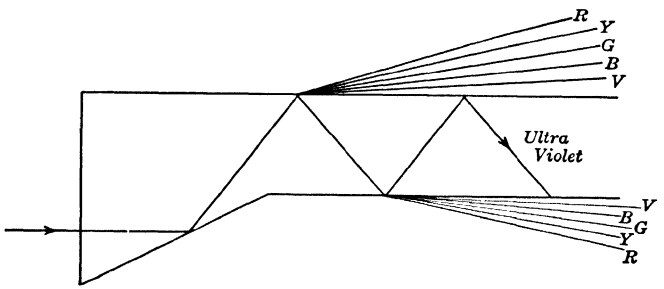


FIG. 201

emergence from the plate is illustrated. This will make the incidence angle on the upper surface less and the ultra-violet emerges. The same thing can, of course, be accomplished by raising the horizontal slit.

For studying the fine structure of spectrum lines, the best disposition of apparatus, is to focus an image of the fringes on the slit of a quartz spectrograph, the slit of which should be opened as wide as is consistent with the separation of the spectrum lines under investigation. The widened spectrum lines will then, in general, be found to be made up of narrow *horizontal* lines, some

strong and some faint, the faint lines being the companions or "satellites," as indicated in Fig. 202.

We will now derive the equation for the relative retardation of two adjacent rays (order of spectrum) incident at angle r within the plate (of refractive index μ), and refracted out into air at angle i. We derived this expression as a function of r in the section on interferences by thick plates, but in the present case we require it in terms of i, and the thickness of the plate t.

If $n\lambda = 2\,\mu t \cos r$ (see page 196)

n will be the "order" of the bright fringe formed by rays leaving the plate at angle i. This corresponds to the "order" of spectrum in the case of a grating

$$n\lambda = 2\,\mu t \sqrt{1 - \sin^2 r}$$

or, since $\mu \sin r = \sin i$,

$$n\lambda = 2t\sqrt{\mu^2 - \sin^2 i} \qquad (1)$$

the fundamental equation of the Lummer-Gehrcke interferometer

The angular separation between two consecutive orders is given by

$$\Delta i = -\frac{\lambda\sqrt{\mu^2 - \sin 2i}}{t \sin 2i} \text{ (for } \Delta n = 1)$$

and is inversely proportional to the plate thickness, and increases with λ, and also as grazing emergence is approached.

The dispersion $\Delta\lambda = -\dfrac{n\lambda^2}{n^2\lambda - 4t^2\mu\dfrac{\delta\mu}{\delta\lambda}}$ (for $\Delta n = 1$).

For rays near grazing emergence this expression can be given a more convenient form, by dividing it by $4t^2$ and multiplying by λ.

$$\frac{\dfrac{1}{2t}\cdot\dfrac{n\lambda^2}{2t}}{\dfrac{n^2}{4t^2}\lambda - \mu\dfrac{\delta\mu}{\delta\lambda}} = \frac{\lambda^2}{2t}\cdot\frac{n\dfrac{\lambda}{2t}}{\dfrac{n^2\lambda^2}{4t^2} - \mu\lambda\dfrac{\delta\mu}{\delta\lambda}}$$

$$\text{or } \Delta\lambda = \frac{\lambda^2}{2t}\cdot\frac{\sqrt{\mu^2 - \sin^2 i}}{\mu^2 - \sin^2 i - \mu\lambda\dfrac{\delta\mu}{\delta\lambda}} \qquad \left(\text{since } \frac{n\lambda}{2t} = \sqrt{\mu^2 - \sin^2 i}\right)$$

and for grazing emergence, since then $\sin i = 1$

$$\Delta\lambda = +\frac{\lambda^2}{2t}\cdot\frac{\sqrt{\mu^2 - 1}}{\mu^2 - 1 - \mu\lambda\dfrac{\delta\mu}{\delta\lambda}}.$$

In this expression $\Delta\lambda$ is the difference of wave-length which a companion line must have in order to coincide with the main line of the next order.

The range of dispersion (except for the thickness of the plate) is the same for all quartz plates. The two following tables give $\Delta\lambda$ for the ordinary and extraordinary rays for a plate 5 millimetres thick.

For a plate of any other thickness D these values are to be divided by $2D$ expressed in centimetres.

SEPARATION OF ORDERS

ORD. RAY * WAVE-LENGTH	5-MM. PLATE	EXTRAORD. RAY * 5-MM. PLATE (CAL. BY A. SCHRAMMEN, *Ann. d. Phys.*, *83*, 1164)
2000	.0246	—
2500	.0431	.042
3000	.0669	.065
3500	.0959	.093
4000	.1276	.125
4500	.1627	.162
5000	.2024	.202
5500	.2481	.246
6000	.2991	.295

* Values for ordinary ray reduced to fit 5-mm. plate, from MacNair's table for a 4.4-mm. plate (*Phil. Mag.*, *2*, 614, 1926). Values for extraordinary ray calculated by Annelise Schrammen, *Ann. der Phys.*, *83*, 1164.

Curves are to be drawn from these tables from which the separation for any intermediate λ can be determined.

To determine the main line to which a given companion belongs, *i.e.* whether it is on the long-wave-length side of the adjacent order, it is necessary to compare photographs made with plates of different thicknesses. The short-wave-length components appear on the side towards which the distance between the orders is increasing. In Fig. 202, if component B is of shorter λ than the main component A, the appearance with the thin plate will be as indicated by the middle diagram. If, however, it is a longer λ component belonging to C, the appearance with the thin plate will be as figured in the diagram to the right. We measure the distances a, b and c of the middle diagram. The average of the distances b and c is taken as the separation of orders in millimetres. Then $a \Big/ \frac{b+c}{2}$ times the separation of orders in Ångström units (from the curve

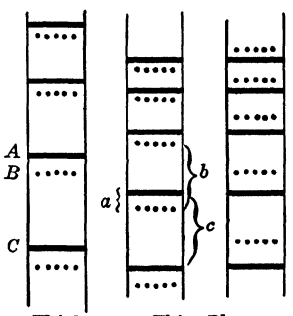


FIG. 202

plotted from the table) gives us the difference of λ between the main line and its companion. This approximation method was the one used by MacNair.[1]

Another and more accurate method was used by McLennan and McLeod [2] who photographed the two sets of fringes (formed on opposite sides of the plate simultaneously). This enabled them to determine the distance of a given fringe from the central plane which divided the two systems, which, taken together with the focal length of the lens, gives the emergence angle e of the ray system forming the fringe. In making use of this method, the plate is employed standing on edge, an advantageous position, as there is less danger of flexure and less liability of a temperature difference between the two surfaces. If $a_1, a_2, a_3 \cdots a_n$ are the distances of successive fringes of the main line from the median plane, and $b_1, b_2, b_3 \cdots b_n$ the distances for a companion line, the wave-length difference between the two is given by

$$d\lambda = \frac{(b_n + a_n)(b_n - a_n)}{(a_{n+1} + a_n)(a_{n+1} - a_n)} \Delta\lambda.$$

More accurate values are obtained by measuring the distances between corresponding orders of the main lines and of the satellites on photographs of the complete fringe system, made by rays issuing from both sides of the plate. These values, divided by 2, give us the distances of the fringes from the median plane. The following formula was used by McLennan and McLeod,

$$d\lambda = \frac{a_{m_1}^2 - a_s^2}{a_{m_1}^2 - a_m^2} \lambda$$

for a companion line of longer wave-length than that of the main line, while for a line of shorter wave-length, the numerator becomes $a_s^2 - a_m^2$, the shorter λ companions falling on the side towards the median plane.

In this expression a_m is the distance of the main line (of any given order) from the centre of the pattern or median plane. a_{m_1} is the distance of the next higher order of the main line, *i.e.* the one further removed from the median plane, and a_s is the distance of the satellite of same order as a_m. λ is the change of wave-length which would be necessary to cause a fringe of given order to drift to the position of the next higher or lower order. The resolving-power, for plate of length L

$$\frac{\lambda}{\Delta\lambda} = \frac{L}{\lambda \sin i}\left(\mu - \sin i - \mu\lambda \frac{\delta\mu}{\delta\lambda}\right).$$

[1] *Loc. cit.*

[2] McLennan and McLeod, *Proc. Roy. Soc.*, *90*, 243, 1914.

In this expression $\Delta\lambda$ is the smallest difference of wave-length which two lines of wave-length λ can have and still be seen separated.

The expression shows us that the resolving-power is proportional only to the length of the plate, and is independent of the thickness. This at first sight perhaps strikes us as curious, as increasing the thickness increases the retardation and order of spectrum. We must note, however, that, for a plate of given length if we increase the thickness we cut down at once the number of possible internal to-and-fro reflections, and the number of emergent rays upon which the narrowness of the fringes depends.

The derivation of these expressions will be found in a paper by Simeon [1] in which the following values are given for plates of two different thicknesses and two angles of emergence, for the green mercury line $\lambda = 5461$ and $\mu = 1.51$.

t	i	n	$\Delta i \cdot 10^3$ Radians	$\frac{\lambda}{\Delta\lambda}$
4.5 mms.	80	18,865	0.406	305,900
	88	18,654	1.968	294,500
10 mms.	80	41,923	0.183	705,900
	88	41,454	0.885	679,700

It is clear that, with a single plate, we cannot determine the real structure of a line, for there is always the possibility that some companion lines or satellites may be missed as a result of their coincidence with the main line of the next or higher orders.

The method of crossed plates, introduced by Gehrcke solves this difficulty.

Interference Points of Crossed Plates. — If the light which emerges from the upper surface of the plate is received by the prism of a second plate set on edge and pointing down slightly towards the surface of the horizontal plate, the horizontal sharp maxima formed by the first, will be cut across by the dark minima formed by the second, and the illuminated field at the focus will be reduced to a multitude of bright points or dots, arranged in a rectangular pattern. If the plates have unequal thickness, the satellite dots will lie along a diagonal line which does not cut through the dots of higher order. If the thickness of the plates is the same, nothing is gained, as the satellite and main dot again coincide. This will be clear from Fig. 203. Suppose the real structure to be as shown at *a*, the satellites being represented by dotted lines. The structure as shown by one plate mounted hori-

[1] *Journal Sci. Instruments*, i, 298.

zontally, is indicated by *b*; here the third satellite coincides with the next order of the main line, and the fourth satellite coincides with the second of the next order. A second plate (mounted vertically) shows the structure as represented by *c*, the fourth satellite coinciding with the main line of the next order. This is approximately the state of things with the 2536 line of mercury, in the case of the two quartz plates used by the author. Now suppose the two plates to act simultaneously in the "crossed" position: pattern *b* will be furrowed by the broad interference minima of pattern *c*, as indicated by *d*.

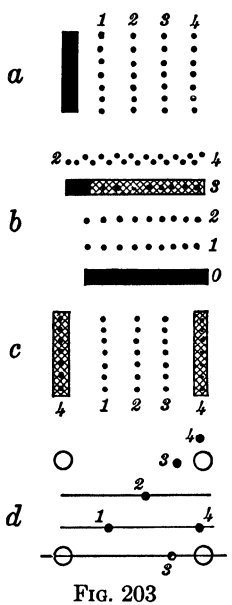


FIG. 203

The lower bar of *b* is made up of the wave-lengths of the main line and the third satellite. The main line maximum formed by the second plate is the lower left-hand circle of *d*. The maximum for the wave-length of the third satellite will fall to the right of this circle at a distance equal to the distance of satellite 3 from the main line in *c*, followed by the circle representing the main line of the next order. Dotted lines 1 and 2 of *b* will give maxima in the position shown by the next two lines to the lower line of *d*, being displaced to the right by the distance separating line 1 of *c* from the main line. Line 3 of *b* is broken up in the same way as the lower line, and line 4, which is made up of satellites 1 and 4 gives dots at distances from the main dot equal to the distances of lines 1 and 4 in *c*.

The satellites are now spread out along a diagonal line, and the true structure is at once revealed. If we make the same construction for plates of equal thickness, we shall find that the same coincidence occurs for "dots" as obtained with the fringes.

The technique of working with crossed plates in the ultra-violet is a little involved. The following hints may prove useful.

The first plate, in horizontal position, is so adjusted that only one of the polarized sets of rays emerges. The second plate should be mounted as close as possible to the first in the vertical position (*i.e.* "on edge") on some sort of turn-table (preferably a small graduated circle) and pointing down slightly so as to receive, on its prism, the rays leaving the first plate at nearly grazing emergence. Obviously the direction of the electric vector, if favorable for the first plate, will be unfavorable for the second, and it would seem desirable to have one plate cut with the optic axis parallel to

the breadth. The gain, however, is not very great, and perfectly satisfactory dots can be obtained with two plates cut in the same way. A quartz lens of about 30-cm. focus is next mounted close to the second plate and the slit of the spectrograph brought up into the interference pattern at its focal plane. On bringing the eye into the position of the photographic plate we adjust the instrument, by raising or lowering its feet with blocks or fragments of glass, until the brilliant spot of light is at the centre of the iris diaphragm. The slit should be opened to a width of one or two millimetres. We now examine the pattern of the interference dots in the green Hg line with a short-focus lens, adjusting the plates in such a way that the dots form a square pattern, with say four or five dots in the vertical rows. We may have to raise or lower the quartz lens, which should be mounted at the end of a tube forming a telescope without an eye-piece. We now slip pieces of thin glass under the foot of the first plate, abolishing first the green and then the violet line. The first plate is now probably transmitting the 2536 line, but as no visible light gets through, we have no means of adjusting the second plate. To do this we remove the glass plates from under the foot of the first plate, and then rotate the second plate in a clockwise direction until the green line disappears, the violet remaining: a further rotation of about one degree will bring the plate into the correct position, and we now replace the glass plates and make a trial exposure.

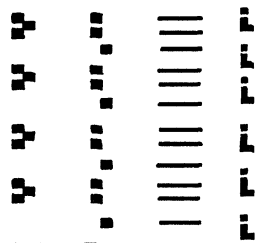


FIG. 204

In studying the fine structure of lines slightly or barely separated by the spectroscope, it is advantageous sometimes to employ the less efficient of the two polarized rays. The reason for this is made clear by Fig. 204. With the ordinary ray the single component of the right-hand line comes midway between the two components of the left-hand one, and will not appear separated from them unless the slit of the spectroscope is made very narrow. This makes intensity measurements with a density photometer difficult. With the extraordinary ray, the component of the right-hand line falls (in this case) midway between the orders of the two component lines, and will be clearly separated from them with a wide slit as shown. It is often possible to see, by a close inspection which fringes belong to the respective components of a close double line not actually resolved by the spectroscope, the edges having a saw-tooth appearance as shown at the right of Fig. 204.

If exposures of several hours are necessary, it is absolutely essential to hold the temperature of the plate constant to 0.01° or less by means of a thermostat; mounting the plate in a double-walled metal box, with water between the walls, a method often employed, is not recommended, as the water temperature rises almost as fast as the room temperature if the latter is increasing at the rate of only one or two degrees per hour.

Photographs of the fringes obtained with a quartz plate of a close group of mercury lines in the ultra-violet are reproduced on Plate 8*c* in the Chapter on Resonance Spectra of Atoms. These were taken by projecting the fringes on the wide-open slit of a quartz spectrograph, incapable of resolving the group even with a fine slit.

CHAPTER IX

THE POLARIZATION OF LIGHT

The simplest type of wave-motion is one in which the direction of vibration, or of electric displacement, is to-and-fro along a line. In the case of transverse, as distinguished from longitudinal waves, there will be lack of symmetry around their direction of propagation.

Natural or ordinary light shows no such lack of symmetry, and can be regarded as a wave-motion in which the direction of the electric vector suffers changes in orientation too rapid to be detected by any means at our disposal.

It is possible, however, to obtain from natural light a radiation having vibrations of fixed type and orientation, and the behavior of such light when it encounters matter depends upon this orientation.

For example, it is possible to obtain light which a glass or water surface refuses to reflect at a certain angle of incidence. Such light is said to be polarized, and the polarization may be one of three types, plane, elliptical or circular according to the type of vibration.

Plane of Polarization. — What was formerly known as the plane of polarization was the plane of incidence in which the polarized light was most copiously reflected. This definition was given at a time when little was known about the nature of the vibrations constituting light, and we now know that the plane thus specified is at right angles to the direction of the electrical vibrations of the light. We shall, therefore, employ the term to define the direction of the electric vector, and have nothing more to do with the old definition.

Discovery of Polarization. — The polarization of light was discovered by Huygens in 1690, while experimenting with Iceland spar. He found that a ray of light was, by passage through the crystal, divided into two separate rays of equal intensity, except when the light traversed the crystal in a direction parallel to the crystallographic axis. He found furthermore that if one of these emergent rays was passed through a second crystal, it was divided into two rays of equal or unequal intensity, or not divided at all, according to the orientation of the crystal. Though this single

experiment was sufficient to establish the existence of light which was asymmetrical around its line of propagation, and though many other crystals exist having similar properties, Huygens was ignorant of the nature of the phenomenon, and the discovery remained an isolated fact for more than a century. The polarization is in this case produced by double refraction, which we shall study in detail in a subsequent chapter.

Polarization by Reflection. — The discovery was made by Malus in 1810 that light, which had suffered reflection at a certain angle from a surface of water or glass, exhibited the peculiarities, previously observed only in the case of light after its passage through a crystal of Iceland spar. The polarization of light by reflection can be exhibited by means of the easily improvised apparatus shown in Fig. 205. The reflectors A and B are made of ordinary plate glass. The polarizing plate A is mounted on an iron stand on a hinged support so that it can be set at the polarizing angle. The other reflector is mounted on the vertical axis of an ordinary turntable, in such a way that the light reflected down from the polarizer meets the surface of the glass at an angle of 57° with the normal, *i.e.* the plate must make an angle of 33° with the axis. A cylindrical ring of parchment paper or other translucent medium surrounds the revolving plate, and receives the light reflected from it. The upper plate is so adjusted that its plane is parallel to the plane of the lower plate, in which position it will be found to reflect light capable of reflection from the latter: if, however, the lower plate is turned through an angle of 90°, it will be found that the light is no longer reflected from it, while in intermediate positions of the plate the reflection is partial. If the plate is rotated a ring of light is seen on the translucent screen with two maxima and two minima, corresponding to the positions in which the light is most and least copiously reflected.

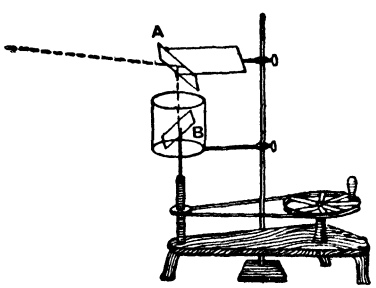


FIG. 205

Angle of Polarization. Brewster's Law. — If the angle of the upper mirror is varied, it will be found that the reflected light is less completely polarized, and the maxima and minima obtained by revolving the lower plate are less marked. In general, as we increase the angle of incidence from normal to grazing, the polarization increases, passes through a maximum, and then decreases. The angle at which the polarization is most complete varies with

the nature of the substance, and is known as the polarizing angle. Jamin found that only a few substances with a refractive index of about 1.46 completely polarize the reflected light. For all other substances the polarizing angle is merely the angle at which the polarization is a maximum.

The relation between this angle and the refractive index of the substance was investigated by Brewster, who discovered the remarkable law that the index of refraction was the tangent of the angle of maximum polarization. When the light is incident at this angle, the refracted ray makes an angle of 90° with the reflected ray, for

$$\frac{\sin i}{\sin r}=\tan i=\frac{\sin i}{\cos i};$$

$$\therefore \cos i=\sin r \text{ and } i+r=90°.$$

If this law is true, the angle of maximum polarization will be different for the different colors owing to dispersion. In the case of most transparent media the dispersion is too small to affect the angle appreciably, as can be shown by examining the image of the sun reflected in a glass plate through a Nicol prism so orientated as to cut off most completely the reflected light. The image of the sun appears uncolored, which would not be the case if the angle of polarization varied for different parts of the spectrum. The Nicol prism, which will be presently described, takes the place of the second reflector, having the property of completely cutting off light polarized in a certain plane, and transmitting with greater or less facility light polarized in all other planes. In the case of substances having very high dispersion, the variation of the angle with change of wave-length becomes very marked.

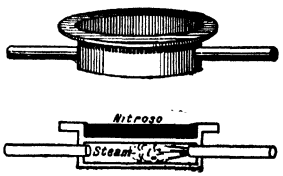


FIG. 206

The fact that light can be obtained having a lack of symmetry around the direction of propagation is one of the most direct and convincing proofs which we have of the transverse nature of the waves, for we cannot very well conceive of a pressural or longitudinal wave, having different properties in the different directions perpendicular to the line of propagation.

The organic compound nitroso-dimethyl-aniline, which has been found by the author to have, in the brighter parts of the visible spectrum, the highest dispersion of any known substance, is admirably adapted for the exhibition of what may be termed the dispersion of the angle of polarization. A little of the substance is fused on a glass or metal plate, or better in a small brass cell heated

by steam (Fig. 206).[1] A brilliant source of light of small dimensions is arranged so that its light is reflected from the liquid surface at an angle which can be varied.

On examining the light reflected at a fairly large angle with a Nicol prism, it will be found to vary from light blue to deep violet and purple, as the angle of incidence is increased, the Nicol being held in such a position as to refuse transmission to the light polarized by reflection. If a small direct-vision spectroscope is placed behind the Nicol, a dark band will be seen crossing the spectrum, which shifts its position as the incidence angle varies. The centre of this dark band evidently marks the wave-length for which the angle of incidence happens to be the angle of maximum polarization, or in other words the refractive index of the substance for this wave-length is the tangent of the angle of incidence. In the case of glass and substances of low dispersion, the different colors are polarized at nearly the same angle, *i.e.* very little color effect is observed when the reflected light is examined with the Nicol. In these cases the dark band is so broad as to occupy practically the entire visible spectrum. In the case of a substance with as high a dispersion as that of the nitroso, the angle of maximum polarization is quite different for the different colors; consequently the Nicol prism only extinguishes a portion of the spectrum for a given angle of incidence. This gives us a reflection method of determining the refractive index of a substance, for by determining the angle of incidence for which the centre of the dark band is at a given point in the spectrum, we have only to look up the tangent of the angle in order to get the refractive index for the wave-length in question. The band will be found to be very sharp and quite narrow when it occupies a position in the green and greenish-blue, but on attempting to drive it into the red, we shall find that it broadens and becomes much less sharply defined. This is of course due to the fact that the dispersion is much less in the red and orange portion of the spectrum. If the nitroso cannot be obtained, selenium plates, made by pressing the molten substance between glass plates, which are to be separated by a blow from a hammer when cold, can be used for the exhibition of the dark band in the spectrum.

Polarization by Refraction. — If we examine the light transmitted through a plate of glass placed at the polarizing angle, we shall find that the light is partially polarized; *i.e.* its intensity varies slightly when examined by means of an analyzer. Arago discovered

[1] Nitroso-benzyl aethyl aniline, which can be obtained from the Berlin Aniline Co., is better than the nitroso-dimethyl compound, as after fusion it remains liquid for some hours at ordinary temperatures

that the reflected and refracted portions of the light contained equal quantities of polarized light, and that the planes of polarization were at right angles.

The greater intensity of the transmitted light is responsible for the incompleteness of the polarization. If the light transmitted through a plate placed at the polarizing angle is received upon a second plate, the unpolarized portion suffers a further resolution into two polarized components, one of which is reflected out through the upper plate and the other transmitted. By increasing the number of plates we can increase the intensity of the reflected polarized light, and consequently the completeness of the polarization of the transmitted light, seven or eight being sufficient to give us nearly complete polarization in the transmitted, as well as in the reflected beams.

A simple polariscope can be constructed on this principle. The glass plates used should be as thin as possible, in order to avoid loss of light by absorption. The large-sized rectangular cover-glasses used for microscopical preparations are best for the purpose, though the thin glass employed for lantern slides is almost as good. The plates should be carefully cleaned and freed from dust, and mounted in two piles, of eight plates each, in tubes of wood or pasteboard at an angle of about 33° with the axis of the tube. It is best to determine the angle experimentally, as it varies slightly with the nature of the glass. The two piles of plates should almost completely cut off light when the planes of incidence are at right angles. Bundles of plates thus mounted form very fair substitutes for the more expensive Nicol prisms, and are well adapted to lantern experiments.

Law of Malus. — When a beam of light, polarized by reflection at one plane surface, is allowed to fall upon a second, at the polarizing angle, the intensity of the twice-reflected beam varies as the square of the cosine of the angle between the two planes of reflection. The assumption was made that the incident vibration, polarized in a plane making, say, an angle θ with the plane of incidence, was resolved into two components, one perpendicular, the other parallel to the plane of incidence, the former being partially reflected, and the latter wholly transmitted. This will make the reflected amplitude $a \sin \theta$, if a is the reflected amplitude when $\theta = \pi/2$, and the intensity will be $a^2 \sin^2 \theta$, or the maximum reflected intensity multiplied by the $\sin^2$ of the angle between the plane of polarization and the plane of incidence. On the old definition of the plane we have cos instead of sin.

The law of Malus is therefore simply a statement of the resolution of a vibration into two rectangular components, the direction

of the vibration being considered the plane of polarization, on the notation given above.

It is important to distinguish between the behavior of vibrations parallel to the plane of incidence, and vibrations perpendicular to the plane, when they meet a reflecting surface at the polarizing angle. If the light is so polarized that the vibrations are perpendicular to the incidence plane, *i.e.* parallel to the reflecting surface, a portion of the energy will be reflected, and a much larger part transmitted, the directions of the vibrations remaining parallel to the incident vibration (Fig. 207*a*). If, on the other hand, the direction of vibration is parallel to the incidence plane, practically no energy is reflected, the light being refracted without loss of intensity (Fig. 207*b*). If now the vibration takes place in a direction making, say, an angle of 45° with the plane of incidence, it will be resolved into two components, one parallel to the reflecting surface and the other parallel to the plane of incidence.

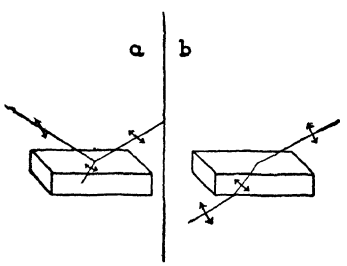


FIG. 207

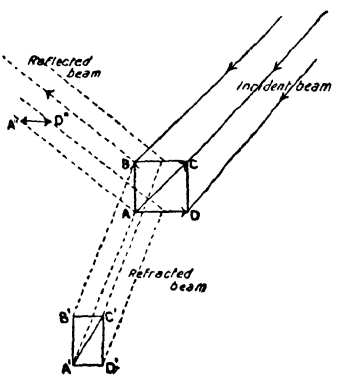


FIG. 208

Let *ABCD* (Fig. 208) represent a portion of the wave-front of the incident beam, which is coming towards us, the direction of vibration being *AC*. We have resolution into the components *AB* parallel to the incidence plane, and *AD* parallel to the reflecting surface. The former is wholly transmitted (*A'B'*), the latter in part reflected and in part transmitted (*A'D'*). The reflected light is therefore polarized with its vibration parallel to the surface, since only this component is reflected; the refracted light is made up of the completely transmitted component *A'B'*, and the partially transmitted component *A'D'*, their resultant being *A'C'*, a polarized vibration, rotated counterclockwise with respect to the incident vibration *AC*. If now the resultant *A'C'* be received on a second reflecting surface, the same resolution will take place, and there will be a further rotation of the plane. The effect of a pile of plates will therefore be to bring the plane of vibration of the transmitted light

into coincidence with the plane of incidence, since $A'B'$ is transmitted each time without loss, while $A'D'$ is reduced in intensity by the partial reflection. This rotation of the plane of polarization is clearly the result of the reduction in the intensity of one of the rectangular components, and may be shown best by means of a pair of Nicol prisms, so oriented as to refuse transmission. If a glass plate, or better, a pile of three plates, be placed between the prisms at the polarizing angle and so oriented that the plane of incidence is inclined at 45° to the principal planes of the Nicols, there will be a partial restitution of light, and the analyzing Nicol will have to be turned counterclockwise to produce complete extinction.

It is clear now why the transmitted light is only partially polarized, when a ray is refracted at a single surface, and completely polarized by refraction at a large number of parallel surfaces. We may consider ordinary light as consisting of vibrations polarized in all possible planes. Each vibration is therefore transmitted with a slight rotation towards the plane of incidence, and the light will not differ greatly in its properties from ordinary light. By every succeeding surface there is a further rotation, and eventually all are brought into the plane of incidence and the transmitted light is plane-polarized. If the vibrations of the incident light be represented by A (Fig. 209), the effect of successive refractions may be represented by the succeeding diagrams B, C, D, E, the transmitted light being plane-polarized in the latter case. The foregoing elementary treatment of polarization by reflection will suffice for the present. The theoretical treatment of the subject will be given in a subsequent chapter.

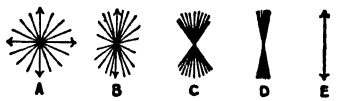


FIG. 209

Polarization by Double Refraction. — The double refraction of light by crystals of Iceland spar was first noticed by Erasmus Bartholinus, and subsequently more completely investigated by Huygens. It occurs whenever light enters the crystal in a direction not parallel to the optic axis, and is due to the fact that the incident vibration is decomposed into two mutually perpendicular components which travel through the crystal with different velocities. One of the two rays obeys the ordinary laws of refraction and is called the ordinary ray, while the other behaves in a most peculiar manner and is called the extraordinary ray, for it is bent away from the normal *even at perpendicular incidence.* At first sight this appears impossible, for even if the two rays have different velocities there seems to be no reason why there should be a change of *direction* for normal incidence. As we shall see when we come to the

Chapter on Double Refraction, this is due to the fact that the wave-front of the extraordinary disturbance is not spherical but ellipsoidal. For the present, however, we are not concerned with this question, and merely consider the crystal as a means of resolving ordinary light into two polarized beams. One of these may be cut off by any suitable device, leaving plane-polarized light. Crystals of tourmaline have the remarkable property of absorbing the ordinary ray and transmitting the extraordinary, consequently a thin section of a crystal transmits only polarized light, and may be used either as a polarizer or analyzer. The sections are cut parallel to the optic axis, and when superposed with their axes parallel transmit light quite freely. If one is rotated through a right angle, the combination becomes opaque, since the polarized vibrations transmitted by one are absorbed by the other. The tourmaline polariscope is a very simple instrument, consisting of a pair of crystal sections mounted in a pair of wire tongs in such a way that one of them can be rotated in front of the other. The object to be examined, for example, a mica or selenite film, is placed between the two tourmalines, and the instrument directed towards a bright light; owing to the deep color of the tourmaline crystals, this form of polarizer and analyzer is very inferior to the Nicol prism, which only cuts down the intensity of the original light one-half.

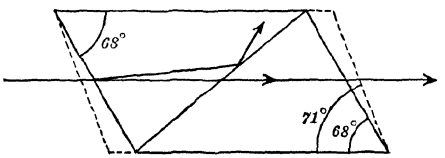


FIG. 210

The Nicol Prism. — Iceland spar, on account of the large size in which the crystals occur, and their great transparency, is especially suitable for the construction of polarizing prisms. Since both the ordinary and extraordinary rays are transmitted with equal facility, it becomes necessary to eliminate the one or the other by some optical contrivance. The method employed by Nicol was to stop the ordinary ray within the crystal by total reflection.

The two end faces of a natural rhomb of spar are cut down sufficiently to reduce the normal angle of 71° between the faces to 68° as shown in Fig. 210 and the crystal is then cut in two along the solid line which is perpendicular to the new end faces. After polishing the cut surfaces they are cemented together with Canada balsam. The refractive index of the balsam is intermediate between the indices of the spar for the ordinary and extraordinary rays, and the former meeting the balsam film at an angle greater than the critical angle is totally reflected to one side and absorbed by a coating of black paint.

The prism therefore transmits only the extraordinary ray, which is plane-polarized, the direction of the vibration being parallel to the short diagonal of the prism, *i.e.* perpendicular to the balsam surface. In some prisms it is impossible to tell by mere inspection in which direction the emergent light vibrates, owing to the way in which the crystal has been cut. It is always possible, however, to determine in a moment the direction in question. We have only to reflect ordinary light from a glass surface in the neighborhood of the polarizing angle and examine it through our prism, holding it in such a position that the transmitted light has its maximum intensity. The vibrations of the light are parallel to the glass surface, consequently the diagonal of the prism which is parallel to this surface is the direction in which the transmitted light vibrates. If any difficulty is experienced in fixing in the mind the direction of vibration or electric vector in the case of reflected and transmitted light, the following analogy may prove useful. If a cylindrical wooden rod is thrown in an oblique direction and with great force upon the surface of water, it will bounce off if parallel to the surface, the rod being supposed to move in a direction perpendicular to its length. If, however, the rod is perpendicular to the surface, the lower end will enter the water first and the rod "cut down" into the fluid, without suffering reflection. Consider our polarized vibrations as parallel to the rod, and we have the optical analogy, which is only of use, however, in enabling us to remember the direction of the vibration in the reflected and refracted components.

The Nicol prism is sometimes made with end surfaces perpendicular to the axis. This necessitates a more oblique section, and consequently a greater length in proportion to the width, the ratio being nearly 4 : 1. The Foucault prism is similar to the Nicol, except that the balsam film is replaced with an air film. This device reduces the ratio of length to width to 1.5 : 1, but the prism is less efficient than the Nicol, owing to multiple reflections in the air film. It is useful for work in the ultra-violet region, for balsam absorbs these rays.

Angular Aperture of Polarizing Prisms. — Since many experiments require the passage of a convergent or divergent beam through the prism, it is of some interest to consider the maximum angular aperture of the prism, or maximum divergence which a cone of rays may have, and still be completely polarized by passage through a prism of given type. If the divergence exceeds a certain amount, it is obvious that some of the ordinary rays will not suffer total reflection. The angular aperture of a prism of the Nicol type is about 30°, while that of the Foucault prism is

only 8°. A prism was devised by S. P. Thompson [1] in which the optic axis was perpendicular to the long axis of the prism, which had an aperture of 39°. Glazebrook [2] constructed one along similar lines, but with end surfaces perpendicular to the long axis.

The question of the best construction for a prism with large aperture and end surfaces perpendicular to the long axis was investigated by Feussner,[3] who found that the optic axis should be perpendicular to the section, and the refractive index of the cementing film the same as that of the crystal for the extraordinary ray. Such a prism has an aperture of 42° and a ratio of length to width of 4 : 1.

In another and quite different type of polarizing prism, the doubly refracting substance acts as the rarer medium, the extraordinary ray being totally reflected from a thin plate of Iceland spar immersed in a liquid of higher refractive index. The first prism constructed on this principle was made by Jamin, who immersed a thin plate of spar in a glass trough filled with carbon bisulphide. Zenker improved the device by substituting prisms of flint glass for the liquid, while Feussner suggested the use of a plate of sodium saltpetre instead of Iceland spar, on account of the greater difference between the ordinary and extraordinary refractive index. Such a prism would have an aperture of 56°.

Double Image and Polarizing Prisms for the Ultra-Violet. — The Nicol prism is useless for work in the ultra-violet on account of the absorption of these rays by the film of Canada balsam.

The most satisfactory prisms for work in this region are quartz double prisms of the Rochon or Wollaston type, cemented together with a film of glycerine, or better, castor oil, as the former is decomposed by intense ultra-violet rays. These two types are shown in Fig. 211 the directions of the optic axes being indicated by lines and points. The Rochon transmits the ordinary ray without deviation, the image being achromatic, while the Wollaston deviates both rays and consequently gives a greater separation of the two images, which however are colored. As both rays are transmitted a lens or lenses must be used, and one of the images of the source screened off. It frequently happens, however, that

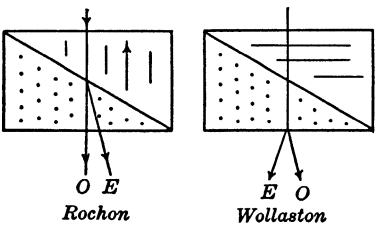


FIG. 211

[1] *Phil. Mag.*, 5, 12, page 349, 1881.
[2] *Phil. Mag.*, 5, 15, page 352, 1883.
[3] *Zeitsch. für Instrkdé*, 4, 41, 1884.

both images are utilized, for example, in determining the percentage of polarization in partially polarized light, which will be considered presently. The direction of the electric vector in the two images can be determined by observing which one is the brighter when a glass plate, reflecting light at angle of about 60°, is viewed through the prism, remembering that the electric vector in this light is chiefly parallel to the glass surface; or we may examine the two images through a Nicol, which transmits light with its electric vector parallel to the short diagonal. In using the Rochon prism it is important that the light enter the component in which it travels along the optic axis, as shown in the figure, otherwise the natural rotation of the quartz will give all sorts of orientations of the plane in the transmitted light.

Detection of Polarized Light. — If the amount of polarized light present in a beam is too small to be detected by the slight changes in intensity produced by passing it through a slowly revolving Nicol, some more delicate method must be adopted.

The Nicol used alone will not give evidence of the presence of less than about 20% of polarization, consequently in cases where the polarization is not considerable (the solar corona, for example), we cannot rely upon its testimony. By the use of the "bi-quartz" (two plates 3.75 mm. thick side by side of R and L quartz $\perp$ to axis), the presence of 5% of polarization may be detected by the slight coloration of the two segments of the plate when it is placed in front of a Nicol and directed towards the light.

Savart's plate and Babinet's compensator are still more sensitive detectors of small amounts of polarization, and are generally employed in the study of the polarization of the sun's corona during total eclipses. The presence of polarized light is shown by a system of colored fringes which appear when the plate is used in the same manner as the bi-quartz. Savart's plate is made by cutting a plane-parallel plate from a quartz crystal at an angle of 45° with the optic axis. The plate is then cut into two halves, which are mounted the one above the other, but rotated through 90° with respect to each other. The sensitiveness is at a maximum, *i.e.* the fringes are most distinct when the direction of the polarized vibration is perpendicular to the fringes, the instrument then being capable of showing the presence of 1% of polarization.

If one is working with monochromatic light a single thin wedge of quartz (1° to 3° angle) cut parallel to the optic axis may be employed in place of the Savart plate. Such a wedge shows interference bands when properly oriented between a source emitting traces of polarized light and a Nicol prism or other analyzer. The distance between the fringes increases with decreasing angle of the wedge.

Degree of Polarization. — In the case of partial polarization we can define the degree of polarization as the ratio of the polarized to the total intensity. If we represent the partially polarized light as a mixture of plane-polarized light of intensity I_p and natural or unpolarized light of intensity I_n the degree of polarization P is given by

$$P=\frac{I_p}{I_p+I_n},$$

i.e. the ratio of the polarized to the total light. The ratio can be measured in the following way. The source of light, say of a small strip of glossy white paper reflecting light at such an angle as to show a glare partially polarized with electric vector in the vertical plane (glare reduced or completely suppressed by a Nicol with its long diagonal vertical), is mounted at such a distance behind a double-image Wollaston prism as to show two images one above the other and in contact, the upper with its electric vector vertical, and the lower with its vector horizontal. These two images are viewed through a Nicol set to extinguish the upper image. By rotating the Nicol in either direction from this position through an angle θ the two images may be equalized in intensity. For greater accuracy we read the double angle 2θ. We call r the ratio of the (weak) horizontal component of intensity to the (strong) vertical component, in the light reflected from the paper, and shown separated in the double image prism. This ratio r is the measure of the imperfection of polarization, or the degree of depolarization studied in the Chapter on the Scattering of Light.

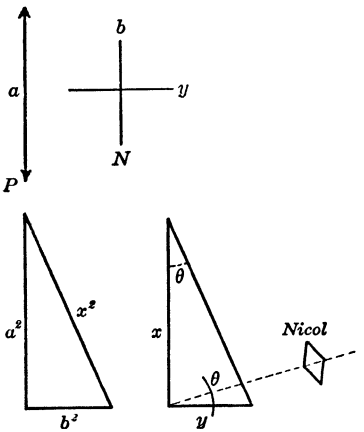


FIG. 212

The light of the partially polarized source can be regarded as a mixture of plane-polarized light P and unpolarized or natural light N; the latter can be regarded as equivalent to a mixture of two equal parts of amplitudes b and y, the former polarized parallel, and the latter perpendicular to the polarization of P, the intensity being b^2+y^2. The amplitudes we can represent by a, b and y in Fig. 212.

If a double-image Wollaston prism is oriented so as to give, say, an upper image containing ligh' P, this image will also have

one-half of light N, the other half going to the lower, or fainter image. We cannot, however, add amplitude b to a, as there is no phase relation between them. We must add their intensities and thus derive the equivalent amplitude. This addition must be made in such a way as to give us the square of the desired amplitude, as shown in Fig. 212 (lower left) where $a^2+b^2=x^2$. The equivalent amplitude of the upper image is thus x, and the amplitude of the lower image is y (the other half of N). With the Nicol oriented to equalize the intensities of the two images as shown in Fig. 212 (lower right) we have the resolutions of x and y on the line (parallel to short diagonal of Nicol) inclined to y at angle θ, equal to each other. Therefore the two right triangles are similar and the upper angle is also θ. We thus have $y^2/x^2 = \tan^2\theta$, which is the degree of depolarization or imperfection of polarization referred to above. We will now deduce the ratio in which P is the percentage of polarization if the denominator is put as 100.

$$\frac{P}{P+N}=\frac{a^2}{a^2+(b^2+y^2)}=\frac{x^2-y^2}{x^2+y^2}=\frac{1-\frac{y^2}{x^2}}{1+\frac{y^2}{x^2}}=\frac{1-\tan^2\theta}{1+\tan^2\theta}, \text{ and}$$

$$(\text{multiplying by } \cos^2\theta)=\frac{\cos^2\theta-\sin^2\theta}{\cos^2\theta+\sin^2\theta}=\cos 2\theta.$$

Other ratios in terms of r or $\tan^2\theta$ are as follows:

$$\frac{\text{Unpol.}}{\text{Pol.}}=\frac{2r}{1-r} \qquad \frac{\text{Unpol.}}{\text{Total}}=\frac{2r}{1+r} \qquad \text{and} \qquad \frac{\text{Pol.}}{\text{Total}}=\frac{1-r}{1+r}=$$

$$1-\frac{2r}{1+r}=1-2\sin^2\theta=\cos 2\theta.$$

Another method frequently used in studying the polarization of the solar corona consists in compensating the polarization by means of one or more inclined glass plates. The same method was used by the author in studying the polarized fluorescence of sodium and potassium vapor. One or more glass plates are placed between the partially polarized source of light and the Savart plate with its analyzing Nicol. If the electric vector of the polarized portion of the light is vertical, the plates are to be rotated on a vertical axis. The plates are rotated until the Savart fringes disappear, *i.e.* until the polarization produced by transmission through the oblique plates exactly compensates the opposite polarization originally present in the source. The angle through which the plates have been turned is read. To avoid setting the plates

for normal incidence, we rotate first to the right and then to the left, thus measuring the double angle. If the angle through which the plates are turned from normal incidence is Φ the angle of refraction x can be calculated if we know the refractive index n of the glass

$$\sin x = \frac{\sin \Phi}{n}.$$

If r and R are the reflection coefficients for rays with their electric vectors in and perpendicular to the plane of incidence, the corre-

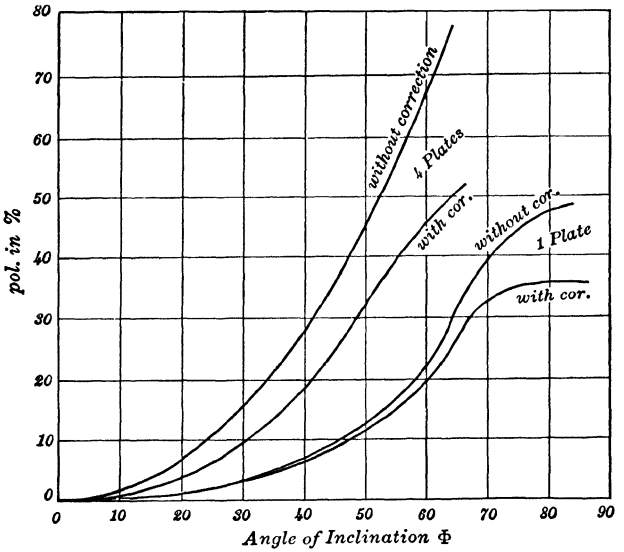


FIG. 213

sponding transmitted energies are $d=1-r$ and $D=1-R$, and the degree of polarization is, for k reflecting surfaces

$$p=\frac{d^k-D^k}{d^k+D^k}=\frac{1-\cos^{2k}(\Phi-x)}{1+\cos^{2k}(\Phi-x)}$$

since by the Fresnel formula

$$d=\frac{\sin 2\Phi \sin 2x}{\sin^2(\Phi+x)\cos^2(\Phi-x)} \qquad \text{and} \qquad D=\frac{\sin 2\Phi \sin 2x}{\sin^2(\Phi+x)}$$

$$\frac{D}{d}=\cos^2(\Phi-x)$$

$$p=\frac{d-d\cos^2(\Phi-x)}{d+d\cos^2(\Phi-x)}=\frac{1-\cos^2(\Phi-x)}{1+\cos^2(\Phi-x)}.$$

Gaviola and Pringsheim [1] have shown that, if the percentage of polarization is considerable, this formula does not give correct results, since, with the plates inclined at the large angle necessary to secure compensation, we have multiple reflections of considerable intensity which must be taken into account. Fig. 213 is reproduced from their paper and gives the percentages of polarization for various angles of inclination for four plates and a single plate ($n=1.52$) computed with and without the correction for multiple reflections.

The Eye as an Analyzer: Haidinger's Brush. — The polarization of light can sometimes be detected by the eye alone. If we look through a Nicol prism at a white cloud, and slowly revolve the prism, a faint blue and yellow double brush appears at the point upon which the eye is fixed, which revolves with the prism. It is not easy to see it at the first attempt, but once noticed, it is easily recognized on subsequent occasions. It consists of four quadrants, colored blue and yellow alternately, and is usually very faint. Various explanations of the phenomena have been given, most of them based on the laminary structure of the lens of the eye. If this were the case, however, the centre of the brush ought to appear a little to one side of the point observed, since the point on the retina which receives the image of the point upon which the eye is fixed lies to one side of the axis of the eye. This fact led Helmholtz to look for the cause of the phenomenon in the structure of the "yellow spot" of the retina, which is the point of the retina alluded to above. He found that the brush was due to the fact that the yellow elements of the spot were doubly refracting, and that the extraordinary rays of blue color were more strongly absorbed than the ordinary rays. For a more complete treatment of the subject the reader is referred to Helmholtz's *Physiological Optics*.

Polarization by Oblique Emission. — Arago found that the light emitted in an oblique direction from the white-hot surface of a solid or liquid was partially polarized in a plane perpendicular to the plane of emission, *i.e.* it resembled light refracted at an oblique surface. From this we may infer that the emitted light comes not only from the surface molecules, but also from those lying below the surface, and that the polarization is due to the refraction of this light when it emerges into the air.

The state of polarization of the light coming from an oblique surface of a body gives us a clew as to the origin of the light and the state of the body. The light of the moon, for example, is partially polarized in a plane passing through the sun, moon and earth, which shows us that the moon shines by reflected light.

[1] *Zeit. für Phys.*, *24*, 24.

If it shone by emission the light would be polarized slightly in the opposite plane. The absence of polarization in the light coming from near the edge of the sun led Arago to infer that it was emitted by an incandescent gas, which is in accord with Schmidt's theory of the solar disk. (See Chapter on Refraction.)

A convenient way of showing this effect is to fuse the end of a rather thick platinum wire in the arc, or oxy-coal-gas flame, forming a sphere 2 mms. in diameter on the end. This sphere is then heated to incandescence in a small blast-lamp flame and a small Nicol prism mounted close to it followed by a very short focus lens such as a pocket triplet magnifying-glass. This projects an image of the ball about 4 cms. in diameter on a sheet of paper distant say 50 cms. The image shows two bright areas near the edge at opposite ends of a diameter, which rotate as the Nicol is rotated. Another method is to wind a narrow strip of platinum foil around a small piece of glass tubing forming a close spiral of three or four turns, which, when removed from the tube gives us the equivalent of a cylinder. This is heated white hot by a current and its image projected in the same manner as that of the ball. The edges are abnormally bright with the Nicol properly oriented.

The theory of polarization by oblique emission was given by Kirchhoff [1] who deduced the polarization by considering the difference in the reflecting power of metals at large angles of incidence for radiation polarized in and perpendicular to the plane of incidence. This in turn depends upon the optical constants of the metals. Czerny [2] has used a platinum strip at 850° C. at a very oblique angle of emission, as a source of almost completely polarized infra-red radiation.

Stationary Polarized Waves. — Wiener found, in repeating his experiments with plane-polarized light, that if the light was incident at an angle of 45° the effects of stationary waves were obtained only when the plane of polarization (term used in its old sense) was parallel to the plane of incidence. Stationary waves can be formed only when the effective vector in the reflected disturbance is parallel to the vector of the incident light, from which we infer that the vector which is effective in producing photographic action is perpendicular to the plane of polarization. The same was found to hold true for fluorescence action. Wiener having already determined that a node of the stationary wave system occurred at the reflecting surface, the inference to be drawn from the experiments with oblique light was that the electric vector was the one concerned in photographic and fluorescent action, and that it was

[1] *Wied. Ann., 109,* 299, 1860.
[2] *Zeit. für Phys., 26,* 182, 1924.

perpendicular to the plane of polarization, or parallel to the plane of polarization as we now use the term. This will be made clearer in the Chapter on the Theory of Reflection.

Landolt's Fringe. — If a brilliant source of light is viewed through a pair of Nicol prisms, so oriented that their principal planes are at right angles, it will be seen that the whole field does not become absolutely dark when exact adjustment is reached, but that the darkened field is crossed by a black fringe which changes its position if either prism is rotated ever so slightly. This fringe, which was first noticed by Landolt, was explained by Lippich,[1] who showed that it was due to the fact that the directions of vibration in the different parts of the field were not strictly parallel, a circumstance resulting from the varying obliquity of the rays. Lippich showed further that in the case of polarizing prisms with end faces perpendicular to the prism axis the direction of vibration was represented by a system of converging lines which met in a point outside of the prism, as shown in Fig. 214*a*. Complete darkness will occur with crossed Nicols only in those parts of the field where the directions of vibration in the two prisms are perpendicular. If we draw two fields similar to the one represented by *a*, at right angles to each other, we shall find that the small areas formed by the intersection of the converging lines are in general diamond-shaped, but that they are approximately rectangular along a curved line represented by the dark band in Fig. 214*b*. This is the region where the directions of vibration are strictly perpendicular, and it in consequence appears black. A slight rotation of either field will obviously change the position of this locus, the squares becoming diamond-shaped, and the adjacent diamonds straightening out into squares. Rotation of one of the prisms through 180° will be found to give a locus of squares perpendicular to the one shown in the figure, and since the fringe moves broadside across the field, the directions of motion in the two cases are perpendicular.

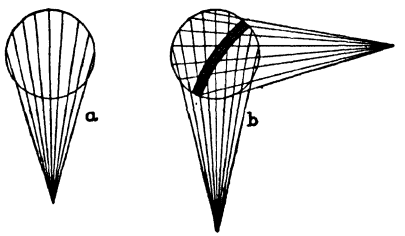


FIG. 214

Lippich made use of the fringe in the construction of a polarimeter, in which the position of the fringe was determined with reference to a pair of cross hairs. It was possible in this way to set the analyzing Nicol with an error not exceeding two or three sec-

[1] F. Lippich, *Wien Akad. Ber.*, III, Bd. lxxxv, page 268, 1882.

onds of arc. A very brilliant source of light must be used, such as the sun or electric arc.

In the case of the prisms of Glan, Thompson and Glazebrook, the directions of vibration in the field form a system of parallel lines and the Landolt fringe is not present.

Interference of Polarized Light. — The study of the interference of polarized light was taken up by Fresnel and Arago in 1816. Young's explanation of the colors of thin plates of doubly refracting substances in polarized light was not wholly satisfactory to Fresnel. The non-appearance of color in the absence of the polarizer and analyzer had not been accounted for, Young's explanation being simply that the color was produced by interference between the ordinary and extraordinary rays emerging from the thin plate. They performed an extensive series of experiments, polarizing the light passing through two slits (arranged so as to give interference fringes normally) in mutually perpendicular planes, and found no trace of interference. Employing thin plates of selenite they then arranged matters so that two rays, polarized at right angles obtained from initially polarized light could be brought into the same plane, in which case fringes were observed. If, however, the two polarized rays were obtained from ordinary light, no interference occurred. A full account of these experiments will be found in earlier editions of this book.

The Fresnel-Arago Laws. — Their results can be summed up in what are known as the Fresnel-Arago laws.

1. Two rays polarized in the same plane interfere in the same manner as ordinary light.
2. Two rays polarized at right angles do not interfere.
3. Two rays polarized at right angles (obtained from ordinary light), and brought into the same plane of polarization, do not interfere in the ordinary sense.
4. Two rays polarized at right angles (obtained from plane-polarized light) interfere when brought into the same plane of polarization.
5. In the latter case, under certain conditions, half a wavelength, corresponding to the phase-difference of 180°, must be added to the path-difference.

As we shall see in a later chapter two beams of light, polarized at right angles, are capable of uniting into a circular or elliptical vibration, so that we can consider interference as taking place in this sense. Some experiments illustrating this effect were described by the author [1] which we will now consider.

[1] *Phil. Mag.*, April, 1904.

On the Polarized Fringes Produced by the Interference of Two Streams of Light Polarized at Right Angles.[1] — In the case of ordinary thin-film interferences the planes of vibration of the disturbances reflected from the two surfaces of the film are parallel.

It is possible, however, to prepare a film which shall fulfil the requirement that the vibrations reflected from its upper surface make any desired angle with those coming from the lower surface. The path-difference between the two streams will vary with the thickness of the film; and if the amplitudes be equal we shall have the vibrations compounding into circular, elliptic or plane ones, according to their phase-difference.

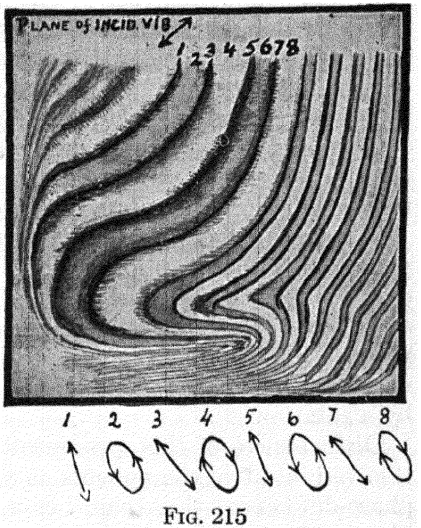


FIG. 215

A thin glass or gelatine film, backed by a metallic reflecting surface, is all that is necessary. The incident sodium light should be polarized at an angle of 45° with the plane of incidence by passage through a Nicol, and the reflected light examined with an analyzing Nicol. The fringes obtained in this way present a most curious appearance, reminding one forcibly of a spectrum line with a fainter component seen in the Fabry and Perot interferometer. Their general appearance is shown in Fig. 215, which represents the fringes obtained by flowing a plate of speculum metal with a rather dilute solution of gelatine, and allowing it to dry in a slightly inclined position.

The easiest way to get them, however, is to blow out the end of a rather large glass tube into a large thin balloon of tissue glass, picking out a portion, by the light of a sodium flame, which shows fairly straight interference fringes one or two millimetres apart. A small piece of the thin glass is laid, with its slightly convex side down, upon a clean mercury surface, and sodium light, polarized in azimuth 45°, reflected from the surface at an angle of about 60°. On viewing the reflected light through a Nicol, the curious double fringes can be easily found by slowly turning the Nicol. The light will be found to be plane-polarized along the lines 1, 3, 5, 7, though in general the planes of polarization along one set of lines

[1] Wood, *Phil. Mag.*, April, 1904.

is inclined to the plane of polarization along the alternate lines, as is indicated below the figure, the arrows representing the direction of the vibration (electric vector). Between the lines of plane-polarized light, which appear as dark fringes when the Nicol is so oriented as to extinguish the light, we have either elliptically or circularly polarized light, as can at once be shown by the introduction of a quarter-wave plate, which enables us to extinguish the light along the lines 2, 4, 6 and 8, by suitable adjustment of the mica plate and analyzer. The direction of revolution of the vibration along lines 2 and 6 is opposite to that along lines 4 and 8. The fringes, and their states of polarization were completely accounted for by examining separately the states of polarization of the light reflected from the glass and the glass-mercury surface, by laying a thick plate of glass on the metal and using a narrow beam of light, the two reflected beams being thus separated. A full explanation is given in earlier editions.

Elliptical and Circular Polarization. — In the case of plane polarization the direction of the electric vector is fixed in space, or adopting the nomenclature of the elastic solid theory, the vibration is along a line. We shall now take up two other types of polarization in which the path of the terminal point of the electric vector is an ellipse or circle, or the to-and-fro vibration is replaced by a rotation.

Disturbances of this type result whenever a plane vibration is resolved into two components at right angles to each other with a phase-difference other than 0 or π in which case the resultant is a plane vibration.

If the amplitudes are the same and the phase-difference an odd number of quarter periods, we shall have a circular vibration which is right- or left-handed according to the circumstances. This can be easily shown by means of the circular pendulum: suspend a weight by a string and strike the weight a blow in any direction. a linear vibration results. Strike a second blow, at right angles to the direction of the first, and a quarter period later, *i.e.* when the weight has reached its position of greatest displacement, and the linear vibration will be replaced by a circular one. If we delay the second blow until a half period has elapsed, the resultant motion will be linear, but in a direction making an angle of 45° with the original direction, while if we wait until three-quarters of a period have elapsed, we again get the circular vibration, but in an opposite direction.

We have cases similar to the above when plane-polarized light is transmitted through a thin crystalline plate which is doubly refracting. The incident vibration is in general resolved into two

components at right angles to each other which traverse the plate with different velocities, and consequently emerge with a phase-difference depending on the thickness of the plate. If the plate is very thick and the cross section of the ray small, the two components are completely separated and emerge plane-polarized, but in the case of very thin plates the components emerge without appreciable separation, and compound into a vibration which may be circular, elliptical or linear according to the path-difference within the plate, and the amplitudes of the two components. The circular vibration results only when the amplitudes are equal and the path-difference is $(2n+1)\lambda/4$. If plane-polarized monochromatic light is passed through a plate of some doubly refracting crystal cut parallel to the axis with its electric vector making an angle of 45° with the axis, it will emerge plane-polarized parallel to the original plane for certain thicknesses of the plate, and plane-polarized at a right angle to this plane for other thicknesses. For intermediate thicknesses it will be elliptically or circularly polarized.

The plates most commonly employed are of mica or selenite, both of which can be easily cleaved, or of quartz cut in the form of thin wedges or plates. Owing to dispersion the phase-difference of the emergent components depends not only on the thickness of the crystal plate but also on the wave-length of the light, consequently if white light is employed the state of polarization of the transmitted light varies as we pass along the spectrum, and if a plate of moderate thickness is used and the light passed through a Nicol and into a spectroscope, the spectrum will be found to be furrowed by dark bands which represent the wave-lengths quenched by the Nicol.

Polarizing Monochromator for the Separation of Close Spectrum Lines. — An instrument for separating and utilizing the light of close spectrum lines, based on the dispersion of the plane of polarization described above, was designed by the author in 1904 and used in the study of the dispersion of sodium vapor. Later on [1] it was improved and utilized for the excitation of the fluorescence of sodium vapor by a single D line, since which time it has been extensively employed in all cases in which the separation of very close lines is required.

If the light of a sodium flame or vacuum tube is polarized and allowed to fall on a plate of quartz 32 mms. thick, cut parallel to the optic axis, with the plane of polarization making an angle of 45° with the axis the emergent waves of D_1 and D_2 of sodium will be plane-polarized at right angles to each other, and either

[1] *Phil. Mag.*, March, 1914.

can be quenched by a Nicol suitably oriented. If white light is used, and analyzed by a spectroscope, the spectrum will be furrowed by dark bands, the distance between a bright and a dark band being, in the yellow region, 6 Ångström units, the distance between the D lines. As it is necessary to employ a large condenser and work with very divergent and convergent cones of light, a block of quartz of large size must be used, placed between the two halves of the condenser, since the rays which traverse the block must be parallel. If this is not the case, different pencils will traverse different thicknesses, and will be differently polarized. Moreover, one-half of the light is lost at the start by the polarizing Nicol. This difficulty was overcome by employing a large double-image prism, and subsequently analyzing by a double-image prism. In this way, with proper orientation of the prisms, the two images containing only D_2 light were superposed, the D_1 images (of one-half the intensity) lying to the right and left. By this expedient the D_2 image had the full intensity, except for the loss by reflection.

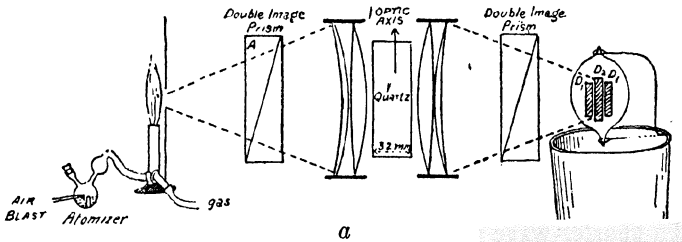
A rotation of 90° of the plane of polarization is produced by a quartz plate .032 mm. in thickness for sodium light, consequently the plate must be plane-parallel to within considerably less than this distance, otherwise D_1 will be passed by some parts of the field and D_2 by others. If the difference in thickness changes by .032 mm. in passing from one edge of the block to the other, one edge will transmit D_2 only, and will appear brighter than the other edge which transmits only D_1, while the centre of the plate will transmit both D_1 and D_2 in a state of circular polarization.

The extinguishing of one of the D lines can be shown with a natural uncut crystal of quartz, if the surfaces are fairly good. The crystal is to be placed between crossed Nicols, utilizing two opposed surfaces, which are separated by a distance of two or three centimetres.

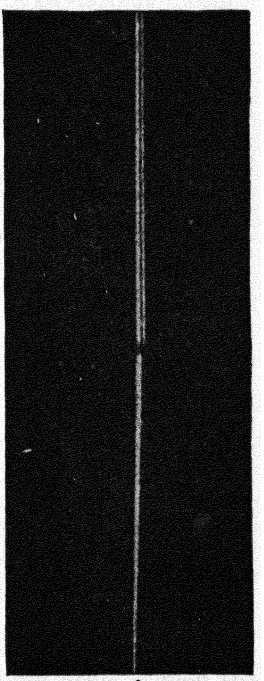
An image of the crystal is projected upon the slit of the spectroscope, and D_1 will be found absent at certain points, D_2 absent at others. It is another matter if a large amount of light is to be used, as in experiments upon fluorescence, for in this case we must use a large block of uniform thickness *free from all traces of crystalline irregularities.*

The arrangement of the apparatus is shown in Fig. 216*a* as arranged for the study of the resonance radiation of sodium vapor. A photograph of the sodium lines with and without the use of the instrument is shown in Fig. 216*b*. Only D_2 appears in the former. The double-image prism A is so oriented as to have the vibrations of the two transmitted beams at 45° with the vertical.

The second double-image prism is oriented *in the absence of the quartz block*, so as to give two images of the source side by side,



a



b

FIG. 216

and separated by a distance equal to the width of each image. Each image contains both D_1 and D_2. If now the quartz block is placed between the two halves of the condenser, all of the D_2 light (for example) leaves the original images, and unites into a third image between the other two. This is the image utilized. If it is desired to have D_1 light in this image, it is necessary only to rotate the block a degree or two, thereby increasing the optical path in quartz. If we rotate the analyzing double-image prism the central image splits up into two, and the correct position of the prism can be determined by bringing these two images into exact coincidence. As has been said previously, this method obviates the 50% loss of light which results from the use of Nicol's prisms. As the double-image prisms had apertures considerably less than that of the quartz block, they were placed midway between the condenser lenses and the image and source, as shown in the figure. With prisms of too small aperture, however, placed close to the source and image, the separation of the images would be insufficient.

Polarization of Emergent Light. — The state of polarization of the light which emerges from a crystal plate, cut parallel to the axis for two azimuths of the incident plane-polarized light and various thickness of the plate is shown in Fig. 217 in which x and y are the directions in which the electric vector must lie for propaga-

tion without resolution, a disturbance parallel to y being transmitted at the lower velocity. The initial disturbance A is resolved along x and y and for azimuth, $i=45°$, gives elliptical, circular, elliptical, plane, as shown, for thicknesses giving the path-differences indicated while for azimuth 20° we have only elliptical and plane vibrations.

Production and Properties of Circular Light. — The easiest method of producing circularly polarized light is by means of a plate of mica of such thickness that the path-difference between the ordinary and extraordinary rays is a quarter of a wave-length, the proper thickness for yellow light being .032 mm. Such a plate is called a **quarter-wave plate**, which we shall hereafter speak of as a $\lambda/4$ plate. These plates can be prepared without difficulty by splitting a good quality of mica by means of a needle into the thinnest possible sheets, and selecting such as completely restore the light when they are placed in the proper azimuth between a pair of crossed Nicols.

The thickness of the plates should be measured with a micrometer caliper or spherometer, as otherwise the mistake may be made of getting the plate three times too thick, the optical effects produced (with sodium light) by a $3\lambda/4$ plate being similar in appearance. The two directions on the plate parallel to vibrations which are propagated without change should be marked. They can be easily found by holding the plate between two crossed Nicols, in such an azimuth that the field appears dark. The directions in question will then be parallel and perpendicular to the longer diagonal of the field of the analyzing Nicol. It is also important to know which of the two directions corresponds to the greater retardation.

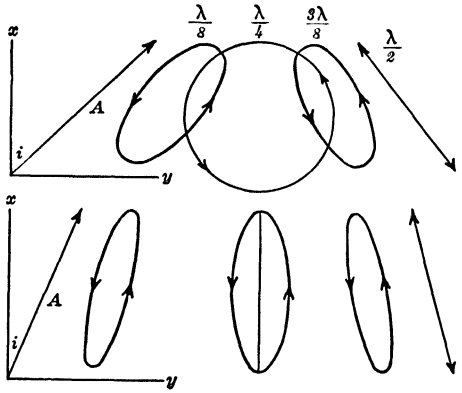


FIG. 217

If the plate is mounted with its principal directions vertical and horizontal respectively, in front of one of the silvered mirrors of a Michelson interferometer and the fringes found with white light, it is not difficult to determine the direction corresponding to the faster propagation. The central black fringe is brought upon the cross hair of the telescope in which the fringes are viewed, and the

light passed through a Nicol before it reaches the instrument. It will be found that a shift of ½ a fringe width occurs when the vibration plane is changed from horizontal to vertical. If this shift is in the same direction as the shift originally produced by the introduction of the mica plate, it means that the retardation has been increased by changing the direction of the vibration from horizontal to vertical, consequently the vertical direction in the plate is the direction in which the slower component vibrates. This direction should be marked "Slow," the other "Fast."

As this method involves some trouble, the following, based on observations made with a plate previously tested as above, will be found simpler:

A Nicol prism is mounted in front of a sodium flame with its short diagonal turned in the direction in which the hands of a clock move, through an angle of 45° from the vertical. The light polarized in azimuth 45° is then reflected from a polished metal surface, *e.g.* silver or speculum metal, at an angle of about 60°, which introduces a phase-difference between the components of nearly a quarter of a period (the component perpendicular to the plane of incidence being retarded). If the light is then passed through the quarter-wave plate and an analyzing Nicol, it will be found that it can be extinguished by the latter; that is, the quarter-wave plate reduces the nearly circular vibration into a plane vibration. The plane of this vibration, which is given by the *long diagonal* of the analyzing Nicol when set for complete darkness, makes an angle of 45° with the two directions of vibration which we have marked on the λ/4 plate; this direction is indicated by the dotted arrow in Fig. 218, the directions of vibration of the fast and slow disturbances being as shown. We have then merely to fit this diagram to our λ/4 plate, making the dotted arrow coincide with the direction of vibration of the plane-polarized emergent ray.

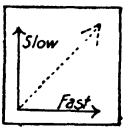


FIG. 218

For the present we will not concern ourselves with the explanation of why this method enables us to distinguish between the ordinary and extraordinary rays. The reason will become clear after the study of the direction of revolution of circularly polarized light. If plane-polarized light is passed through the λ/4 plate, with its plane of vibration making an angle of 45° with the two principal directions, it will be found to suffer very little change in intensity when examined with a slowly rotating analyzer. In this respect it resembles ordinary unpolarized light. It may be distinguished from the latter, however, by passing it through a second λ/4 plate, which, by bringing the retardation between the

components up to $\lambda/2$, converts it into plane-polarized light, which can be extinguished with a Nicol. It also shows brilliant colors in the designs made up of thin flakes of selenite, when the latter are viewed through a Nicol or other analyzer.

Circular Polarization by Total Reflection. Fresnel's Rhomb. — When light polarized in a plane making an angle of 45° with the plane of incidence is totally reflected at an angle of 54°, the two reflected components have a phase-difference of one-eighth of a period (for glass-air reflection). Two such reflections give the required $\lambda/4$ difference, and produce circular polarization. In the case of total internal reflection, the phase of the component of vibration parallel to the plane of incidence is retarded 135°, or a total retardation of 270° for two reflections. This is virtually the equivalent of an *acceleration* of 90°, and we can so consider it in all experimental work. (See Lord Kelvin's *Baltimore Lectures*, page 401.)

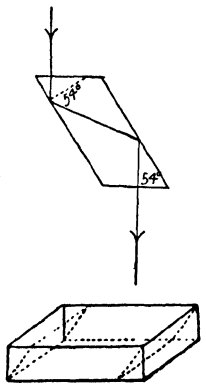


FIG. 219

This phenomenon will be more completely discussed in the article on the theory of reflection, and for the present we shall merely assume the fact to be true. Fresnel constructed a rhomb of glass to verify his calculations of the effect of total reflection upon plane-polarized light, and found that after two internal reflections at an angle of 54°, as shown in Fig. 219, the light emerged circularly polarized.

A rhomb of this description can be made easily out of a rectangular piece of thick plate glass, the dimensions of which should be in about the proportion 1:2:3. The plate glass employed should be as thick as possible. It is usually possible to get strips of glass an inch or two in width and an inch thick, which have been trimmed from large plates. These make excellent rhombs, though equally good results on a smaller scale can be obtained with pieces cut from quarter-inch plate. The ends of the block are to be ground down on a grindstone to an angle of 54°, as shown by the dotted lines in Fig. 219. With a quarter-inch plate this can be done in a short time, but if the very thick plate is employed it is better to saw off the ends with a mineralogist's saw, as the slow grinding is very tedious. Small pieces of thin plate glass, cemented to the rough ground ends of the rhomb with "boiled-down" Canada balsam, make an excellent substitute for polished faces, and save several hours of labor.

If the light entering one of the oblique faces of the rhomb is polarized in a plane making an angle of 45° with the plane of inci-

dence, the emergent light will be freely transmitted by a Nicol in every azimuth. If, however, thin mica or selenite films are interposed between the rhomb and the analyzer, they will show brilliant colors, which is not the case when ordinary light is used.

The Fresnel rhomb has an advantage over the $\lambda/4$ plate, for the phase-difference between the rectangular components is nearly independent of the wave-length, which is not so in the case of the mica film.

Production and Properties of Elliptical Light. — Elliptically polarized light can be produced in a number of ways: by the transmission of plane-polarized light through a quarter-wave plate, the plane of vibration making an angle of say 20° with the principal plane of the plate; by decreasing this angle the ellipse becomes more eccentric, degenerating into a line when $i=0$. By increasing i the ellipse becomes less eccentric, and passes through the circular condition when $i=45°$. If we use a plate of some other thickness, we obtain an ellipse with its axis inclined to the original direction of vibration.

Elliptically polarized light, when examined through a Nicol, shows fluctuations in brilliancy as the prism is rotated, the change in intensity becoming more marked as the eccentricity of the ellipse is increased. It thus resembles partially polarized light, but can be distinguished from it by introducing a retardation of a quarter of a period by means of a $\lambda/4$ plate, which converts it into plane-polarized light. The directions of the axes can be determined by the $\lambda/4$ plate, for they are parallel and perpendicular to the principal section of the plate when it is so oriented as to give plane-polarized light.

The ratio of the axes can be determined by observing the angle between the principal plane of the analyzing Nicol when it extinguishes the light, and the principal plane of the $\lambda/4$ plate. The tangent of this angle is the ratio of the axes of the ellipse, for when two rectangular vibrations compound into a linear vibration the tangent of the angle which the resultant makes with one of the components is the ratio of the components.

When the $\lambda/4$ plate and the Nicol are in such positions as to extinguish the light, we have the arrangement shown in Fig. 220, in which the elliptical disturbance A (with components a and b) approaches the observer, passing through the $\lambda/4$ plate B, which decomposes it into its components. On emerging, the resultant linear vibration c is extinguished by a Nicol oriented as shown, the tangent of the angle c giving us the ratio b/a. The constants of elliptical polarization cannot be very accurately determined

with the $\lambda/4$ plate, owing to the difficulty of making accurate settings of the mica plate and Nicol.

Babinet Compensator. — A better contrivance is Babinet's compensator, which has been adapted by Jamin to the study of elliptically polarized light. It consists of two acute prisms of quartz, which, when placed in contact, form a plate the thickness of which can be varied by moving one prism with a micrometer screw. The optic axes are parallel to the surfaces of the plate, but perpendicular to each other, as shown in Fig. 221. If plane-polarized light falls normally on the face of the compensator, the plane of vibration not coinciding with either of the principal planes, it will be broken up into two components parallel and perpendicular to the optic axis. When these vibrations enter the second prism, their directions will remain unaltered, but they

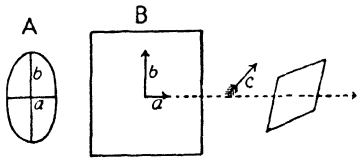


FIG. 220

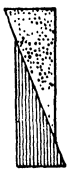


FIG. 221

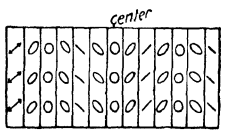


FIG. 222

will exchange velocities; *i.e.* the ordinary ray in one becomes the extraordinary ray in the other. If μ_ϵ and μ_o be the refractive indices for the two polarized disturbances, and if a ray traverses a thickness ϵ in one prism, the relative retardation of the two disturbances is $\epsilon(\mu_\epsilon - \mu_o)$, and for a thickness ϵ' in the second prism it is $-\epsilon'(\mu_\epsilon - \mu_o)$, for the disturbance which is the faster in the first prism is the slower in the second. The retardation produced by the plate as a whole at the point in question is obviously

$$(\epsilon - \epsilon')\ (\mu_\epsilon - \mu_o).$$

The retardation is zero for the central ray, for at this point $\epsilon = \epsilon'$, and the light emerges polarized in the original plane. On either side of this point we shall have points at which the relative retardation is π, 2π, 3π, etc.; the plane of polarization of the light emergent at the points at which the retardation is an even multiple of π is parallel to the original plane of the incident light. At intermediate points, where the retardation is an odd multiple of π, the transmitted light will be polarized in a plane inclined to the original plane by an amount $2a$, where a is the angle between the plane of the original vibration and the plane of vibration of the retarded component. There will thus be a system of lines across the face of the compensator along which the light is polarized in the original

plane, and another system midway between them where the light is polarized at angle $2a$ with the original plane.

If the incident vibration makes an angle of 45° with the principal planes of vibration of the compensator, the plane of vibration along this second set of lines will be at right angles to the vibration along the first system. At points between the lines the light will be elliptically or circularly polarized, the condition over the surface of the compensator being roughly represented in Fig. 222. If the surface of the compensator be viewed through a Nicol with its planes of vibration parallel to the planes of the linear vibrations along the surface, the light will be extinguished along one set of lines, and the field will appear traversed by equidistant dark bands.

The bands of circular polarization may be detected by bringing a quarter-wave plate between the compensator and the analyzing Nicol; this gives us plane polarization along the lines which were previously circularly polarized, and a new system of dark bands results. The mica plate should be so oriented as not to affect the appearance of the systems of plane-polarized lines.

The dark bands previously alluded to are of course most distinct when $a=45°$.

A new and superior modification which allows compensation over a large field instead of along a narrow strip, consists of two wedges with axes parallel instead of perpendicular combined with a plane-parallel plate of the same thickness as the wedge combination, but with its axis perpendicular to the axes of the wedges.

The two wedges are equivalent to a plane-parallel plate the thickness of which can be varied by a measurable amount, and this in combination with the plate makes it possible to have elliptical polarization of any axis ratio over the entire field.

Determination of the Constants of Elliptical Polarization. — When the elliptical polarization is produced by a quarter-wave plate we can calculate the position and ratio of the axes, but in cases in which the ellipticity is the result of reflection, it becomes necessary to determine the constants experimentally. These determinations are of importance in connection with the theory of reflection, as we shall see in a subsequent chapter.

The compensator in its original form was provided with a fine cross wire moved by a micrometer screw, by means of which the distance between the bands could be measured, and the displacement of the bands determined. As modified by Jamin for the study of elliptically polarized light, the instrument has a fixed cross wire, one of the quartz wedges being moved by the screw. The relative retardation $e-e'$ is increased or diminished at a given point ac-

cording to the direction of the motion of the wedge, consequently the dark bands are displaced by a corresponding amount.

The wedge must, however, be moved through double the distance moved by the wire in the old form of instrument, in order to displace the system by the width of a band, since in this case e varies while e' remains constant, while in the case of the movable wire both e and e' vary, the one increasing and the other diminishing; the difference between e and e' consequently increases twice as rapidly with a moving wire as it does with a fixed wire and moving wedge. Let $2a$ be the distance between two dark bands as measured by the wire, and $2b$ the distance through which the wedge is moved in order to produce the same shift — from the foregoing $b=2a$.

The retardation δ at distance x (measured by moving wedge) from the central band is $\delta=(x/b)(\lambda/2)$, since moving the wedge a distance b changes the retardation by $\lambda/2$. We have now the necessary data for the study of an elliptical vibration, and will first determine the phase-difference between the two components. The components into which the incident vibration is resolved at the quartz surface will differ in phase by an amount $\alpha-\beta$, if we represent them by

$$x=A\cos(\omega t+\alpha),\ y=B\cos(\omega t+\beta).$$

Transmission through the plate alters this phase-difference by an amount $\delta=(2\pi/\lambda)(e-e')(\mu_e-\mu_o)$, and there will be a system of lines along which the total phase-difference $\alpha-\beta+\delta$ will be multiplies of π and the transmitted light plane-polarized.

We first adjust the wedges so that with plane-polarized light to start with the central dark band is bisected by the cross wire. The phase-difference at this point is zero. Substituting elliptically polarized light we find the central band shifted to a point, so situated that the phase-difference between the components of the elliptical vibration is compensated exactly by δ, the phase-difference resulting from transmission through the plate. The quartz wedge is now to be moved by the micrometer screw until the central band is again bisected by the wire. If this distance is denoted by x, we have

$$\frac{\alpha-\beta}{\pi}=\frac{x}{b}\text{ or }\alpha-\beta=\pi\frac{x}{b},$$

λ having been previously determined.

Position of the Axes. — The phase-difference of the component vibrations along the axes is 90°. We set the compensator as before, so that with plane-polarized light the central band falls under the

wire, and then move the wedge a distance $\frac{1}{2}b$. There is now a phase-difference of 90° along the line under the wire. Substituting the elliptical light we rotate the compensator until the central band is again bisected by the wire. The axes of the elliptical vibration are now parallel to the axes of the quartz wedges.

Ratio of the Axes. — If the compensator is set so that its axes are parallel to the axes of the elliptical vibration, the tangent of the angle between one of its principal planes and the principal plane of the analyzer is the measure of the ratio of the axes. The compensator acts in this case in the same way as the $\lambda/4$ plate, the use of which in the determination of the ratio of the axes has already been given.

Elliptical Polarization by Reflection. — We have seen that when unpolarized light is twice internally reflected at an angle of 54°, it emerges as circularly polarized light if the original plane of polarization made an angle of 45° with the plane of incidence; each reflection in this case introduces a phase-difference of $\frac{1}{4}\pi$ between the reflected components, consequently a single internal reflection from glass will give us elliptically polarized light. This can be shown with an ordinary right-angle prism. In general, when plane-polarized light is reflected at an azimuth of 45°, *i.e.* with its plane of vibration inclined at 45° to the plane of incidence, the reflected light will be, to a greater or less extent, elliptically polarized. In the case of glass and other transparent media the eccentricity of the ellipse is very great; in other words, the reflected light is very nearly plane-polarized, but in the case of metals the elliptical polarization is very marked. If plane-polarized light is reflected from a silvered mirror it will be found to be quite freely transmitted by a Nicol prism in all positions, if the plane of polarization originally made an angle of 45° with the plane of incidence. These cases will be more fully discussed in the Chapter on the Theory of Reflection.

Direction of Revolution in Circularly Polarized Light. — The direction of revolution of the circular vibration depends on the thickness of the crystalline plate, its orientation, and its nature, *i.e.* whether it is a positive or negative crystal. The positions of the ordinary and extraordinary ray, and the directions of vibration in each, are shown for positive and negative crystals in Fig. 223. In the former the extraordinary component travels slower than, and consequently lags behind, the ordinary; in the latter the reverse is true. We will now take the case of a $\lambda/4$ plate of mica, which is a negative crystal, and determine the direction of revolution for two different orientations. First suppose the direction of vibration of the incident light to make an angle of $+45°$

with the principal section (optic axis in Fig. 224). It is decomposed into the components O and E, the former lagging behind the latter by one-quarter of a wave-length. The E component consequently carries the ether particle to the right, and when it is at its point of greatest displacement, the O component acts in a vertical direction, consequently the direction of rotation is from right to left, as shown by the arrow. If we now rotate the plane of vibration of the incident light through 180°, making the angle between

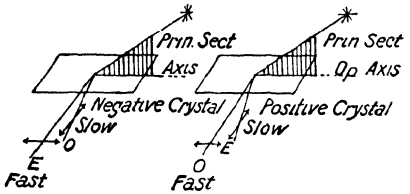


FIG. 223

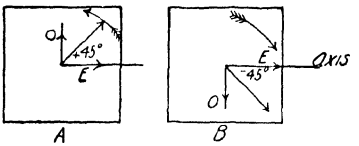


FIG. 224

it and the principal section −45°, we have the condition shown in B, and applying the same reasoning we find that the direction of rotation is now from left to right. In the case of positive crystals we apply the same construction, considering, however, that the O component acts first, since it is in advance of the other. The directions of revolution will be found to be the reverse of those in the former cases.

We can determine experimentally the direction of revolution with the quarter-wave plate. Suppose the light to be coming towards us, and the direction of revolution clockwise. It can be decomposed into two rectangular components A and B, B being a quarter of a period behind A. We will now suppose it transmitted through the $\lambda/4$ plate (placed as shown in Fig. 225) and examined with an analyzer. The component A, which is a quarter period ahead, will traverse the plate at the slower velocity and be brought into the same phase as the component B, the resultant plane vibration having the direction CD. If the direction of revolution be reversed, component A will be a quarter period behind B, and will experience a further quarter period relative retardation in traversing the $\lambda/4$ plate, the resultant having the direction EF. The direction of revolution is thus determined by observing whether the plane vibration makes an angle of + or −45° with the direction designated "fast."

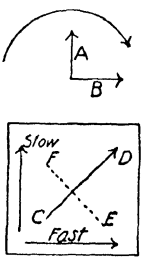


FIG. 225

If we have a Nicol prism and quarter-wave plate so oriented as to give us a right-handed circular vibration, by turning the Nicol

through 180° we reverse the direction of rotation. This can be readily understood by constructing two diagrams representing the two conditions.

Let us now return to the method which we use for determining the "fast" and "slow" directions in our $\lambda/4$ plate. With the arrangement of the Nicol and metallic reflector which we employed, we obtained a clockwise circular vibration. Obviously, if the direction of rotation is known, the fast and slow directions of the $\lambda/4$ plate can be determined by observing the direction in which the plane-polarized disturbance vibrates on leaving the plate.

Direction of Revolution in the Case of Fresnel's Rhomb. — As has been stated before in the case of total reflection, the component perpendicular to the plane of incidence virtually lags behind the other. If the rhomb is placed in a vertical position, and the incident light polarized in a plane turned clockwise 45° from the vertical, the direction of revolution will be clockwise. A convenient way of determining the direction of revolution when the rhomb is set for circular polarization is to notice the direction in which it must be turned in order to bring the plane of incidence into coincidence with the plane of vibration. This is the direction of revolution of the circularly polarized light.

Natural and Partially Polarized Light. — By natural light we mean ordinary unpolarized light, which is characterized by showing no change of intensity when passed through a $\lambda/4$ plate and Nicol prism, no matter how oriented, and by being doubly refracted by certain crystals, the intensities of the two refracted rays being independent of the orientation of the crystal. Partially polarized light, such as we obtain by reflection from a glass surface at some other angle than the polarizing angle, is characterized by showing fluctuations of intensity when it is examined through a rotating Nicol, never being completely extinguished, however. In this respect it resembles elliptically polarized light, but the difference between the two can be shown with the $\lambda/4$ plate, which converts the latter into plane-polarized light. Partially polarized light is doubly refracted by crystals, the relative intensities of the two rays varying with the orientation.

We have now to consider the simplest forms of vibration which are consistent with the above results, and we will begin with a discussion of natural or unpolarized light. Brewster explained natural light by assuming it to be made up of two plane-polarized disturbances perpendicular to each other and independently propagated. A disturbance of this nature, if it could exist, would undoubtedly have the properties of natural light, but there are of

course objections to the conception of a disturbance in which it is necessary to assume that the adjacent ether particles on the wave-front move in totally different directions. Fresnel accordingly, in 1821, advanced another hypothesis, namely, that natural light was in reality plane-polarized light, the azimuth of which changed with exceeding rapidity. Fresnel considered that a ray which came from a single centre of disturbance was plane-polarized, but that the azimuth varied rapidly. If we could isolate such a ray and experiment with it, we should find that it was alternately transmitted and cut off by a Nicol prism in a fixed azimuth. The isolation of a ray coming from a single centre of disturbance is, however, impossible, and even if it could be done there would be no way of verifying the hypothesis experimentally owing to the faintness of the light. Fresnel probably included the ellipse and circle in his conception of the motion of ordinary light, though he does not mention them specifically. The elliptical vibration is the most general form, and we may, on Fresnel's hypothesis, consider natural light as an elliptical vibration, the form and orientation of which changes with great rapidity, passing through the circle and straight line as special forms.

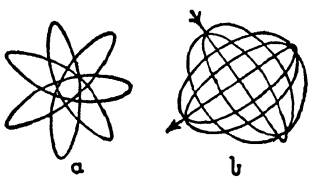


FIG. 226

The change in the orientation of the ellipse cannot be conceived as taking place suddenly. If it takes place gradually, the curve ceases to be an ellipse and takes the complicated form shown in Fig. 226*a*. If the ratio of the axes changes as well as the orientation, we have a curve of the form shown in Fig. 226*b*, except that the change from the straight line to the circle is much more gradual, and the whole curve must be conceived of as constantly changing its orientation.

Rotation of the Planes in Unpolarized Light. — A. Langsdorf and L. Du Bridge[1] have described an interesting experiment which may be regarded as showing the presence of the variously oriented polarized components of ordinary unpolarized light. They formed a set of interference fringes with a Fresnel bi-prism and introduced into the path of the two interfering streams of light two cells filled with a dextro- and laevorotatory liquid respectively, each giving a rotation of 45°. The interfering disturbances which were originally parallel were thus placed at right angles and gave rise to circular polarized light instead of darkness, and the fringes disappeared. The introduction of a quarter-wave plate of mica caused their reappearance.

[1] *J. O. S. A.*, 1933.

Obviously two quartz plates, one of right- and the other of left-handed quartz, cut perpendicular to the axis, and a little over 2 mm. thick (see Chapter on Rotatory Polarization) would be a more convenient way of performing the experiment.

Light-Beats. — Airy's conception that revolving plane-polarized light was merely the resultant of two circularly polarized disturbances, of different periods and opposite directions of rotation, was tested experimentally by Righi, who passed light from a slit through a rotating Nicol and after reflecting it from a pair of Fresnel mirrors, focussed it by a lens, thus forming two adjacent sources of revolving plane-polarized light. By means of two quarter-wave plates and a double image prism he arranged matters so that the interference fringes were in motion, passing a given point at the rate of $2n$ per second, n being the speed of rotation of the Nicol. The moving fringes could be regarded as "light-beats" of the two circular vibrations of opposite direction and altered frequency which were the equivalent of the rotating plane-polarized vibration. It should be observed, however, that a statical treatment can be given as well as a kinematical, the fringes occupying positions corresponding to the position of the Nicol. The moving fringes seen in an interferometer are just as much an evidence of light-beats, the change in wave-length in this case resulting from reflection from a moving mirror.

Relation between Amplitude and Intensity. — That the intensity of light is proportional to the square of the amplitude is by no means self-evident. The relation can be demonstrated in a number of ways, however. Taking a mechanical analogy we know that the instantaneous energy of a pendulum is proportional to the square of its velocity and its average energy to the square of its amplitude. We also know that two rays of light polarized at right angles to each other do not interfere, or rather that the resultant intensity is equal to the sum of the separate intensities, for interference of a certain type (circular or elliptical polarization) takes place. If a and b are the amplitudes, the resultant intensity will be $f(a)+f(b)$ in which f is the function to be determined. The phase-difference has no effect on the intensity, however, for it influences only the form of the vibration, circular, elliptical, or for the case of $\pm\pi/2$, plane with amplitude equal to $\sqrt{a^2+b^2}$. In this case the intensity will be $f(\sqrt{a^2+b^2})=f(a)+f(b)$ since we know from experiment that the intensities are additive. This equation holds only if f represents the square.

A third proof, free from any special theory, is based on the law of Malus. If the intensity of the light reflected from two parallel glass mirrors at the polarizing angle (as in Nörrenberg's apparatus)

is I_0 and one mirror is turned through angle Φ in azimuth (*i.e.* on the vertical axis), the reflected intensity becomes $I = I_0 \cos^2 \Phi$. The electric vector from the second mirror, of amplitude A is, by the rotation, resolved into two components $a \cos \Phi$ and $A \sin \Phi$, of which the first only is reflected and if its square is equated to $I_0 \cos^2 \Phi$ we see at once that $A^2 = I_0$ or that the original intensity is equal to the square of the amplitude.

Intensity of Circular and Elliptical Light. — The transformation of plane-polarized light into elliptical or circular light, as with a quarter-wave plate produces no change of intensity. This is not only an experimental fact, but also almost self-evident, as all of the energy must be transmitted (except the small loss by reflection) if there is no way of showing a loss of energy in the transmitted beam, and this could occur only if the reflecting power of the plate depended on its orientation, a supposition for which no justification exists.

If the intensity of plane-polarized light is represented by the square of the amplitude, and we resolve this amplitude A into two amplitudes a and b parallel to the principal directions of the mica plate, a and b are the semi-major and minor axes of the elliptical vibration, or the radius of the circular vibration, and since $a^2 + b^2 = A^2$ we see that the intensity of the circular light is represented by twice the square of the radius of the circle, or that of the elliptical light by the sum of the squares of the semi-major and minor axes. When therefore we add two rectangular disturbances together to produce circular light we get double illumination, exactly as when we add the effects of two independent sources of light.

CHAPTER X

DOUBLE REFRACTION

We have seen in the Chapter on Polarization that crystals of Iceland spar have the property of dividing a ray of light into two plane-polarized rays, the directions of polarization being at right angles to each other. In the present chapter we shall study in detail the propagation of light in crystalline media, and the laws which govern it.

The division of a ray of light by a crystal of Iceland spar, or double refraction, was observed in 1669 by Erasmus Bartholinus, and the polarization of the two rays was subsequently discovered by Huygens, though he was unable to explain the phenomenon, since at the time light was supposed to consist of waves in which the displacement was parallel to the direction of propagation. It was not until Young and Fresnel introduced the idea of transverse waves that the true nature of polarization was understood. Double refraction can be easily observed by laying a crystal of Iceland spar over an ink dot on a sheet of paper. Two images are seen which can be quenched in succession by the rotation of a pile of glass plates held at the polarizing angle. On revolving the crystal of spar, one of the images is seen to remain stationary, while the other revolves around it. The distance between the two images is independent of the position of the eye, showing that the rays of light, after refraction through the crystal, emerge parallel to one another, as shown in Fig. 227. At first sight it may appear strange that rays of light, incident normally upon the refracting surface, should be deviated away from the normal, since we are sometimes accustomed to associate the bending due to refraction with oblique incidence, the light-waves turning through an angle as they enter the denser medium. We shall see presently, however, that the phenomenon is easily explained by Huygens's principle, when applied to the peculiar type of waves which we have in doubly refracting media.

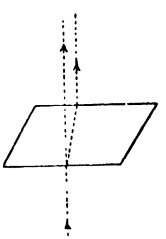


FIG. 227

Experiments have shown that in crystals belonging to the hexagonal and tetragonal system one of the rays obeys the ordinary laws of refraction, *i.e.* the refracted ray lies in the plane of incidence,

and the sine of the angle of incidence bears a constant ratio to the sine of the angle of refraction. This ray is called the ordinary ray; the other ray in general conforms to neither of these two laws, though in certain cases it may conform to one or to both of them. This ray is called the extraordinary ray. In the case of all other crystals except those of the cubic system, neither of the two refracted rays conforms of necessity to the ordinary laws of refraction. Crystals belonging to the cubic system do not exhibit the phenemenon of double refraction, the light being propagated as in isotropic media.

Physical Explanation of Double Refraction. — On the elastic solid theory we can explain double refraction in crystalline media by assuming that there are three directions called axes of elasticity, which have the distinctive properties of the two planes of vibration of Blackburn's pendulum, which consists of a weight suspended by strings as shown in Fig. 228*a*. If the weight is displaced either in, or perpendicular to, the plane of the paper it will oscillate in a straight line, the period being greater for vibrations perpendicular to this plane than for those parallel to it. If, however, it is displaced in an oblique direction, the force acting upon it will no longer be directed towards the position of equilibrium, and the weight will move in a curved orbit. In the case of crystals a particle displaced parallel to any one of the axes of elasticity will be acted upon by a force directed towards the equilibrium position, and the vibration will be plane-polarized. If displaced in any other direction and released, it does not return to its original position, but moves in a curved path in a manner analogous to that of the pendulum. We require an explanation of the splitting of a beam of light into two polarized components, and for their unequal velocities of propagation.

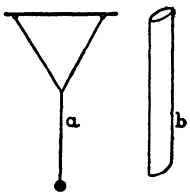


FIG. 228

The vibrations of a thin wooden rod clamped at one end in a vice form a useful analogy. If the rod has a circular cross section, and the free end is struck, the rod will vibrate in a plane with a frequency independent of the direction of the blow. This vibration can be regarded as a stationary wave with a node at the clamp, the frequency giving us the measure of the velocity of propagation. If however the rod has an elliptical cross section, it will execute a plane vibration only when struck in a direction parallel to one or the other of the two axes of the ellipse, the frequency being different in the two cases. If struck in a direction oblique to an axis, the force of restitution will not be directed towards the centre, and the end of the rod will execute a Lissajou figure, which shows the pres-

ence of two frequencies, and consequently two velocities of propagation.

We have in doubly refracting media a somewhat similar condition, the elasticity being different in different directions. Luminous vibrations will be decomposed into two polarized components which traverse the crystal with different velocities.

Wave-Surface in Uniaxial Crystals. — The fact was established by Huygens that, in isotropic media, the form of the wave-surface was spherical, and as one of the rays in Iceland spar was found to obey the ordinary laws of refraction, he assumed that the corresponding wave was a sphere. In the case of the extraordinary ray, which does not obey such simple laws, he made the assumption that the wave-surface was a spheroid, *i.e.* an ellipsoid of revolution. The velocity of the extraordinary ray in any direction is therefore given by the following construction: "Let an ellipsoid of revolution be described around the optic axis, having its centre at the point of incidence, and let the greater axis of the generating ellipse be to the lesser, in the ratio of the greater to the lesser index of refraction. Then the velocity of any ray will be represented by the radius vector of the ellipsoid which coincides with it in direction."

The optic axis may be defined as the direction in the crystal in which a ray of light may be propagated without double refraction. The law just given was found to apply to Iceland spar and many other crystals, but in all of these there was but a single optic axis. Brewster, however, discovered that in many crystals there were two directions in which light could be propagated without double refraction. Such crystals are termed biaxial, and the law of Huygens was found not to apply in these cases. Fresnel then established a theory which not only conformed to all of the known facts, but made possible predictions which were afterwards verified by experiment. This theory we shall take up a little later.

According to the theory of Huygens the wave-surface in uniaxial crystals consists of two sheets, one a sphere, the other a spheroid, which touch each other at two points. The direction of the line joining these points of contact is called the optic axis of the crystal. This conception applies, however, only to uniaxial crystals. In the case of Iceland spar and all so-called negative crystals, the sphere lies within the spheroid. In such crystals the angle of refraction of the extraordinary is greater than that of the ordinary ray. In the case of quartz and other positive crystals, the spheroid lies within the sphere, and the angle of refraction of the extraordinary ray is less than that of the ordinary. This will be clearer when we come to the construction of the refracted ray.

Huygens's Construction. — Suppose a luminous disturbance to start within a uniaxial crystal. The wave will spread out in two sheets, a sphere and a spheroid, which touch each other at two points. In the direction of the line joining these two points both waves travel with the same velocity. If we apply Huygens's construction to crystalline refraction, giving to the secondary wavelets, which originate on the refracting surface, the forms of spheres and ellipsoids, we can account for, and calculate the position of the two refracted rays. In all other directions the velocities will be unequal and we shall have a division of the ray, as may be seen by Huygens's construction. Consider a wave-front AB incident in an oblique direction upon the surface of a uniaxial crystal (Fig. 229). The direction of the optic axis is represented by the dotted line. The point A becomes the centre of two secondary

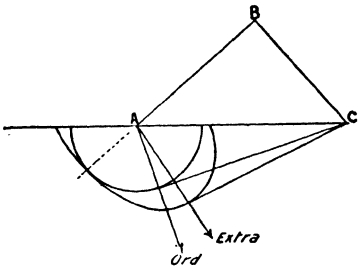


FIG. 229

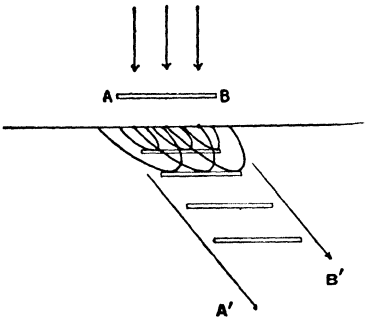


FIG. 230

wavelets which are propagated with different velocities. Making use of the same construction which we applied in the case of isotropic media (see Chapter III), we draw tangent planes from the point C to the two wave-surfaces; the directions of the refracted rays are given by joining the point A with the points of tangency. In the case of the ordinary wave the refracted ray lies in the plane of incidence. This is also the case with the extraordinary ray, *provided the optic axis lies in the plane of the paper.* If, however, the optic axis is not in the plane of the paper, the point of tangency for the extraordinary wave will lie above or below the plane of the paper, and the refracted ray will no longer be in the plane of incidence. In the latter case neither of the ordinary laws of refraction is obeyed, for the sine relation only holds when the section of the secondary wave is circular. If the optic axis is perpendicular to the plane of incidence, the section of the spheroid is equatorial and therefore circular, the extraordinary refracted ray in this case lying in the plane of incidence and obeying the sine law. The ratio of

the sines of the angles of incidence and refraction in this case is termed the extraordinary index of refraction.

We will next consider the case in which rays of light are incident in a normal direction upon the crystal. As we have seen, double refraction occurs in this case, one of the rays passing straight through, while the other is deflected away from the normal. It is obvious that we cannot apply in this case the simple explanation of refraction which assumes successive portions of the wave-front retarded upon entrance into the refracting medium. For the wave-fronts originally parallel to the surface must remain so after refraction. What we have actually, if our original waves be plane, are two plane-waves travelling through the crystal with unequal velocities but parallel always to the surface. The deflection of the extraordinary ray is obvious if we apply Huygens's construction to the present case. Assume that the points on the surface of the crystal become simultaneously the centres of ellipsoidal wavelets as indicated in Fig. 230. If the incident wave-front is limited to the region AB, the refracted wave-front will be the tangent plane of the ellipsoidal wavelets, and the refracted rays will be the lines AA', BB'. What actually happens may be described as follows: The refracted wave-front travels in the medium in a direction normal to its surface, but any limited portion of it bears away constantly to one side, and the ray is defined as the direction in which a limited portion of the wave-front travels. We see that, in this case, the ray is not perpendicular to the wave-front, which is, in general, true in doubly refracting media.

Verification of Huygens's Construction. — The assumptions made by Huygens regarding the form of the wave-surfaces in uniaxial crystals were speedily verified by experiment. That the ordinary wave-front is a sphere, was shown by constructing a prism formed of pieces cut in all possible directions from a crystal of Iceland spar and cemented together. The spectra formed by the extraordinary rays were deviated by different amounts, whereas a single spectrum only was formed by the ordinary rays which traversed the different elements of the prism. To verify the construction of the extraordinary wave-front we will consider several cases.

(1) **The refracting face is parallel to the optic axis, and the plane of incidence perpendicular to it** (see Fig. 231). — In this case the axis is perpendicular to the plane of the paper. The sections of the two wave-surfaces will in this case be circles, as we have seen. The tangent planes touch the sphere and spheroid at C and C'. Let the velocity in air be 1, then the velocity of the ordinary and the extraordinary rays will be proportional to b and a

the radii, and the refractive index of the extraordinary ray will be

$$\frac{\sin i}{\sin r}=\frac{1}{a}=\mu_e.$$

By cutting a prism of Iceland spar with its refracting edge parallel to the optic axis we obtain two spectra, and by measuring the deviation with a spectrometer we can calculate in the usual manner the refractive indices μ_o and μ_e for the ordinary and extraordinary rays. It can easily be shown that both rays are propagated through the prism according to the same law which holds in the case of a glass prism. This indicates that the section of the wave-surface is a circle for both rays, the radius for the extraordinary ray being $1/\mu_e$ and for the ordinary $1/\mu_o$. The extraordinary wave is therefore

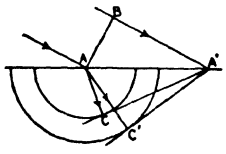


FIG. 231

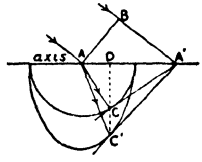


FIG. 232

a surface of revolution around the optic axis, and to determine the form of the generating curve we shall consider the refraction which takes place under different conditions.

(2) **Optic axis parallel to the surface of the crystal and to the plane of incidence.** — The sections of the wave-surfaces in this case are shown in Fig. 232. *Assume* the extraordinary wave-section to be an ellipse, the minor axis of which lies in the surface. The section of the sphere will be a circle touching the ellipse at the extremities of the minor axis. Drawing tangent planes from A', as before, to the two wave-surfaces, and joining the points of tangency with A, we obtain the refracted rays. A line joining the two points of contact and produced will cut the minor axis at a right angle, since the polar of any point in the chord of contact of a circle and ellipse having double contact is the same with regard to both curves. We have then

$$\frac{\tan r}{\tan r'}=\frac{AD}{CD}\times\frac{C'D}{AD}=\frac{C'D}{CD}=\frac{a}{b}=\frac{\mu_o}{\mu_e},$$

or the ratio of the tangents of the angles of refraction is equal to the ratio of the two indices of refraction.

This relation, which was deduced on the assumption that the wave-section was an ellipse, was verified by Malus in the following manner: Two scales AC and BC (Fig. 233) were engraved on a

plate of polished steel, and a thick plate of crystal with its faces parallel to the optic axis was laid on the scale and viewed through a telescope mounted on a graduated vertical circle. The crystal was brought into the horizontal position by means of levelling screws, the correct position being that in which the image of a distant point of light was not changed by rotation of the platform. Two images of the scale were seen in the telescope, and if we denote these by AC, $A'C'$, BC, $B'C'$, there will be some point of BC coinciding with some point of $A'C'$. We call this point h. If the axis of the telescope is directed towards this point it will cut the surface of the crystal at h, the position of which can be determined with reference to the scales. The divisions at E and D which appear to coincide can be read off and the distance ED determined by actual measurement. If e is the thickness of the crystal, we have

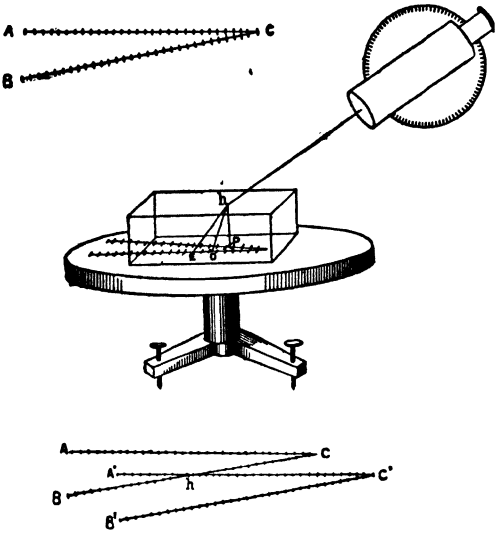


FIG. 233

$$ED = EP - DP = e(\tan r' - \tan r),$$

in which $\tan r$ is known, for the angle of incidence (considering the ray reversed) is equal to the angle which the axis of the telescope makes with the vertical. Moreover, $\sin i = \mu_o \sin r$ (since the ordinary ray obeys both laws of refraction for all conditions), therefore r is known, and r' may be determined by the above formula. If the value of r' thus found agrees with the value determined by the formula $\tan r'/\tan r = \mu_e/\mu_o$, the experiment will have proved that the section of the extraordinary wave is, in the present case, an ellipse. Inasmuch as we have already proven that the wave-front is a surface of revolution, the experiment will prove that it is a spheroid of axes A and B. The experiments made by Malus completely verified this theory.

The method employed by Malus was not very accurate. A method based upon prismatic refraction was used by Stokes,

which gave results correct to the fourth place of the decimal, verifying Huygens's construction, and completely disproving the law resulting from the theory that the double refraction results from a difference of inertia in different directions.

Fresnel's Theory of Double Refraction. — In the foregoing discussion we have considered only uniaxial crystals, making certain assumptions regarding the form of the wave-surfaces, and showing that certain relations deduced from them were verified by experiment.

We will now consider the phenomenon of double refraction in its more general aspect, following the treatment of Fresnel.

As we have seen, the velocity of a transverse wave in an isotropic medium is proportional to $\sqrt{e/d}$, in which e is the elasticity of the medium.

In doubly refracting media e is assumed to vary with the direction of the displacement, and there will be two directions in every possible plane for which e has its maximum and minimum values The corresponding velocities of propagations $\sqrt{e_1/d}$, $\sqrt{e_2/d}$ are for vibrations parallel to these two directions. If the displacement is in any other direction, the wave is not propagated with an intermediate velocity, as might at first be supposed, but is decomposed into two waves, which travel with the above velocities, the directions of their vibrations being perpendicular to each other. If we are dealing with trains of waves, as is always the case, the actual motions of the vibrating particles will not be along straight lines, for they are the resultants of the two sets of disturbances which are travelling with different velocities. Until the rays become completely separated by the double refraction, we must regard the vibration as changing its type from point to point, changing from plane to elliptical and circular, and then back again to plane, as the relative phases of the two perpendicularly polarized disturbances alter.

If the direction of displacement coincides with one or the other of the two directions of maximum or minimum elasticity, a single plane-polarized wave will be propagated in the medium. From this it is clear that in the case of the changing type of vibration alluded to above, the vibration along a line will never occur in either one of these directions, for if it did, it would be propagated from that point on, as a plane-polarized vibration without further change.

Fresnel arrived at a conception of the wave-surface by considering it as the envelope of an infinite number of plane-waves, which have passed simultaneously in all possible directions, through a given point in the doubly refracting medium.

Consider now the following construction. Through the point in

question imagine an infinite number of planes, in all possible orientations, and draw through the point, on each plane, two lines at right angles to each other, and coinciding with the directions of maximum and minimum elasticity, and of lengths proportional to velocities of propagation of disturbances vibrating parallel to the lines in question.

If the two lines are made to bisect each other at the point, the terminal points of the lines for all the planes taken collectively will lie upon an ellipsoid. This fact can be deduced theoretically, by making certain specifications regarding the medium, but as the deduction will not help us much in understanding the phenomena, we will simply consider it as representing experimental facts. Having the ellipsoid given, it is possible to find the direction of vibration and the velocities of propagation of a plane-wave, by drawing a central section of the ellipsoid parallel to the plane wave-front.

The ellipsoid is called the ELLIPSOID OF ELASTICITY.

Let its equation be $a^2x^2+b^2y^2+c^2z^2=V^2$, in which V is the velocity of light in vacuo.

The constants, a, b and c, are related to the elastic properties of the medium, and represent the velocities of waves vibrating parallel to the axes of elasticity, which may be defined as the three directions at any point, along which we can displace the ether, and have the force of restitution parallel to the displacement. In any given *plane* there are but two such directions; in space, however, there are three.

If we take as our unit of time the time occupied by a wave in travelling unit distance in vacuo, then $V=1$. If we put $x=0$ in our equation, we obtain the equation of the intersection of the ellipsoid with the yz plane, which is an ellipse having $1/b$ and $1/c$ as semi-axes, and a plane-polarized wave will be propagated along the x axis with a velocity b if the direction of vibration is parallel to y, or with a velocity c if it is parallel to z.

The reciprocals $1/a$, $1/b$, $1/c$ correspond to refractive indices, and are called the principal refractive indices. If we designate them by μ_1, μ_2, μ_3, we can write the equation of the ellipsoid in the form

$$\frac{x^2}{{\mu_1}^2}+\frac{y^2}{{\mu_2}^2}+\frac{z^2}{{\mu_3}^2}=1.$$

The deduction of the equation of the ellipsoid from a consideration of the elastic properties of the medium is generally accomplished by considering the potential of the medium. The following simple method is taken from Schuster's *Optics:*

"Fresnel's method of treating double refraction which led him to

the discovery of the laws of wave-propagation in crystalline media, though not free from objection, is very instructive, and deserves consideration as presenting in a simple manner some of the essential features of a more complete investigation. Consider a particle P attracted to a centre O with a force a^2x when the particle lies along OX, and a force b^2y when it lies along OY. The time of oscillation, if the particle has unit mass, is $2\pi/a$ or $2\pi/b$ according as the oscillation takes place along the axis of X or along the axis of Y. When the displacement has components both along OX and OY, the components of the force are a^2x and b^2y, and the resultant force is

$$R=\sqrt{a^4x^2+b^4y^2}.$$

"The cosines of the angles which the resultant makes with the coördinate axes are a^2x/R and b^2y/R. The direction of the resultant force is not the same as that of the displacement, the direction cosines of which are x/r and y/r. The cosine of the angle included between the radius vector and the force is found in the usual way to be

$$\frac{a^2x^2+b^2y^2}{Rr},$$

and the component of the force along the radius vector is

$$\frac{(a^2x^2+b^2y^2)}{r}.$$

"If we draw an ellipse $a^2x^2+b^2y^2=k^2$ (Fig. 234), where k is a constant having the dimensions of a velocity, the normal to this ellipse at a point P, having coördinates x and y, forms angles with

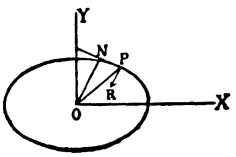


FIG. 234

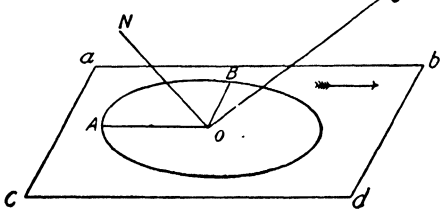


FIG. 235

the axes, the cosines of which are in the ratio a^2x to b^2y, hence the force in the above problem acts in the direction of ON of the line drawn from O at right angles to the tangent at P. The component of the force along the radius vector is k^2/r, and the force per unit distance is k^2/r^2, so that if the particle were constrained to move

on the radius vector OP, its period would be $2\pi r/k$. Since the ratio r/k depends only on the direction of OP, our result is independent of the particular value we attach to k.

"If we extend our investigation to three dimensions, the component of attraction along OZ being c^2z, we obtain the same result, and the component of force acting along any radius vector OP per unit length is k^2/r^2, where r is the radius drawn in the direction of OP to the ellipsoid

$$a^2x^2+b^2y^2+c^2z^2=k^2."$$

For a plane-wave to be propagated without alteration it is essential that the effective force of restitution shall be parallel to the displacement.

Though in general this force does not even lie in the plane of the wave-front, we can always resolve it, however, into two components, one in, the other perpendicular to, the front. Fresnel neglected the latter component, as it contributes nothing towards the propagation of a transverse wave. The longitudinal disturbance which, in the case of elastic solids, is produced by the normal component, is considered as non-existent in the case of light, owing to the incompressibility of the medium.

The direction of the component of force parallel to the wave-front is along the radius vector of the ellipsoid which is perpendicular to the section conjugate to the direction of the displacement. This will be made clearer by reference to Fig. 235. Let $abcd$ be a plane-wave travelling within the crystal, the direction of the displacement being parallel to ab. The ellipsoid is assumed constructed around a point lying on the wave-front, which cuts it in the elliptical cross section as indicated. The displacement is along AO, which we will assume to be the semi-major axis of the ellipse, while the direction of the force of restitution is along the radius ON, perpendicular to the plane BOC. If the projection of ON on the plane of the wave-front coincides with the direction of the displacement OA, the plane AON must be perpendicular to the wave-front, and since ON is perpendicular to OB, OB must be perpendicular to OA; in other words, OA and OB are the axes of the elliptical section. This is the condition which we assumed at the start. If the direction of the displacement is not along one of the axes, the effective force of restitution will not be directed parallel to the displacement, and *two* plane-polarized waves will result as we have seen. Two sections of the ellipsoid will be circular, and plane-waves parallel to these sections will be propagated without alteration, whatever may be the direction of the displacement, though there may be a division of the ray, as we shall see

presently. These circular sections of the ellipsoid of elasticity are perpendicular to the optic axes of the crystal. We may sum up as follows:

In any given direction in the crystal two systems of plane-waves can be propagated normally, the vibrations being along the axes of the elliptical cross section, and the velocities of normal propagation inversely proportional to the lengths of the axes. Two directions exist, however, in which but a single wave-front is propagated, known as the axes of single wave-velocity or optic axes. In these directions the velocity of normal propagation of a plane-wave is independent of the direction of vibration, although the direction in which a limited portion of the wave-front travels (the ray direction) depends upon the nature of the vibration, for the ray is not necessarily perpendicular to the wave-front in crystalline media.

We will now investigate the form of the wave-surface, which we can do by considering a geometrical construction known as the normal velocity surface.

The Normal Velocity Surface. — Around any point O within a crystal construct the ellipsoid of elasticity, and consider a system of plane-waves passing simultaneously through O in all possible directions. We have seen that, in general, a crystal has the property of transmitting only vibrations polarized in a definite direction, and that all other types of vibrations are resolved into two components which travel with unequal velocities. We shall thus have two systems of plane-waves passing through the point. To determine the velocities of these waves in different directions we proceed as follows. Let any one of the plane-waves passing through O cut the ellipsoid in the section AOB (Fig. 236), of which the axes are OA and OB. Draw a normal to the plane at O and measure off on it distances ON and ON', inversely proportional to the axes OA and OB. If now planes are drawn through N and N' parallel to the original plane of the section, they will represent the positions of the two waves which passed through the point O simultaneously, the one having its vibrations parallel to OA and the other parallel to OB. If we rotate the plane AOB around O in every possible direction, the points N and N', as defined above,

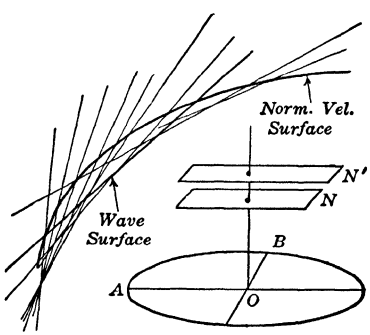


FIG. 236

will trace out a surface consisting of two sheets termed the surface of normal velocities, any radius vector of which determines the normal velocity of the plane-wave propagated in that direction. Since for two positions of the plane AOB the section of the ellipsoid is circular, it is obvious that the points N and N' will coincide, when the waves are parallel to these sections. In other words, the inner sheet will touch the outer at four points.

This surface is not, however, identical with the wave-surface, which is the surface enveloped by the plane-waves which we have just considered. This family of planes is represented by the equation

$$lx+my+nz=v$$

in which l, m, n are the direction cosines of the direction in which the wave travels with a velocity v, which is, however, a function of l, m and n. We require a relation connecting these quantities. The following treatment is taken from Rayleigh's *Wave-Theory:*

"If v be the velocity of propagation in the direction l, m, n, the wave-surface is the envelope of planes $lx+my+nz=v$, where v is a function of l, m, n, the form of which is to be determined. If $(\lambda\mu\nu)$ be the corresponding direction of the vibration, then

$$l\lambda+m\mu+n\nu=0."$$

According to the principles laid down by Fresnel, we see at once that the force of restitution $(a^2\lambda, b^2\mu, c^2\nu)$ corresponding to a displacement unity is equivalent to a force v^2 along $(\lambda\mu\nu)$ together with some force (P) along (lmn).

Resolving parallel to the coördinate axes, we get

$$lP=a^2\lambda-v^2\lambda, \quad mP=b^2\mu-v^2\mu, \quad nP=c^2\nu-v^2\nu,$$

or

$$\lambda=\frac{lP}{a^2-v^2}, \qquad \mu=\frac{mP}{b^2-v^2}, \qquad \nu=\frac{nP}{c^2-v^2}.$$

Multiplying these by l, m, n respectively, and remembering the relation $l\lambda+m\mu+n\nu=0$, we obtain

$$\frac{l^2}{a^2-v^2}+\frac{m^2}{b^2-v^2}+\frac{n^2}{c^2-v^2}=0,$$

an equation which we shall use presently.

The Wave-Surface. — If for every possible position of the section AOB in the construction which we have just considered, we construct planes through N and N' parallel to the section, these planes will envelop a surface which consists of two sheets, and resembles in its general appearance the normal velocity surface which we have just considered.

The enveloping surface represents the wave-front which we should have if a luminous disturbance started at O and spread out within the crystal. The wave-surface is the surface enveloped by the plane-waves just considered, and this surface will lie within the normal velocity surface, as shown in Fig. 236 above, which was constructed by drawing lines perpendicular to the lines joining the centre of the ellipse with various points on the curve.

The equation which represents the system of plane-waves which envelop the wave-surface is

$$lx+my+nz=v,$$

in which l, m, n and v are subject to the conditions

$$\frac{l^2}{v^2-a^2}+\frac{m^2}{v^2-b^2}+\frac{n^2}{v^2-c^2}=0 \text{ and } l^2+m^2+n^2=1.$$

The equation of the wave-surface was found by Archibald Smith [1] to be

$$r^2(a^2x^2+b^2y^2+c^2z^2)-a^2(b^2+c^2)x^2-b^2(c^2+a^2)y^2 \\ -c^2(a^2+b^2)z^2+a^2b^2c^2=0,$$

from which can be determined the general form of the wave-surface, by studying its sections with the planes xy, xz, yz. This we can do by making $x=0$, $y=0$, $z=0$ in succession in the equation of the wave-surface, when we obtain the equations of the curves of section. Assume $a>b>c$.

If we make $z=0$, we get at once

$$(x^2+y^2)(a^2x^2+b^2y^2)-a^2(b^2+c^2)x^2-b^2(c^2+a^2)y^2+a^2b^2c^2=0,$$

or

$$(x^2+y^2-c^2)(a^2x^2+b^2y^2-a^2b^2)=0,$$

which is separately satisfied by

$x^2+y^2=c^2$, a circle of radius c,

and $a^2x^2+b^2y^2=a^2b^2$, an ellipse of semi-axes a and b.

The circle lies wholly within the ellipse, since we have assumed c less than either a or b. Making $x=0$, we find the section with the yz plane to be

$y^2+z^2=a^2$, a circle of radius a,

and $b^2y^2+c^2z^2=b^2c^2$, an ellipse of semi-axes b and c.

In this case the ellipse lies within the circle.

For $y=0$, the section with the xz plane,

$x^2+z^2=b^2$, a circle of radius b,

$a^2x^2+c^2z^2=a^2c^2$, an ellipse of semi-axes a and c.

[1] *Phil. Mag.*, p. 335, 1838.

In this case the circle meets the ellipse at four points. The three sections are shown in Fig. 237.

A model of the surface can be made by cutting the sections out of cardboard and fitting them together in three perpendicular planes,

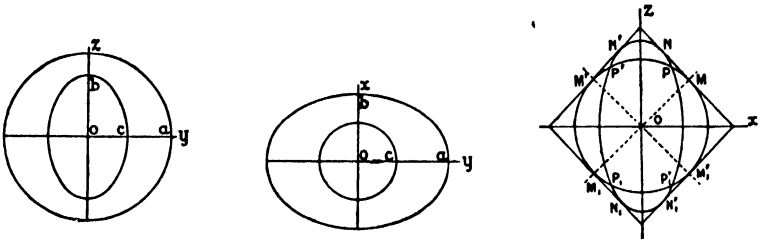


FIG. 237

which can easily be done by cutting them up in a suitable manner, and fastening them together again by means of strips of gummed paper. Such a model is shown in Fig. 238. The inner surfaces should be blackened on both sides of the section as indicated. By a little exercise of the imagination it is easy to see the general form of the inner and outer sheets, though a still better idea can be obtained from the wire or plaster models, which can be procured

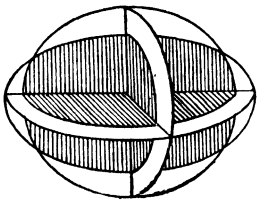


FIG. 238

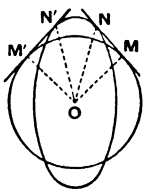


FIG. 239

from dealers in physical apparatus. The outer sheet has the general form of an ellipsoid with four depressions or pits similar to the pit found on an apple around the point where the stem is inserted, only much shallower. At these four points the two sheets come in contact, and some very remarkable optical phenomena are associated with this peculiar condition, which we will now investigate.

The Optic Axes or Axes of Single Wave-Velocity. — Consider now the *xz* section of the wave-surface, in which the curves intersect at four points and have four common tangents, one of which is represented by *MN* (Fig. 239). Planes passing through these tangents and perpendicular to the plane of the section, are tangent planes to the wave-surface. They touch the surface, moreover, not at two points, as was imagined by Fresnel, but all around a circle of

contact, a condition which can be represented by pressing a flat card against the dimple on an apple. This was first proved by Sir William Hamilton, who predicted from it the remarkable phenomena of internal conical refraction, which we shall consider presently. The lines OM, OM' perpendicular to the tangent planes are the directions in which a single wave only is propagated, for the planes MN and $M'N'$ touch both sheets. These directions are therefore the optic axes of the crystal.

Internal Conical Refraction. — Huygens's construction may be applied to determine the direction of the refracted rays, the points on the surface of the crystal becoming centres of wave-surfaces of the form which we have just studied. If light is incident upon the crystal in such a direction that the refracted wave-front is parallel to MN or $M'N'$ (Fig. 239) any line joining the centre O with the circle of contact of MN with the wave-surface is a possible direction of the refracted ray. The direction of the refracted ray will depend on the direction of the vibration in the incident wave-front.

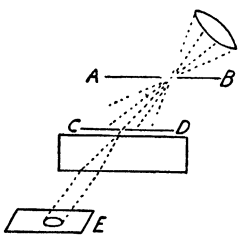


FIG. 240

The type of the vibration will not be altered by the crystal, since the wave-front is moving parallel to an optic axis, but the direction of the ray will depend on the plane of polarization. If the incident light is polarized in all possible planes, *i.e.* unpolarized, the ray upon entering the crystal will open out into a cone, each elementary ray of the cone being plane-polarized. This result was predicted from theory by Sir William Hamilton and verified by Lloyd [1] with a plate of aragonite cut so that its faces were equally inclined to the two optic axes.

A divergent cone of light from a screen AB (Fig. 240) perforated with a very small hole, upon which sunlight is concentrated by means of a lens, is intercepted by a second perforated screen CD. This screen can be moved about over the surface of the crystal, and serves to isolate a narrow pencil from the divergent cone. In general, if the transmitted light is received upon a screen at E, two spots of light appear, but by moving the screen CD about, it is possible to find a position such that the two spots run together into a ring of light, the diameter of which is independent of the distance of the screen E from the lower face of the crystal plate. This proves that the rays leave the plate in a parallel direction, notwithstanding their strong divergence within the crystal. The angle of the cone was found to be 1° 50′, while the value calculated was 1° 55′, a very close agreement between theory and experiment.

[1] *Trans. Roy. Irish Ac.*, vol. xvii, p. 145, 1833.

Axes of Single Ray Velocity. External Conical Refraction. — The directions determined by joining the point O (Fig. 239) with the points at which the two sheets of the wave-surface meet are termed the axes of single ray velocity. At each one of the conical points or pits an infinite number of tangent planes can be drawn to the surface, which collectively form a tangent cone. A rough model of such a cone can be made by cutting a paper circle along a radius and then pasting the edges together, making them overlap a little. This cone fits into the conical depressions of the wave-surface.

Suppose now that a ray is travelling within the crystal along the axis of single ray velocity, and emerges from the surface of the crystal. The direction of the ray after refraction out into the air is determined by the position of the plane tangent to the element of wave-surface corresponding to the ray. For example, suppose we are dealing with a simple spheroidal wave starting at O within the crystal (Fig. 241). We wish to determine the direction of the ray OB after emergence. This direction will be that traversed by

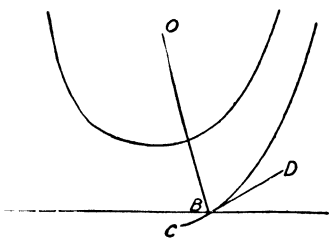


FIG. 241

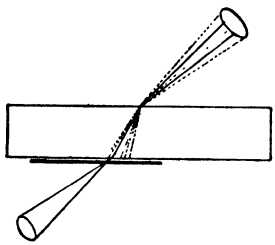


FIG. 242

a plane-wave CD tangent to the spheroid at B. In other words, the small element of the wave at B can be considered as a portion of the tangent plane. The direction of the refracted ray is thus seen to be determined by the position of the plane tangent to the wave-front at the point where it intersects the surface.

Now a ray travelling along an axis of single ray velocity has an infinite number of tangent planes which envelop a cone, and the refracted ray may pursue a direction determined by any one of them.

It will therefore open out into a hollow cone, and if the light be received upon a screen we shall see a ring which increases in diameter as the distance from the crystal face is increased. The phenomenon is exhibited by concentrating a pencil of rays upon the surface of the crystal. This converging system of rays contains the hollow cone of rays which we should have if we transmitted a ray

up through the crystal along the axis of the single ray velocity. The cone is indicated by solid lines (Fig. 242), the superfluous rays which pursue other paths in the crystal being indicated by dotted lines. A screen perforated with a small hole limits the emergent light to the ray which has traversed the plate in the direction of the single ray axis, and if the beam which issues is received upon a white screen it is seen to have the form of a hollow cone.

Colors of Thin Crystal Plates. — In the Chapter on Polarization we have seen that a beam of plane-polarized light falling upon a crystal plate cut perpendicular to the optic axis (for example, a film of mica or selenite) is in general doubly refracted, that is, the incident vibration is resolved into two mutually perpendicular vibrations, one polarized in, and the other at right angles to, the principal plane. The two disturbances traverse the crystal with different velocities, and consequently emerge with a difference of phase depending upon the thickness of the crystal plate. The plane vibration on entering the medium becomes transformed into an elliptical vibration, owing to the different velocities of the two rectangular components. As the disturbance proceeds its type changes, becoming circular, elliptical and plane in succession, each plane phase being turned through 90° with respect to the phase immediately preceding or following. It is obvious that if the plate is thick, and the two rays become separated by double refraction, each ray will be plane-polarized, that is, we shall no longer have a circular and an elliptical type. If white light falls upon the plate the difference of phase at emergence of the two components will vary with the wave-length, certain colors, for example, emerging plane-polarized parallel to the original plane of polarization; other colors polarize at right angles to it. Certain colors will, therefore, be absent when the emergent light is examined with Nicol prism held with its principal plane parallel to the principal plane of the polarizing Nicol. On rotating the Nicol through 90° each color changes to its complementary tint for obvious reasons. The state of polarization for waves of length intermediate between those specified will, in general, be of the circular or elliptical type.

Colors of Crystal Plates in Convergent and Divergent Polarized Light. — A remarkable series of phenomena are presented when we examine crystal plates, cut in various ways, in a beam of strongly convergent or divergent light. Colored fringes of varied forms appear crossed by dark crosses and brushes, the variety being almost as great as in the kaleidoscope. A complete investigation of the forms which occur under all possible conditions is hardly profitable, and we will examine a few typical cases only.

The simplest form of polariscope for viewing the rings and crosses in convergent light is the tourmaline tongs. When the crystal plate is placed between the tourmalines, and the eye brought close to the apparatus, which is directed towards a brilliant light of large size, such as the sky, the rays which enter the pupil have traversed the crystal in the form of a cone of wide aperture, as shown in Fig. 243. Tourmaline crystals are, however, usually so strongly colored, that only an imperfect idea of the color distribution can be obtained in this way, though the general form of the fringes can be made out. It is therefore customary to use some such arrangement as that shown in the lower part of the figure.

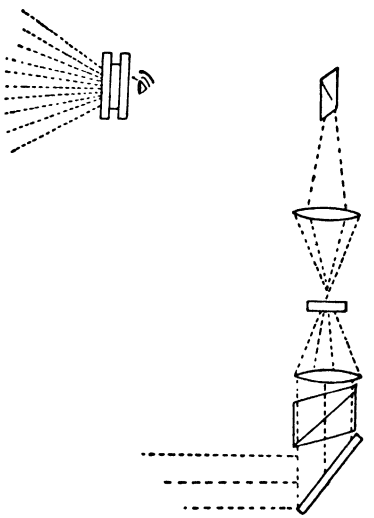


FIG. 243

Uniaxial Crystal Cut Perpendicular to Axis. — Consider what happens when a cone of plane-polarized rays diverging from S' (left hand, Fig. 244) passes through a crystal plate, the central ray SO of the cone coinciding with the optic axis. Consider the source S' in front of the plane of the paper, and let the vibrations be vertical. The ray incident at O passes through the plate in the direction of the optic axis, and its vibration plane remains unaltered. Other rays in general will suffer double refraction, and emerge with a phase-difference between the components of the vibration. This will not be true, however, for certain rays. Consider the ray $S'P$. The direction of vibration is in the principal plane $S'OP$, *i.e.* the plane containing the ray and the optic axis, and it will therefore be transmitted by the crystal without resolution. The same is true for the ray OP', since in this case the vibration is perpendicular to the principal plane. Hence all rays striking the crystal plate along OP or OP' or their prolongations, will not suffer double refraction, and will be wholly transmitted or completely stopped by a Nicol held behind the plate, according as its principal plane is vertical or horizontal. Consider now a ray incident at some other point, say Q (right hand, Fig. 244). The vibration a will be resolved into two components, b and c, one lying in the principal plane SOQ and the other perpendicular to it. The vibrations will traverse the crystal with different velocities.

and emerge with a difference of phase, which will depend upon the thickness traversed, and also upon the wave-length of the light. Now the thickness traversed will increase as we pass from O to Q, owing to the increasing obliquity of the rays. The phase-difference of the emergent components will therefore vary

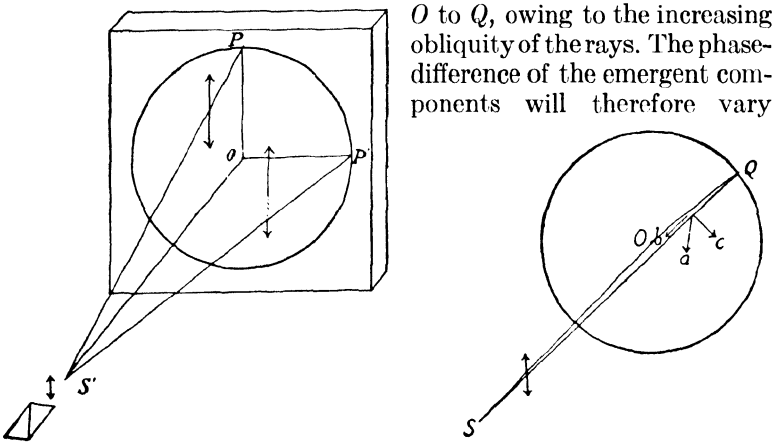


FIG. 244

along the line OQ, and the emergent light at some points on this line will be polarized in the same plane as the incident light, at other points in a plane perpendicular to it. The analyzing Nicol will quench one or the other, according to its position. By symmetry the conditions of equal phase-difference will occur along concentric circles with a common centre at O. We shall accordingly see bright and dark circles surrounding O if the light is monochromatic, and colored fringes if it is white. These circles will, however, be inter-

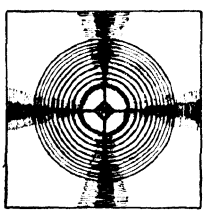


FIG. 245

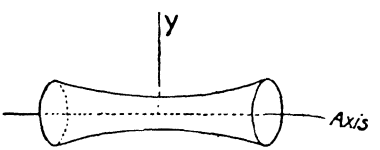


FIG. 246

rupted along the lines OP and OQ (for reasons above specified) by a cross which appears bright or dark according to the position of the analyzing Nicol (Fig. 245).

Isochromatic Surfaces. — The characteristic of a fringe is that the retardation δ is constant along its length, and the locus of points in space for which δ is constant is called an isochromatic surface. For every value of δ there will be a corresponding sur-

face, and if we describe these surfaces around S as an origin, with retardations of 1, 2, 3, 4, etc., half wave-lengths, the intersections of these surfaces with the second surface of the crystal will determine the isochromatic lines or fringes. The form of the surface is shown in Fig. 246. Its sections with the surface of a plate cut perpendicular to the axis are circles, with a plate parallel to the axis hyperbolae.

Isochromatic Surfaces in Biaxial Crystals. — The form of the surface in biaxial crystals is shown in Fig. 247. A section parallel

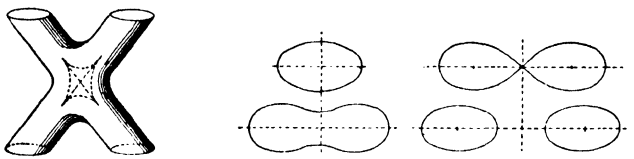


FIG. 247

to the plane containing the axes gives us curves closely resembling hyperbolae. A section perpendicular to the bisector of the angle between the optic axis gives us a family of lemniscates. Sections in planes along a, b, c, d give us fringes of the form shown on the right-hand side of the figure. These different curves correspond to

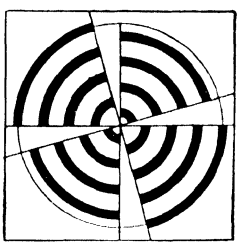


FIG. 248

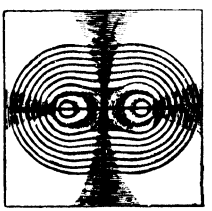


FIG. 249

successive values of δ, and they may all be seen simultaneously, as in Fig. 248.

The region of constant illumination, which in the case of uniaxial crystals had the form of a cross, in the case of biaxial crystals presents the appearance of a double brush of hyperbola form. The subject of the various modifications which the fringes and brushes may undergo is a very large one, but its study teaches us very little regarding the phenomenon of double refraction, the problems being purely geometrical. We shall examine but one other case, the remarkable transformation of a biaxial into a uniaxial crystal resulting from an elevation of temperature.

Position of Axes as a Function of Temperature. — A remarkable phenomenon occurs when certain biaxial crystals are heated, for as the temperature rises, the angle between the optic axes becomes decreased until the axes finally coalesce, the crystal becoming uniaxial. An oblique section of selenite is usually used for exhibiting the phenomenon. As the plate is warmed the lemniscates close in, the centres approaching, and presently meeting, at which stage the isochromatic fringes are circles crossed by a rectangular cross. A further elevation of temperature causes the axes to cross one another, so to speak, the crystal becoming again biaxial. The experiment makes one of the most beautiful lantern demonstrations ever devised.

Phenomena Exhibited by Twin Crystals. — Calcite is sometimes found with one or more layers crystallized in opposite directions. Such crystals sometimes show the rings and crosses without either polarizer or analyzer, the front and back parts of the crystal taking their place, and the oppositely crystallized plane serving as the thin film. A slice from a nitre crystal frequently exhibits four systems of rings.

If a crystal possesses rotating power still further complications result, notable among which are the beautiful spirals described by Airy and named after him. They appear when plates of right- and left-handed quartz cut perpendicular to the axis are superposed and viewed in convergent light. Or a single plate may be made to exhibit them if it is placed on the lower mirror of the Nörremberg polariscope, on account of the reversal of the rotation.

Convergent Circular Light. — If a quarter-wave plate is interposed between the first Nicol and the crystal plate in a converging polariscope, the appearances are completely altered. As we should expect, the black cross disappears almost completely, the arms being replaced by thin lines of nebulous gray, which rotate with the analyzer without changing in appearance. The rings in adjacent quadrants are dislocated as shown in Fig. 249, the light rings in one quadrant being opposite the dark ones in the next.

The explanation of this can easily be found by working out the resolution of the circular vibration in two opposite quadrants, taking care to distinguish between components parallel to the radii and those perpendicular to them.

Double Refraction in Non-Crystal Media. — Many of the phenomena of double refraction can be observed in isotropic substances subjected to strain, or to sudden differences of temperature. Glass plates squeezed in a vise (Fig. 250) and viewed between crossed Nicols exhibit most beautiful colored fringes, the lines of strain being clearly indicated.

Tyndall found that a long strip of glass thrown into sonorous vibration restored the light when placed between crossed polarizing prisms. The experiment has since been modified in a beautiful manner, the transmitted light being examined in a revolving mirror and found to be restored periodically, the band appearing broken up into beads, showing that the double refraction was coincident with the vibration. On inserting a selenite plate the band was found to vary in color.

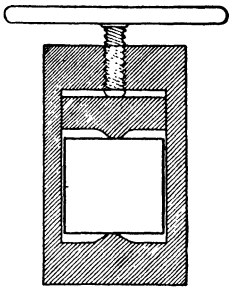


FIG. 250

A permanent strained condition can be established by heating a block of glass nearly to a red heat and cooling it suddenly. Polarized light is an extremely sensitive test for imperfect annealing. Prince Rupert drops make excellent objects. They can be easily prepared by melting the end of a glass rod in a powerful blast-lamp, and allowing the drop to fall into a bowl of water with some filter paper on the bottom. Four out of five fly to pieces, but with a little practice a number can be prepared in a short time. They are best viewed by immersing them in a small rectangular cell of glass filled with a mixture formed by dissolving about ten parts of chloral hydrate in one part of hot glycerine. This mixture has the same refractive index as the glass.

Sheets of cellophane, which have become doubly refracting by strains introduced in the process of manufacture, make brilliant objects between crossed Nicols, especially when two or more are superposed in various orientations.

*

CHAPTER XI

METEOROLOGICAL OPTICS

The refraction and diffraction of light by water drops, ice crystals, dust particles, etc., in the earth's atmosphere, give rise to a great variety of optical phenomena, the study of which is of help in determining the atmospheric conditions producing them.

Of these the commonest example is the rainbow, the complete investigation of which is an extremely complicated problem. The elementary explanation given in 1637 by Descartes accounts for the bow by the laws of geometrical optics; it is wholly inadequate, however, for the explanation of the great diversity of color arrangement, and the fainter secondary bows which are so often seen. Other types of atmospheric halos, parhelia or mock suns, etc., are due to ice crystals, which may or may not have a definite orientation. If the air is quiet, *i.e.* free from turbulent motion, the ice spicules will lie for the most part with their long axes horizontal, the position assumed by a spindle-shaded body falling through a resisting medium. This orientation may give rise to concentrations of light at definite points, as in the case of mock suns. Mirage, scintillation, and related phenomena come under the head of meteorological optics, but they have been sufficiently treated in the Chapter on Refraction. A more comprehensive study of the whole subject will be found in Pernter's *Meteorologische Optik*. We will begin by a study of the rainbow.

The Rainbow. — The first theory of the rainbow was given by Descartes in 1637. He calculated the paths of various rays which together formed the parallel bundle incident upon the raindrop. These rays after suffering one internal reflection emerge from the under side of the drop in different directions. He made calculations, by the aid of Snell's law, of the paths of 10,000 rays incident on different parts of one side of the drop, and discovered that the angle between the extreme issuing rays, from the 8500th to the 8600th, is measured in minutes of arc. These are the so-called least deviated rays, and they leave the drop in the direction RT, Fig. 251, which is the direction followed by the ray figured in the drop, incident at P. Rays incident above and below P leave the drop in directions indicated by the dotted lines. There is thus a concentration of energy, or increased illumination along the direc-

tion RT. This light which suffers minimum deviation by the drop is the light by which the illuminated drops in the bow are seen. Its deviation is 137° 29′ and the angle which it makes with the incident rays is 42° 31′. All drops therefore which lie on a circle (more exactly a cone), the centre of which is opposite the sun, and of which the radius is 42° 31′, appear strongly illuminated. This fact, which was laboriously ascertained by Descartes, can be easily found by the method of maxima and minima. We will make

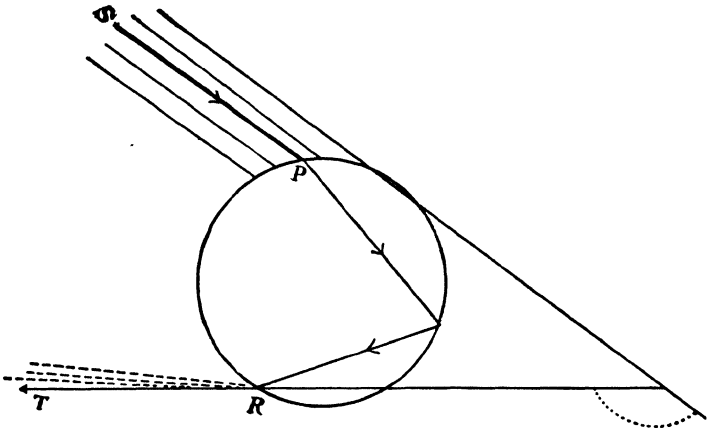


FIG. 251

the calculation for any number k of internal reflections, which will give us the position of other bows.

The deviation D is equal to $2(i-r)+\pi-2r$ for a single internal reflection, or for k internal reflections

$$D=2(i-r)+k(\pi-2r).$$

Differentiating,

$$dD=2di-2(k+1)dr,$$

and equating dD/di to zero gives

$$\frac{di}{dr}=k+1.$$

If μ is the refractive index of water, we have

$$\sin i=\mu \sin r,$$

$$\cos i di=\mu \cos r dr,$$

$$\frac{di}{dr}=\mu\frac{\cos r}{\cos i};$$

or

$$\mu\frac{\cos r}{\cos i}=k+1.$$

Squaring, $\mu^2 \cos^2 r = (k+1)^2 \cos^2 i.$

Adding $\mu^2 \sin^2 r = \sin^2 i,$

gives us $\mu^2 = 1 + (k^2+2k)\cos^2 i$

or $$\cos i = \sqrt{\frac{\mu^2-1}{k^2+2k}}.$$

A second differentiation gives us a positive quantity which shows us that we are dealing with a minimum. Applying the last equation we find, since for water $\mu = 4/3$ for $k=1$,

$$i = 59^\circ\ 23', \quad r = 40^\circ\ 12'$$
$$D = \pi + 2i - 4r = 137^\circ\ 58'.$$

The radius of the bow $\delta = 180^\circ - 137^\circ\ 58' = 42^\circ.$

For $k=2$, $D = 232^\circ\ 30'.$

This gives us a bow radius $232^\circ\ 30' - 180^\circ$ or $\delta = 52^\circ\ 30'$.

If we take dispersion into account it is clear from Fig. 252 that for one internal reflection δ will be larger for the red than for the violet, consequently the bow will be red on the outside and violet

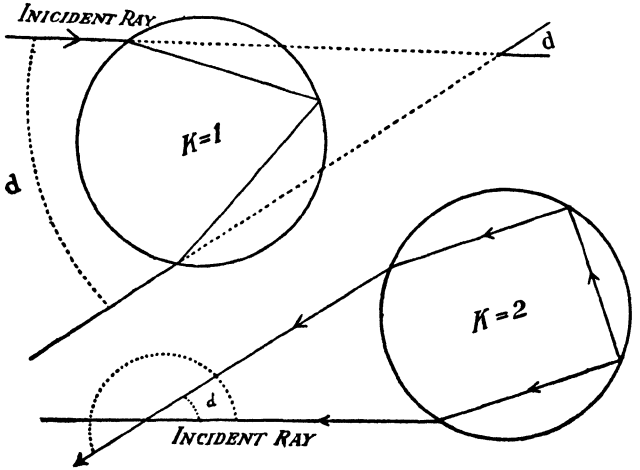


FIG. 252

on the inside. Moreover the space between the bow and its centre will be more or less luminous, for the drops in this region will be seen by the feebler rays for which D is greater and δ consequently less. The bow due to the two internal reflections will appear outside of this, the drops being seen by rays which enter on the under

side and leave on the upper side. The dispersion in this case will make D, and consequently δ, largest for the violet. This bow will accordingly be violet on the outside and red on the inside. The feebler rays will illuminate the region outside of this bow, the region between the two bows being absolutely dark, so far as light from the rain drops is concerned.

Three and four internal reflections give bows which are behind us as we face the primary bow, *i.e.* we should have to face the sun to see them. When the shower is between us and the sun, we see the drops powerfully illuminated by light which is refracted without undergoing internal reflection. This light overpowers the third and fourth bows; no bow corresponds to this directly refracted light, for the intensity falls off gradually as the angle between the drops and the sun increases. The fifth bow coincides very nearly with the second, while the sixth falls inside the first. They are never seen, however, owing to the diminution in the intensity of the light by the refraction which accompanies each internal reflection.

The bow which is sometimes seen inside of the primary bow must not be mistaken for the sixth bow, as it is formed in quite another way, and cannot be accounted for by the elementary theory.

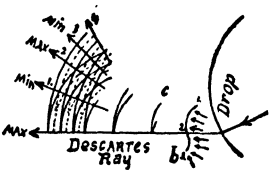


FIG. 253

We will now take up the complete treatment which was first given by Airy.

According to the elementary treatment, the succession of colors in the rainbow should always be the same and the diameter and width constant. This is not the case. Rainbows are frequently seen which are not parts of true circles, and the succession of colors is frequently seen to be different in different parts of the same bow.

This was shown by Airy to depend upon the peculiar form of the wave-front which emerges from the drop. A plane or flat wave, incident upon a transparent sphere (rain drop), can be shown by very elementary methods to acquire, after two refractions and one reflection, the peculiar shape shown in *b*, Fig. 253, where the curved line 1, 2, 3 represents the wave-fronts, the elements of which are travelling in the direction of the arrows. The portion 1, 2 is convex in the direction of propagation, and will of course go on expanding; the part 2, 3 is concave, and converges to a focus. The curvature varies as we pass along the wave-front, being greatest at 3, zero at 2 (where the front is plane), and having a large value again at 1. The element at 3 comes to a focus first, passes through it and becomes convex instead of concave, forming a

"cusp" on the wave as shown at *c*. Successive elements of the wave-front above 3 pass in turn through foci and build the rear front of the cusped wave. A full treatment of the propagation of cusped waves has been given in the Chapter on Reflection.

The caustic surfaces are always bordered by interference fringes, for we are dealing with two sets of wave-trains, formed by the front and rear surfaces of the cusped waves. If we draw the surfaces, as in Fig. 253, representing the crests of the waves by solid lines, and the troughs by dotted lines, we find that crests intersect troughs along the arrows 1 and 3, consequently these represent the positions of the interference minima. This method of looking at the phenomenon is more elementary than that given by Airy, who integrates the effect of the whole wave at a given point in front of it, as in diffraction problems. We thus see that in addition to the least deviated ray of Descartes, indicated by the arrow 0, there will be other directions (arrows 2, 4, etc.) in which there is considerable illumination. With a monochromatic sun we should, therefore, have a number of concentric bows, as can be shown experimentally with a spectrometer and glass cylinder or ball. The distance between these bows will vary with the size of the rain drops, and since with white light we have, theoretically, an infinite number of bows, it is clear that the color of the actual rainbow at a given point can only be determined by computing the "mixed color" due to the superposed bows.

This is merely a qualitative explanation. For a complete solution we require the distance between the successive maxima and minima, and the position with reference to the least-deviated ray of Descartes. Airy solved the problem by integrating the effect of the wave-front at exterior points, and found that the first maximum, which gives the primary bow, fell a little inside of the Descartes ray, which makes the diameter of the bow slightly less than that called for on the geometrical theory. The discrepancy increases as the size of the drops diminishes. Very large drops give a bow of practically the same size as predicted by the elementary method; very small drops, however, give a bow which may have a diameter several degrees less. We thus see how it is possible for the curvature of a bow to change, if the drops which form one part of the bow are smaller than those which form another.

The successive maxima and minima give a series of so-called supernumerary bows within and concentric to the primary bow. With large drops they are closer together than with small. The actual rainbow is therefore a superposition of a number of bows, and the succession of colors will depend upon the spacing of the supernumerary bows.

The color distribution was first worked out by Pernter, who calculated the tints resulting from the mixtures of the primary colors of the superposed bows. If the red of the second bow falls upon the green of the first, we shall have a bow with an abnormally broad yellow band, for red and green lights, when mixed, form yellow. This is a very common type, a bow of red and yellow with green and blue nearly absent. If the drops are smaller, the red of the second may fall upon, or even within, the violet of the first. In this case we see a second bow just within the first, and perhaps a third and fourth.

It is not difficult to investigate these phenomena in the laboratory. Allow the light from an arc lamp to fall upon a vertical glass rod one or two millimetres in diameter, and view the refracted light with a telescope, placed close to the rod, covering the eye with a piece of red glass. A large number of maxima and minima will be seen; these are the fringes which border the caustic. Standing with our back to the arc, in a dark room, we see, if we throw the spray of an atomizer in front of us, not only the primary and secondary bow, but the first supernumerary bow, just inside of the primary.

By suspending a minute drop of water from the end of a very small glass thread (previously greased) and holding it very close to the eye, a little to one side, standing with our back to the arc, or the sun reflected in a mirror, we can see a portion of both the primary and secondary bow most intensely colored, with the very dark region between them, and a set of supernumerary bows within the primary, and outside of the secondary. The best method of getting the drop is to draw out a piece of small glass tubing into a fine capillary, breaking it off at a point where its diameter is about 0.3 mm. Grease the end slightly, introduce a little water, and blow a minute drop on the end of the capillary. This is the most instructive experiment of all, as we can vary the size of the drop and note the effect upon the spacing of the supernumerary bows.

The fact that we can see portions of a number of bows with a single drop presents no difficulty, for the drop is so close to the eye that most of the differently deviated rays enter the pupil. It is interesting to compare this case with that of the halos to be described presently, in which we see a bow resulting from the diffracted rays of light, the drop acting as an obstacle.

Polarization of Light in the Rainbow. — The refraction and reflection of light by the rain drops produces fairly strong polarization and Rinne and Rosch have calculated the intensity ratios of

the perpendicular and parallel components for the different bows. These ratios are the measure of polarization

Bow No	1	2	3	4	5	6
$\frac{I_s}{I_p} =$	21	8.5	8.7	6.3	6.5	7.5

Observation of the bows through a properly oriented Nicol shows a zone-like extinction of some of the light. The calculations were made from the Fresnel formula for reflection.

Halos, Mock Suns and Related Phenomena. — The reflection and refraction of the sunlight by small ice crystals in the air give rise to a very complicated series of phenomena, which, unlike the rainbow, can be seen at all altitudes of the sun. They may be summed up briefly as follows:

A colored circle, surrounding the sun, red on the inside and white on the outside, having a radius of 22°, known as the *22° Halo* or corona.

A larger circle, similar in appearance, but fainter, called the *46° Halo.*

A white horizontal circle passing through the sun, and having a constant height above the horizon, called the *Parhelic Circle.*

Concentrations of light on the parhelic circle, four colored ones at the points where it intersects the halos, are called *Parhelia* or mock suns. One white one at 180° from the sun is called the *Anthelion,* and two at 120° the *Paranthelia.*

Two oval arcs, sometimes joined together forming an ellipse which circumscribes the 22° halo, is known as the *Circumscribing Oval.*

The *Tangential Arcs* touching the 46° halo.

The *Oblique Arcs of Lowitz* which are sometimes seen under the parhelia when they extend beyond the 22° halo. These, the

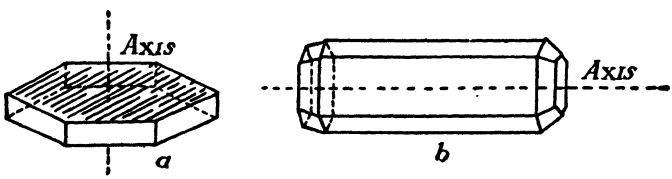


FIG. 254

chief halo phenomena, can all be explained by the laws of reflection and refraction in the hexagonal ice crystals which have the forms shown in Fig. 254. In still air, these crystals will fall slowly in the position shown, since bodies moving through a resisting medium set themselves in the position of greatest resistance. A

mistake has been made by several writers, Bravais and Pernter, for example, who state that the crystals fall in the position of least resistance. It must not be supposed, however, that this error invalidates in any way Pernter's admirable treatment of the optical problems involved, for in each case (fortunately!) we may substitute a horizontal hexagonal plate for a vertical hexagonal prism, and *vice versa*, without affecting the orientation of the angles. It is well to bear this in mind, as the formation of plates or spiculae of ice seems to depend upon the temperature, and it is possible that approximate estimates of the temperatures of the clouds could be made by considering whether the parhelia observed were of the type due to hexagonal crystals with the principal axis vertical (*i.e.* flat plates), or with the axis horizontal (*i.e.* spiculae). Erroneous conclusions would of course be drawn if the mistake here pointed out were not corrected.

This orientation of the crystals taken into account, we can explain all of the above-mentioned phenomena. The hexagonal prisms and plates can transmit light in various ways. The 120° angles are too large of course to act as prisms, but the alternate faces form 60° prisms, which transmit light at a minimum deviation of 22°. It is this transmission that is responsible for the 22° halo. The crystals are not considered as oriented in this case. A crystal situated 22° from the sun can send light at minimum deviation to the eye, and it can do this in six different positions. Crystals nearer the sun can send no light to the eye by this type of refraction and the inner edge of the halo is therefore sharply defined.

As we have seen in the Chapter on Refraction, the maximum quantity of light is transmitted by a prism at minimum deviation, hence the narrowness of the halo. The region outside the halo is slightly luminous, due to the light transmitted by prisms not in this position. The 46° halo is produced in a similar way by refraction through the 90° angles of the crystals. A glass prism of 90° will not transmit light, but ice has a much lower refractive index, and if the crystal is in just the right position, a small amount gets through, which accounts for the faintness of the ring.

These two halos are the only phenomena which can be explained by the action of crystals which have their axes directed in a fortuitous manner. All of the others require still air for their production.

The parhelic circle is white, and is due to reflection from the six sides of the flat plates, and the flat end faces of the elongated ones, the faces in question acting as vertical plane mirrors, which reflect the light of the sun to the eye. The crystals which are at the same elevation above the horizon as the sun are the ones which can send

light to the eye in this way. The light of the parhelic circle can also come from rays which enter the upper surfaces of the flat plates, and suffering total reflection at one of the sides emerge from the under side. The two refractions in this case neutralize the dispersion, consequently no color is produced.

The parhelia are produced by refraction through the 60° and 90° prisms, which are oriented as described and for the moment situated at position of minimum deviation.

When the sun is at a considerable altitude, the rays no longer pass through these prisms parallel to their bases, and the effective angle of the prism is increased, and with it the deviation, which causes the parhelia to appear well outside of the halo, but always on the parhelic circle. Parhelia of 46° have been observed, but only rarely. The 22° ones are frequently observed in winter, even in our latitude. The anthelion is a rarely observed phenomenon, only about thirty appearances of it having been recorded in the past 250 years. The paranthelia are still rarer. Both are formed by two internal reflections combined with two refractions. Pernter's treatment is not very satisfactory in this case, as he makes an erroneous assumption regarding the orientation of the crystals.

There are other phenomena too numerous to mention, such as the vertical pillars of light, seen above or below the sun, formed by reflection from the under surfaces of the horizontal flat plates. They are most frequently seen when the sun is near or below the horizon. Compare these with the parhelic circle. Crosses are sometimes seen, as when the ice cloud is located only at the point of intersection of a halo with the parhelic circle.

The reader is referred to Pernter's *Meteorological Optics* for a fuller account of the almost innumerable phenomena of this sort. There is another type of halo frequently seen surrounding the sun or moon, which is due to diffraction. These are usually of comparatively small diameter. The explanation of the manner in which they are produced has been given in the Chapter on Diffraction, and they can be distinguished from the halos just described from the circumstance that they are always red on the outside.

CHAPTER XII

THEORY OF REFLECTION AND REFRACTION

Maxwell's Electromagnetic Theory. — Clerk Maxwell originated the very fundamental idea that electrical and magnetic phenomena can be propagated in the form of waves, and he formulated a set of equations, known by his name, which form the foundations of the electromagnetic theory of light. He started out with the then revolutionary idea that all electrical currents form closed circuits and that, in the case of a condenser formed of two plates oppositely charged and then connected by a wire, there was a current between the plates, even though separated by an absolute vacuum, equal to the current of discharge flowing in the wire. This current he called a displacement current, without attempting to define the physical process involved. The displacement current exists, however, only when the electric force is changing, as in the process of charging and discharging a condenser. These displacement currents being transient were assumed by Maxwell to give rise to transient magnetic fields, and a changing magnetic field is surrounded by lines of electric force as we know from the induction of electric currents by varying magnetic forces. Energy can thus be propagated in the form of waves by a process which can be simply described as follows:

If a momentary displacement current be started in a plane, lines of magnetic force will be generated above and below the plane at right angles to the direction in which the displacement current flows. These will in their turn give rise to electric forces, the magnetic flux above or below the plane generating a force *in* in the plane opposed to the force which initially caused the displacement current (which will thus be annulled). In the planes above and below the magnetic flux, the force will be in the same direction as the original force and hence cause displacement currents parallel to the original one. This process occurring continuously results in the propagation of a plane-wave of electromagnetic energy, the wave-front being parallel to the plane in which the original displacement current occurred.

Derivation of Maxwell's Equations. — We designate the displacement current by i or i' according as it is expressed in electrostatic or electromagnetic units.

The c.g.s. or electromagnetic system of units starts out with the definition of unit magnetic pole, which is a pole of such strength that, when placed at a distance of 1 cm. from an equal pole, it will exert a force of one dyne upon it. The unit of magnetic field strength is the "Gauss," which is the strength of field which exerts a force of one dyne upon unit pole, *i.e.* it is the strength of the field at a distance of 1 cm. from unit pole.

The unit of current is the current which, when flowing in a wire, across a magnetic field of unit strength, will cause 1 cm. of the wire to be acted upon by a force of 1 dyne. It is roughly 10 amperes. The unit of quantity is the amount of such a current which flows across a given section of the wire in one second. The Faraday, or electrostatic, system of units starts out with unit charge, as the charge which, when placed at a distance of 1 cm. from an equal charge, exerts a force of one dyne upon it. The electrostatic unit of current is the flow of one Faraday unit of charge per second, and the unit of magnetic field is a field which will exert a force of one dyne on a centimetre of wire carrying unit current.

We may also measure the current in the c.g.s. system by the work done in carrying unit magnetic pole once around the current against the lines of magnetic force which surround the current.

The force exerted by a current of intensity i' in a straight wire in a pole at distance r is $2i/r$, and if the pole is carried once around a circle of radius r it moves a distance of $2\pi r$ against this force; the work done upon it is therefore $4\pi i'$.

If the current i' is of unit strength (c.g.s. system) the force will be 1 dyne if $r=1$.

The current i which flows through cross section q is defined as the number of electrostatic units which traverse q in unit time, so that if the quantity of electricity de flows through q in the element of time dt, we have

$$i=\frac{de}{dt}, \qquad (1)$$

and if q is equal to unit cross section, i is equal to j, the current density. The components of j along the x, y, z axes we will designate j_x, j_y, j_z. We will now derive an expression for the current in electromagnetic measure. The current is surrounded by a magnetic field, the lines of force being circles in the case of a current flowing along a wire. An isolated magnetic pole will follow these lines of force, travelling around the wire as long as the current continues to flow.

We have seen that the work done in carrying unit magnetic pole once around a wire carrying a current i' is

$$A = 4\pi i'.$$

Now the work is represented by the force multiplied by the distance through which it acts. Assume that we have a rectangle dx, dy, which is traversed normally by a current $i' = j_z' dx dy$, j_z' being the z component measured in electromagnetic units. If the current flows towards the observer (Fig. 255), a plus magnetic pole will be carried around dx, dy in the direction indicated by the arrows. The total work done by the current in moving unit pole around the rectangle will be

$$A = \alpha dx + \beta' dy - \alpha' dx - \beta dy, \tag{2}$$

in which α and β are the components of magnetic force along AB and AD, and α' and β' are the components along DC and BC. α may not be constant along dx, but if we regard it as variable, for example having the value α at A and $A + \partial\alpha$ at B, the average value will be $\alpha + \partial\alpha/2$, which, when multiplied by dx, gives us αdx + an infinitesimal of the second order. The minus signs occur for the obvious reason that the forces along DC and CB are oppositely directed from the forces along AB and AD. α' differs from α since it works along a line, the y coördinate of which is greater by an amount dy than that of AB. Under certain conditions of course α would be equal to α'.

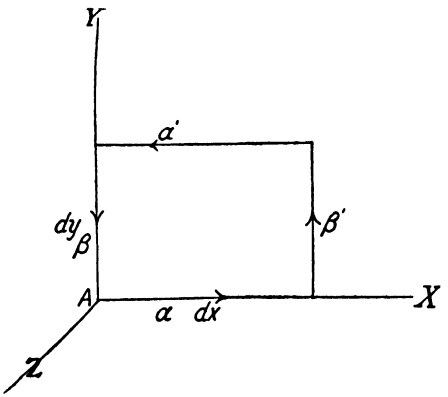


FIG. 255

If dy be taken sufficiently small $(\alpha' - \alpha)/dy$ may be regarded as the partial differential coefficient $(\partial\alpha/\partial y)dy$, and we have

$$\alpha' = \alpha + \frac{\partial\alpha}{\partial y} dy \text{ and } \beta' = \beta + \frac{\partial\beta}{\partial x} dx.$$

We now have for the work, substituting these values in (2) and cancelling,

$$A = \left(\frac{\partial\beta}{\partial x} - \frac{\partial\alpha}{\partial y}\right) dx dy,$$

and since $A = 4\pi i' = 4\pi j_z' dx dy,$

$$4\pi j_z' = \frac{\partial \beta}{\partial x} - \frac{\partial \alpha}{\partial y}, \text{ and similarly, } 4\pi j_x' = \frac{\partial \gamma}{\partial y} - \frac{\partial \beta}{\partial z}, 4\pi j_y' = \frac{\partial \alpha}{\partial z} - \frac{\partial \gamma}{\partial x}, \quad (3)$$

Maxwell's differential equations of the magnetic field.

If c represents the ratio of the two systems of units, *i.e.* $i/i' = c$ and $j_x/j_x' = c$, we can introduce j (defined electrostatically) into the equations, which now become

$$\frac{4\pi}{c} j_x = \frac{\partial \gamma}{\partial y} - \frac{\partial \beta}{\partial z}, \quad \frac{4\pi}{c} j_y = \frac{\partial \alpha}{\partial z} - \frac{\partial \gamma}{\partial x}, \quad \frac{4\pi}{c} j_z = \frac{\partial \beta}{\partial x} - \frac{\partial \alpha}{\partial y}. \quad (4)$$

These equations hold for all media, for it can be shown that the work done in carrying the magnetic pole around the circuit is independent of the nature of the medium. The quantity c has the dimensions of velocity, and can be determined by observing the magnetic effect of discharging a quantity of electricity measured electrostatically through a circuit.

Similar equations connect the magnetic flux s with the changing electric field:

$$\frac{4\pi}{c} s_x = \frac{\partial Y}{\partial z} - \frac{\partial Z}{\partial y}, \quad \frac{4\pi}{c} s_y = \frac{\partial Z}{\partial x} - \frac{\partial X}{\partial z}, \quad \frac{4\pi}{c} s_z = \frac{\partial X}{\partial y} - \frac{\partial Y}{\partial x}. \quad (5)$$

Displacement Currents in Free Ether. — A displacement current will occur in the ether whenever the density of the lines of electric force changes, and the strength of the current will be proportional to the rate at which the change takes place.

We will now derive expressions which connect the displacement current with the electromotive intensity (electric force). Since a charge e sends out $4\pi e$ lines of force, the product of the current density and 4π will be the change in the number of lines of force in unit time. It is obvious that in the case of steady currents there will be no change in the number of the lines, but in the case of displacement currents, where the current strength is changing with the time, the density of the lines of force changes. We can now write

$$4\pi j_x = \frac{\partial N_x}{\partial t}, \quad 4\pi j_y = \frac{\partial N_y}{\partial t}, \quad 4\pi j_z = \frac{\partial N_z}{\partial t} \quad (6)$$

in which the expressions N_x, N_y, N_z represent the components of the density of the lines of electric force (polarization in free ether) parallel to the three axes. Similarly for the magnetic current we have

$$4\pi s_x = \frac{\partial M_x}{\partial t}, \quad 4\pi s_y = \frac{\partial M_y}{\partial t}, \quad 4\pi s_z = \frac{\partial M_z}{\partial t}. \quad (6a)$$

We can form an idea of a magnetic current such as we have in the case of light-waves in the following way: Suppose that we have a soft iron wire with a coil of insulated wire around one end of it, which is traversed by an alternating current. The density of the magnetic lines of force in the iron wire varies periodically, rising from zero to a maximum, and then falling to zero during the first half period, and then rising again to a maximum, with a reversal in the direction of the force, however. The wire is thus traversed by a periodic magnetic current, which is surrounded by circular lines of electric force, which set up alternating induced currents in conductors which are brought into the field.

In the free ether the electric force is considered as numerically equal to the density of the lines of force, so that we may substitute for N_x, N_y, N_z their equivalents X, Y, Z. Our equations now become

$$\left.\begin{array}{lll} 4\pi j_x=\dfrac{\partial X}{\partial t}, & 4\pi j_y=\dfrac{\partial Y}{\partial t}, & 4\pi j_z=\dfrac{\partial Z}{\partial t} \\ 4\pi s_x=\dfrac{\partial \alpha}{\partial t}, & 4\pi s_y=\dfrac{\partial \beta}{\partial t}, & 4\pi s_z=\dfrac{\partial \gamma}{\partial t} \end{array}\right\}. \qquad (7)$$

Substituting these values for the current in equations (4) gives us expressions which connect the variation of the electric force with the magnetic field which results from the displacement current.

$$\left.\begin{array}{lll} \dfrac{1}{c}\dfrac{\partial X}{\partial t}=\dfrac{\partial \lambda}{\partial y}-\dfrac{\partial \beta}{\partial z}, & \dfrac{1}{c}\dfrac{\partial Y}{\partial t}=\dfrac{\partial \alpha}{\partial z}-\dfrac{\partial \gamma}{\partial x}, & \dfrac{1}{c}\dfrac{\partial Z}{\partial t}=\dfrac{\partial \beta}{\partial x}-\dfrac{\partial \alpha}{\partial y} \\ \dfrac{1}{c}\dfrac{\partial \alpha}{\partial t}=\dfrac{\partial Y}{\partial z}-\dfrac{\partial Z}{\partial y}, & \dfrac{1}{c}\dfrac{\partial \beta}{\partial t}=\dfrac{\partial Z}{\partial x}-\dfrac{\partial X}{\partial z}, & \dfrac{1}{c}\dfrac{\partial \gamma}{\partial t}=\dfrac{\partial X}{\partial y}-\dfrac{\partial Y}{\partial x} \end{array}\right\}. \qquad (8)$$

Isotropic Dielectrics. — The equations which we have just derived do not hold for dielectrics, for in media the force exerted by two charges $e'e''$ at distance r is less than the force which would be exerted in the free ether, being represented by $e'e''/\epsilon r^2$, in which ϵ is the dielectric constant of the medium. The dielectric constant is greater than unity for all media, and we shall see that the velocity with which the wave is propagated in the medium becomes less as the dielectric constant increases.

In the same way the force between two magnetic poles in a medium is represented by $m'm''/\mu r^2$, μ being the magnetic permeability, a quantity which differs only slightly from unity except in the case of iron, and we shall see later on that we are justified in writing $\mu=1$ in practically all optical problems. The change in the law of the force which occurs in ponderable media makes it

necessary to modify our last equations, since with the same change in the current intensity the electric force is weaker in the proportion $1/\epsilon$, the current in dielectrics being represented by $4\pi j_x = \epsilon(\partial X/\partial t)$, etc., $4\pi s_x = \mu(\partial\alpha/\partial t)$, etc.

Equations (7) now become

$$\left.\begin{array}{lll} \dfrac{\epsilon}{c}\dfrac{\partial X}{\partial t}=\dfrac{\partial\gamma}{\partial y}-\dfrac{\partial\beta}{\partial z}, & \dfrac{\epsilon}{c}\dfrac{\partial Y}{\partial t}=\dfrac{\partial\alpha}{\partial z}-\dfrac{\partial\gamma}{\partial x}, & \dfrac{\epsilon}{c}\dfrac{\partial Z}{\partial t}=\dfrac{\partial\beta}{\partial x}-\dfrac{\partial\alpha}{\partial y} \\ \dfrac{1}{c}\dfrac{\partial\alpha}{\partial t}=\dfrac{\partial Y}{\partial z}-\dfrac{\partial Z}{\partial y}, & \dfrac{1}{c}\dfrac{\partial\beta}{\partial t}=\dfrac{\partial Z}{\partial x}-\dfrac{\partial X}{\partial z}, & \dfrac{1}{c}\dfrac{\partial\gamma}{\partial t}=\dfrac{\partial X}{\partial y}-\dfrac{\partial Y}{\partial x} \end{array}\right\}, \quad (9)$$

which expressions completely determine all properties of the magnetic field in an isotropic dielectric.

A comparison of equations (6) with the equations preceding (9) shows us that, if we consider the number of lines of force issuing from given charge independent of the surrounding medium, we have the condition within a medium of dielectric constant ϵ, and permeability μ,

$$N_x=\epsilon X,\ N_y=\epsilon Y,\ N_z=\epsilon Z,$$
$$M_x=\mu\alpha,\ M_y=\mu\beta,\ M_z=\mu\gamma;$$

in other words, the densities of the lines of magnetic and electric force are equal to the forces only in a vacuum, for which ϵ and μ both equal unity.

If a charge e is contained in the cube dx, dy, dz, $4\pi e$ lines of force issue from its surface. We can also reckon the number of lines issuing from the cube as the sum of the lines issuing from the six surfaces.

The two squares perpendicular to x contribute the part

$$-(N_x)_1 dydz+(N_x)_2 dydz.$$

By Taylor's Theorem $(N_x)_2=(N_x)_1+\dfrac{\partial N_x}{\partial x}dx$;

the two squares therefore contribute

$$\left((N_x)_1+\frac{\partial N_x}{\partial x}dx\right)dydz-(N_x)_1 dydz=\frac{\partial N_x}{\partial x}dx\cdot dydz.$$

The total contribution of all six faces is

$$\left(\frac{\partial N_x}{\partial x}+\frac{\partial N_y}{\partial y}+\frac{\partial N_z}{\partial z}\right)dxdydz.$$

Equating this to $4\pi e$, and bearing in mind the expression for N^x, M^x, etc., gives us, if we write $e/dxdydz=\rho$ the charge in unit volume,

$$4\pi\rho=\frac{\partial\epsilon X}{\partial x}+\frac{\partial\epsilon Y}{\partial y}+\frac{\partial\epsilon Z}{\partial z}, \qquad (9a)$$

an equation which we shall not make use of for the present.

We shall now develop equations which express the reflection coefficient and the state of polarization of the reflected and refracted light in their dependence on the angle of incidence and the refractive index of the medium.

These equations are very important and will be used frequently in subsequent chapters.

Boundary Conditions. — The whole matter of what happens to light when it falls on the boundary separating two media of different refractive indices hangs upon the values which the electric and magnetic forces of the wave assume on opposite sides of the boundary. The exact determination of the boundary condition for the components of the forces along the x, y and z axes is thus the starting point in the development of the theory of reflection and refraction.

We will begin by considering that the transition is abrupt, *i.e.* that the dielectric constant changes suddenly in crossing a mathematical plane, which we will take parallel to the xy plane of our coördinate system. If lines of electric force of density N_1 cross a boundary separating media of dielectric constants $\epsilon_1>\epsilon_2$ without change of direction a "perpetual motion" would be possible, since the electric force is less in the second medium, being $R_2=4\pi N_1/\epsilon_2$ while that in the first medium is $R_1=4\pi N/\epsilon_1$.

The components along the three axes $X_2Y_2Z_2$ will therefore be less in the second than in the first medium, and an unlimited amount of work could be derived by carrying a charged particle against the lesser force X_2 along the boundary, and after carrying it across the boundary (requiring no work) permitting it to move back in the first medium under the influence of the greater force X_1. This forces us to the conclusion that the x and y components must have the same value on opposite sides of the boundary, that is $X_1=X_2$ and $Y_1=Y_2$, which condition can be reconciled with the condition $R_2<R_1$, only if the lines of force bend on crossing the boundary, thus making the Z component different in the two media, and this relation must be determined.

If the lines bend on crossing the boundary their density (*i.e.* the number crossing unit area in a normal direction will change from N_1 to N_2).

Let the angle which the lines make with the normal in the upper medium be θ_1, and the angle which the refracted lines make with the normal be θ_2, as shown in Fig. 256. We will now determine the value of the z component on the two sides of the boundary. In the upper medium the normal component of the electric force is $Z_1 = R_1 \cos \theta_1$; in the lower medium, $Z_2 = R_2 \cos \theta_2$. The density of the lines of force parallel to the z axis (normal polarization) for the two media is given by

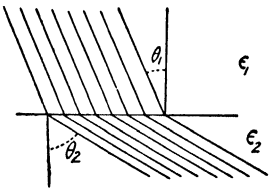


FIG. 256

$$\frac{\epsilon_1 R_1 \cos \theta_1}{4\pi} \text{ and } \frac{\epsilon_2 R_2 \cos \theta_2}{4\pi}.$$

Now the normal polarization is the same in the two media, for the same number of lines pass through a plane of unit area which is perpendicular to the z axis, in whichever medium we consider the plane, consequently we can write

$$\frac{\epsilon_1 R_1 \cos \theta_1}{4\pi} = \frac{\epsilon_2 R_2 \cos \theta_2}{4\pi} \text{ or } \epsilon_1 Z_1 = \epsilon_2 Z_2,$$

which expression determines the boundary conditions of the z component.

In a similar way it may be shown that the boundary conditions for the components of the magnetic force are $\alpha_1 = \alpha_2$, $\beta_1 = \beta_2$, $\mu_1\gamma_1 = \mu_2\gamma_2$. Since, however, $\mu = 1$ in practically all optical problems, we can write $\gamma_1 = \gamma_2$.

Velocity of the Wave. — To find the velocity of the wave we differentiate the first equation of (9) with respect to t,

$$\frac{\epsilon}{c}\frac{\partial X}{\partial t} = \frac{\partial \gamma}{\partial y} - \frac{\partial \beta}{\partial z}, \quad \frac{\epsilon}{c}\frac{\partial^2 X}{\partial t^2} = \frac{\partial}{\partial y}\left(\frac{\partial \gamma}{\partial t}\right) - \frac{\partial}{\partial z}\left(\frac{\partial \beta}{\partial t}\right),$$

and substituting for $\partial\gamma/\partial t$ and $\partial\beta/\partial t$ the values given by the last two equations of (9),

$$\begin{aligned}\frac{\epsilon}{c^2}\frac{\partial^2 X}{\partial t^2} &= \frac{\partial}{\partial y}\left(\frac{\partial X}{\partial y} - \frac{\partial Y}{\partial x}\right) - \frac{\partial}{\partial z}\left(\frac{\partial Z}{\partial x} - \frac{\partial X}{\partial z}\right) \\ &= \frac{\partial^2 X}{\partial x^2} + \frac{\partial^2 X}{\partial y^2} + \frac{\partial^2 X}{\partial z^2} - \frac{\partial}{\partial x}\left(\frac{\partial X}{\partial x} + \frac{\partial Y}{\partial y} + \frac{\partial Z}{\partial z}\right). \qquad (10)\end{aligned}$$

We now differentiate the first three equations of (9) with respect to x, y and z respectively,

$$\frac{\epsilon}{c}\frac{\partial^2 X}{\partial t \partial x} = \frac{\partial^2 \gamma}{\partial y \partial x} - \frac{\partial^2 \beta}{\partial z \partial x}, \quad \frac{\epsilon}{c}\frac{\partial^2 Y}{\partial t \partial y} = \frac{\partial^2 \alpha}{\partial z \partial y} - \frac{\partial^2 \gamma}{\partial x \partial y}, \quad \frac{\epsilon}{c}\frac{\partial^2 Z}{\partial t \partial y} = \frac{\partial^2 \beta}{\partial x \partial z} - \frac{\partial^2 \alpha}{\partial y \partial z}.$$

Addition of these three equations gives

$$\frac{\epsilon}{c}\left(\frac{\partial^2 X}{\partial t \partial x}+\frac{\partial^2 Y}{\partial t \partial y}+\frac{\partial^2 Z}{\partial t \partial y}\right)=0 \text{ or } \frac{\partial}{\partial t}\left(\frac{\partial X}{\partial x}+\frac{\partial Y}{\partial y}+\frac{\partial Z}{\partial z}\right)=0.$$

Since we are dealing with periodic changes of the electric force, the differential coefficient, with respect to the time, of the quantity in the parenthesis, can be considered as proportional to the quantity, with a phase increase of $\pi/2$ (since differentiating the sine gives the cos, the equivalent of a phase change of $\pi/2$).

This gives us
$$\left(\frac{\partial X}{\partial x}+\frac{\partial Y}{\partial y}+\frac{\partial Z}{\partial z}\right)=0. \tag{11}$$

If there are no charges, $\rho=0$ (see 9*a*) and if the medium is homogeneous ϵ does not depend on the coördinates it follows from 9*a* that

$$\frac{\epsilon}{c^2}\frac{\partial^2 X}{\partial t^2}=\frac{\partial^2 X}{\partial x^2}+\frac{\partial^2 X}{\partial y^2}+\frac{\partial^2 X}{\partial z^2}=\Delta X. \tag{12}$$

Similar equations hold for Y, Z, α and β,

$$\frac{\epsilon}{c^2}\frac{\partial^2 X}{\partial t^2}=\Delta X,\quad \frac{\epsilon}{c^2}\frac{\partial^2 Y}{\partial t^2}=\Delta Y,\quad \frac{\epsilon}{c^2}\frac{\partial^2 Z}{\partial t^2}=\Delta Z,$$

$$\frac{\epsilon}{c^2}\frac{\partial^2 \alpha}{\partial t^2}=\Delta \alpha,\quad \frac{\epsilon}{c^2}\frac{\partial^2 \beta}{\partial t^2}=\Delta \beta,\quad \frac{\epsilon}{c^2}\frac{\partial^2 \gamma}{\partial t^2}=\Delta \gamma.$$

Differential equations of this form represent waves travelling with a velocity $v=c/\sqrt{\epsilon}$.

Now the dielectric constant of the ether equals unity, consequently our equation shows us that the velocity of the wave in space is equal to c, the ratio of the two systems of electrical units. This has been confirmed by experiment, the velocity of light determined by optical methods being $2.9989.10^{10}$ cms. per sec., while the value of c determined by electrical methods is 3.10^{10} cm./sec.

In ponderable media the velocity must be smaller in the ratio $1/\sqrt{\epsilon}$. Now the refractive index of a medium is the ratio of the velocity of light in free space to the velocity in the medium, consequently we have

$$n=\sqrt{\epsilon}, \tag{13}$$

or the square of the refractive index equals the dielectric constant. As a matter of fact this result is not confirmed by experiment except in a few cases, for the reason that the dielectric constant is in reality a function of the period of the vibration, or in other

words of the wave-length, and the dielectric constant determined by all electrical methods is the value of n^2 for infinitely long waves, consequently our expression does not hold for the very short waves of light. In the case of some of the gases, in which the dispersion is small, we have fair agreement, as is shown by the following table:

	n	$\sqrt{\epsilon}$
Air,	1.000294	1.000295
CO_2,	1.000449	1.000473
H_2,	1.000138	1.000132

As examples of substances which show especially marked deviations, we may take water and ethyl alcohol:

	n	$\sqrt{\epsilon}$
Water,	1.33	9.0
Alcohol,	1.36	5.0

The cause of these discrepancies we shall see when we come to the study of dispersion.

Deduction of the Laws of Reflection and Refraction for Transparent Media. — If a ray of light is incident upon the boundary separating two media of different optical densities, the percentage reflected and refracted will depend, not only upon the relative refractive indices of the two media, and the angle of incidence, but also upon the nature of the incident vibration as regards its state of polarization.

A plane-wave is represented by the equations

$$\left.\begin{aligned} X &= A_x \cos \frac{2\pi}{T}\left(t - \frac{mx+ny+pz}{V}\right) \\ Y &= A_y \quad " \quad " \quad " \\ Z &= A_z \quad " \quad " \quad " \end{aligned}\right\}, \tag{14}$$

in which X, Y and Z are the values of the components of the electric force at any time and any point in space, A_x, A_y and A_z the maximum values of the components, T the periodic time of the disturbance, V the velocity of the wave, and m, n and p the direction cosines of the normal to the wave-front.

We will now deduce the laws which govern the directions of the reflected and refracted rays, and their relative intensities and states of polarization.

As it is difficult to represent the various components along x, y and z of the incident, reflected and refracted light in two dimensions a three-dimensional model was made from which the

stereoscopic diagram shown in Fig. 257 was prepared by photography.[1]

Resolve the incident electric force into two components along y and x, E_s perpendicular to the plane of incidence and E_p parallel to it. We therefore write for the y component of the electric force, since

$$m = \sin\Phi,\ n = 0,\ \text{and}\ p = \cos\Phi,$$

$$Y_e = E_s \cos\frac{2\pi}{T}\left(t - \frac{x\sin\Phi + z\cos\Phi}{V_1}\right), \tag{15}$$

$$\left.\begin{aligned} X_e &= E_p \cos\Phi \cdot \cos\frac{2\pi}{T}\left(t - \frac{x\sin\Phi + z\cos\Phi}{V_1}\right) \\ Z_e &= -E_p \sin\Phi \quad " \quad " \quad " \end{aligned}\right\}$$

in which V_1 is the velocity of the light in air.

In this expression for Y_e note that the amplitude along the y axis E_s is the y component of the incident light. This is not the case for the amplitudes along the x and z axes, for E_p is parallel to neither and must therefore be resolved into components $E_p \cos\Phi$ along x and $-E_p \sin\Phi$ along z, the minus sign occurring since the direction of E_p is upwards from the boundary. Note that we now have amplitudes which depend on the angle of incidence, the first step in getting an expression for reflected and refracted amplitudes for all angles of incidence and states of polarization.

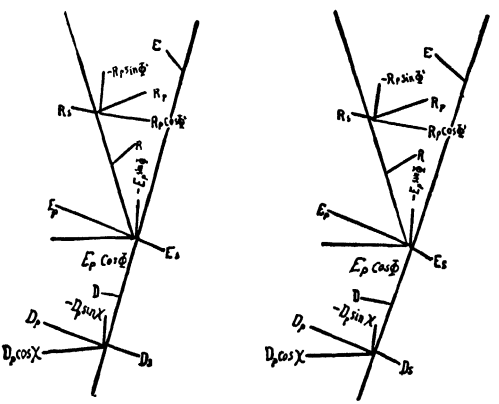


FIG. 257

We use X_e, Y_e and Z_e instead of X_1, Y_1 and Z_1, since the forces in the upper medium are not alone those in the incident wave. We have forces X_r, Y_r and Z_r contributed by the reflected wave, and X_1 is the sum of X_e and X_r.

[1] This, and other stereoscopic diagrams in the book can be seen in stereoscopic relief without a stereoscope with a little practice. Fix the eyes on an object at some distance and quickly slide the diagram into the line of sight. Three diagrams will then appear, and the attention is to be directed to the centre one focussing the eyes for the shorter range while holding the central picture single.

The magnetic forces associated with these are obtained at once by differentiating the above and substituting in equations (9). For example, the x component α is given by

$$\frac{1}{c}\frac{\partial \alpha}{\partial t}=\frac{\partial Y}{\partial z}-\frac{\partial Z}{\partial y}=E_s \sin \frac{2\pi}{T}\left(t-\frac{x \sin \Phi+x \cos \Phi}{V_1}\right)\frac{2\pi}{t}\frac{\cos \Phi}{V_1},$$

$$\frac{\partial \alpha}{\partial t}=cE_s\frac{2\pi}{T}\frac{\cos \Phi}{V_1} \sin \frac{2\pi}{T} \quad " \quad "$$

$$\alpha_s=-cE_s\frac{\cos \Phi}{V_1}\cos\frac{2\pi}{T}\left(t-\frac{x \sin \Phi+z \cos \Phi}{V_1}\right),$$

and since $V_1=\frac{c}{\sqrt{\epsilon_1}}$

$$\left.\begin{aligned} \alpha_e&=-E_s \cos \Phi\sqrt{\epsilon_1} \cos \frac{2\pi}{T}\left(t-\frac{x \sin \Phi+z \cos \Phi}{V_1}\right)\\ \text{Similarly, } \beta_e&=E_p\sqrt{\epsilon_1} \quad \cos \frac{2\pi}{T} \quad " \quad "\\ \gamma_e&=E_s\sqrt{\epsilon_1} \quad \cos \frac{2\pi}{T} \quad " \quad " \end{aligned}\right\}. \quad (16)$$

Writing for the refracted wave,

$$\left.\begin{aligned} X_2&=\mathrm{D}_p \cos \chi \cos \frac{2\pi}{T}\left(t-\frac{x \sin \chi+z \cos \chi}{V_2}\right)\\ Y_2&=D_s \quad \cos \frac{2\pi}{T} \quad " \quad "\\ Z_2&=-\mathrm{D}_p \sin \chi \cos \frac{2\pi}{T} \quad " \quad " \end{aligned}\right\} \quad (17)$$

in which D_p and D_s are the components of amplitude parallel and perpendicular to the plane incidence as shown in Fig. 257.

The magnetic components of the refracted wave are obtained in the same way as those of the incident.

They are given by

$$\left.\begin{aligned} \alpha_2&=-D_s \cos \chi\sqrt{\epsilon_2} \cos \frac{2\pi}{T}\left(t-\frac{x \sin \chi+z \cos \chi}{V_2}\right)\\ \beta_2&=D_p\sqrt{\epsilon_2} \quad " \quad " \quad "\\ \gamma_2&=D_s \sin \chi\sqrt{\epsilon_2} \quad " \quad " \quad " \end{aligned}\right\}. \quad (18)$$

If now the boundary conditions are to be complied with, there will be a reflected wave except when $\sqrt{\epsilon_1}=\sqrt{\epsilon_2}$. Let us take the simplest possible case of a plane-polarized vibration parallel to the x axis at normal incidence.

At the boundary we have

$$\left.\begin{aligned} &X_1 = X_2, \text{ that is, } E_p = D_p, \text{ since } \Phi = \chi = 0 \\ &\beta_1 = \beta_2 \quad \text{or} \quad E_p\sqrt{\epsilon_1} = D_p\sqrt{\epsilon_2} \\ &\qquad\qquad \sqrt{\epsilon_1} = \sqrt{\epsilon_2} \end{aligned}\right\}. \tag{19}$$

The boundary conditions will hold for the incident and refracted wave only under the above condition, that is, when the refractive indices of the two media are the same.

If ϵ_2 in the lower medium is greater than ϵ_1 in the upper, we shall have a reflected wave, and the sum of the forces of the incident and reflected wave constitute the force at the boundary in the upper medium, which is to be equated to the lesser force in the lower medium. The direction of the force in the reflected wave is opposite to that in the incident, for as the reflecting power increases, the force in the lower medium must diminish.

We now write for the electric and magnetic components of the reflected wave, writing $\Phi' =$ angle of reflection

$$\left.\begin{aligned} X_r &= R_p \cos\Phi' \cos\frac{2\pi}{T}\left(t - \frac{x\sin\Phi' + z\cos\Phi'}{V_1}\right) \\ Y_r &= R_s \cos\frac{2\pi}{T} \qquad\quad \text{“} \qquad\quad \text{“} \\ Z_r &= +R_p \sin\Phi' \cos\frac{2\pi}{T} \quad \text{“} \qquad\quad \text{“} \end{aligned}\right\}, \tag{20}$$

and

$$\left.\begin{aligned} \alpha_r &= -R_s \cos\Phi'\sqrt{\epsilon_1} \cos\frac{2\pi}{T}\left(t - \frac{x\sin\Phi' + z\cos\Phi'}{V_1}\right) \\ \beta_r &= R_p\sqrt{\epsilon_1} \qquad\quad \cos\frac{2\pi}{T} \quad \text{“} \qquad\quad \text{“} \\ \gamma_r &= R_s \sin\Phi'\sqrt{\epsilon_1} \quad \cos\frac{2\pi}{T} \quad \text{“} \qquad\quad \text{“} \end{aligned}\right\}. \tag{21}$$

From these equations we can deduce the laws of reflection and refraction, as well as the relation between the intensities of the reflected and refracted rays for various states of polarization.

The relations between the angles of incidence, reflection and refraction follow at once from the boundary conditions, which are only fulfilled when for $z = 0$ we put all of the forces proportional to the same function of t, x and y.

This gives us $$\frac{\sin\Phi}{V_1} = \frac{\sin\Phi'}{V_1} = \frac{\sin\chi}{V_2},$$

or $$\frac{\sin\Phi}{\sin\chi} = \frac{V_1}{V_2} = n, \text{ the refractive index.} \tag{22}$$

We will now deduce expressions for the intensities of the reflected and refracted components, the relations between the phases of the vibrations, their dependence upon the angle of incidence and the state of polarization. The force X_1, on the upper side of the boundary, is equal to the sum of the forces in the incident and reflected waves, $X_e + X_r$, which is to be equated to the force on the lower side of the boundary.

$$\left.\begin{array}{lll} (1)\ X_e+X_r=X_2 \text{ or} & (E_p-R_p)\cos\Phi=D_p\cos\chi \\ (2)\ Y_e+Y_r=Y_2 \text{ or} & E_s+R_s=D_s \\ (3)\ \alpha_e+\alpha_r=\alpha_2 \text{ or} & (E_s-R_s)\sqrt{\epsilon_1}\cos\phi=D_s\sqrt{\epsilon_2}\cos\chi \\ (4)\ \beta_e+\beta_r=\beta_2 \text{ or} & (E_p+R_p)\sqrt{\epsilon_1}=D_p\sqrt{\epsilon_2} \end{array}\right\} \cdot (23)$$

The positive directions of the components R_p and D_p are shown in Fig. 258.

$$\left.\begin{array}{l} \text{Add (2) and (3)}, \qquad 2E_s=D_s\left(1+\dfrac{\sqrt{\epsilon_2}\cos\chi}{\sqrt{\epsilon_1}\cos\Phi}\right) \\ \text{Elim. } D_s \text{ from (2) and (3)}, \\ \qquad E_s\left(\dfrac{\sqrt{\epsilon_1}\cos\Phi}{\sqrt{\epsilon_2}\cos\chi}-1\right)=R_s\left(\dfrac{\sqrt{\epsilon_1}\cos\Phi}{\sqrt{\epsilon_2}\cos\chi}+1\right) \\ \text{Add (1) and (4)}, \qquad 2E_p=D_p\left(\dfrac{\cos\chi}{\cos\Phi}+\dfrac{\sqrt{\epsilon_2}}{\sqrt{\epsilon_1}}\right) \\ \text{Elim. } D_p \text{ from (1) and (4)}, \\ \qquad E_p\left(\dfrac{\cos\Phi}{\cos\chi}-\dfrac{\sqrt{\epsilon_1}}{\sqrt{\epsilon_2}}\right)=R_p\left(\dfrac{\cos\Phi}{\cos\chi}+\dfrac{\sqrt{\epsilon_1}}{\sqrt{\epsilon_2}}\right) \end{array}\right\} \cdot (24)$$

Substitute for $\sqrt{\epsilon_2}/\sqrt{\epsilon_1}=n$ its equivalent $\sin\Phi/\sin\chi$ and we obtain equations identical with the formulae of Fresnel, from which the phrases and intensities of the reflected and refracted waves can be calculated.

$$\left.\begin{array}{l} \text{Reflected amplitudes:} \\ \qquad R_s=-E_s\dfrac{\sin(\Phi-\chi)}{\sin(\Phi+\chi)}, \quad R_p=E_p\dfrac{\tan(\Phi-\chi)}{\tan(\Phi+\chi)}. \\ \text{Refracted amplitudes:} \\ \qquad D_s=E_s\dfrac{2\sin\chi\cos\Phi}{\sin(\Phi+\chi)}, \quad D_p=E_p\dfrac{2\sin\chi\cos\Phi}{\sin(\Phi+\chi)\cos(\Phi-\chi)} \end{array}\right\} \cdot (25)$$

Perpendicular Incidence. — The formula for the reflecting power at normal incidence is the most important, and repeated use of it will be made in the following chapter. It will therefore be derived first. The above equations (25) cannot be used in this case, for $\Phi=\chi=0$ and the expressions become indeterminate.

Substituting in equations (24) for $\sqrt{\epsilon_2}/\sqrt{\epsilon_1}$ the refractive index n, we have, since $\cos \Phi/\cos \chi = 1/1$,

$$R_s\left(\frac{1+n}{n}\right) = E_s\left(\frac{1-n}{n}\right),$$

$$R_s = E_s\left(\frac{1-n}{1+n}\right) = -E_s\left(\frac{n-1}{n+1}\right),$$

$$R_p = E_p\left(\frac{n-1}{n+1}\right).$$

From the first equation it is clear that if $n>1$, the reflected electric amplitude is oppositely directed from the incident, since the direction of the vector depends on its sign. The second equation shows the same condition, for when $\Phi=0$, similar signs mean opposite directions, as will be seen by referring to Fig. 258. The presence of the reflected wave will thus reduce the amplitude of the incident wave *at the reflecting boundary*, and, if the intensity of the reflected wave is equal to that of the incident, the amplitude will be reduced to zero. This was found to be the case by Wiener in his experiments upon stationary light-waves, the node occurring at the reflecting surface. These experiments were fully described in the Chapter on Interference. The opposite condition will be found for the magnetic vectors, which are similarly directed in the incident and reflected waves. They will therefore add their effects at the boundary.

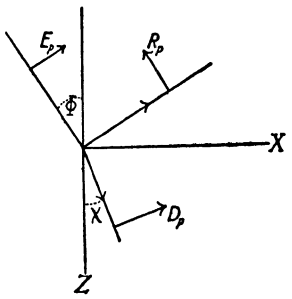


FIG. 258

It must be remembered that the above formulae express the *amplitudes* of the reflected vibrations; the *intensity* of the reflected light, if the intensity of the incident light is I, is given by

$$R_s^2 = I\frac{(n-1)^2}{(n+1)^2}$$

in the case of a medium in air, or for a medium of refractive index n_2 immersed in a medium of index n_1

$$R = \frac{(n_2-n_1)^2}{(n_2+n_1)^2}$$

in which R is the reflection coefficient or fractional part of the incident light which is reflected.

It is clear that we have thus a means of determining the refractive

index by measuring the reflecting power of a substance. This method was employed by Rubens for the determination of n in the infra-red, as will be described in the Chapter on Dispersion Theory.

Oblique Incidence. — We will now take the more general case. It is evident that the component perpendicular to the plane of incidence in the reflected light never vanishes, whatever be the values of Φ and χ in the formula for R_s.

It is different, however, in the case of the formula for R_p, the parallel component. As we increase the angle of incidence from 0, it is evident that we shall eventually reach a point at which the reflected and refracted rays are at right angles, for the angle between them is greater than 90° near perpendicular incidence, and less than 90° at grazing incidence. At the angle in question it is evident that $(\Phi+\chi)=90°$ and $\tan(\Phi+\chi)=\infty$, that is, $R_p=0$, or the component parallel to the plane of incidence is wholly absent. This means simply that the reflected light is plane-polarized, and the angle in question is known as the angle of polarization or the Brewsterian angle.

The refractive index $n=\sin\Phi/\sin\chi$, consequently if Φ' be the angle of polarization, we have

$$\sin\chi'=\sin\left(\frac{\pi}{2}-\Phi'\right)=\cos\Phi' \text{ and } \frac{\sin\Phi'}{\cos\Phi}=\tan\Phi'=n,$$

a relation which has been fully discussed in the Chapter on Polarization.

Polarizing Power of Oblique Plates. — A method has been given, in the Chapter on Polarization, for the determination of the percentage of polarization, by compensation with inclined plates. It is important, therefore, to develop the formula for determining the polarization produced by one or more glass plates at any given angle. For the ratio of the amplitudes transmitted through a single surface we have from (25)

$$\frac{D_s}{D_p}=\frac{E_s}{E_p}\cos(\Phi-\chi),$$

or for two surfaces (*i.e.* a plate),

$$\frac{D_s}{D_p}=\frac{E_s}{E_p}\cos^2(\Phi-\chi).$$

If $E_s=E_p$, the condition equivalent to natural or unpolarized light, we have $D_s<D_p$, that is, the transmitted light is partially polarized, the plane of the vibration which is in excess being in the plane of incidence. If the incident vibration makes an angle of

45° with the plane of incidence, $E_s = E_p$ and the transmitted light is plane-polarized, the plane, however, being more or less rotated towards the plane of incidence, on account of the fact that the component D_p is larger than the component D_s. The percentage of polarization for a single plate increases with the angle of incidence. At the polarizing angle, $\tan \Phi = n$ and $\Phi + \chi = 90°$; therefore, if $E_s = E_p$, we have

$$\frac{D_s}{D_p} = \sin^2 2\Phi = \frac{4n^2}{(1+n^2)^2}.$$

The polarization is not a maximum, however, at this angle, but increases with the angle of incidence, in marked contrast to polarization by reflection. The case of polarization by transmission through one or more glass plates has been treated in the Chapter on Polarization.

Change of Phase. — We have seen that in the case of reflection from a rare to a denser medium the directions of the electric forces in the incident and reflected light are opposed, which means a phase change of 180°. In the reverse case, in which the force is greater in the second medium (lesser value of n) the forces of the incident and reflected wave are in the same direction, their sum equalling the force on the other side of the boundary, which means no phase change. This is illustrated by the Newton black spot of soap-films.

Change of Phase by Passage through the Polarizing Angle. — The formula for the reflected amplitude, parallel to the plane of incidence, shows us that the phase changes suddenly by 180° on passage through the angle of polarization: for $(\Phi + \chi)$ is obtuse or acute according as the angle of incidence is less or greater than the polarizing angle. Suppose now that the incident light is plane-polarized at an azimuth of 45°. At the polarizing angle the component of the vibration, which is parallel to the surface, will be the only one reflected, and it can be completely quenched by means of a Nicol prism held with its short diagonal vertical. On either side of the polarizing angle we shall have a reflected component perpendicular to the other, but the directions of the vectors will be opposite on opposite sides of the angle. The resultant will be in each case a plane-polarized vibration, which will, however, be turned slightly towards the plane of incidence, the direction of the rotation from the plane parallel to the surface being opposite in the two cases. This will be readily understood by drawing the horizontal component, and compounding it first with a small vertical component directed upwards, and then with one directed downwards, the change of direction corresponding to the phase change of 180°.

It was found by Jamin and others that in the majority of cases the light was not completely polarized by reflection at the Brewsterian angle. Moreover, if the incident light was polarized, and reflection occurred in the neighborhood of this angle, the reflected light, instead of being plane-polarized, as the formulae indicate, showed traces of elliptical polarization. This indicates that the phase change, instead of occurring abruptly at the polarizing angle, enters by degrees; Drude observed in 1889 that the elliptical polarization produced by a freshly split surface of rock salt was very small, but that it increased rapidly on the exposure of the surface to the air. Shortly afterward Lord Rayleigh found that the ellipticity produced by reflection from water could be completely eliminated by removing the surface film of grease, which is always present unless special precautions are taken.

These experiments indicate that the disagreement with the formulae is caused by surface films having optical properties different from those of the body of the substance. We will now take up the investigation of the effects of these films, and the calculation of their probable thickness.

Elliptical Polarization. Surface Films or Strains? — The theory of reflection applied to boundaries between media of different optical densities has led us to the conclusion that plane-polarized light should always be reflected as plane-polarized light. As a matter of fact, we find that this is seldom the case. If light incident on a metallic mirror is polarized at an angle of 45° with the plane of incidence, almost no change in intensity is seen when the reflected light is examined through a slowly revolving Nicol, while even in the case of transparent substances it is seldom possible to completely extinguish the reflected light with the Nicol in any position. The cause of this has been supposed to be the almost universal presence of a so-called "surface-layer" within which the optical density changes gradually from that of the upper medium to that of the lower. In the previous treatment we considered that the change at the boundary was abrupt, and deduced our boundary conditions on this assumption.

The electromagnetic theory of the action of surface films was first developed by van Ryn van Alkemade [1] and by Drude.[2]

If ϵ_1 is the dielectric constant of the first medium, and ϵ_2 that of the second (or reflecting substance), and ρ is the amplitude ratio, the equations show that ρ has a positive sign when $\epsilon_2 > \epsilon_1$, and the dielectric constant ϵ of the film has an intermediate value. If the incident light is considered as coming towards us and the plane of

[1] *Wied. Ann.*, *20*, 22, 1883.
[2] *Wied. Ann.*, *36*, 865, 1889.

vibration makes an angle of 45° with the vertical (the rotation away from the vertical being clockwise), the direction of the elliptical vibration will be clockwise for positive values of $\bar{\rho}$ and counterclockwise for negative values, a change of sign of the amplitude ratio, amounting to the same thing as a phase-difference of 180° between the components. We can easily determine the value of $\bar{\rho}$ experimentally by either of the methods given in the section on Elliptical Polarization. In the case of reflection at a glass surface in air it has a value not far from .007, though for flint glass with a high refractive index the value may be as high as .03.

Negative ellipticity occurs in the case of water and certain solids with very low refractive indices, such as fluor spar, in which case the surface film has a refractive index higher than n_2. These relations are inverted when the reflection is from a dense to a rarer medium.

In the case of water Lord Rayleigh has shown that the ellipticity is due to a very thin film of grease, which naturally has a higher refractive index than water. In the case of solids, a higher refractive index of the surface film may perhaps be explained by some action of the polishing material upon the surface.

By carefully cleaning the surface of the water, Lord Rayleigh was able to destroy almost completely all traces of elliptical polarization, the value of ρ being not more than .00035. He also found that the positive ellipticity of glass can be changed into negative by repolishing the surface.

Glycerine gives no trace of elliptical polarization, and is supposed to have a surface film of higher refractive index, rather than no film at all.

Drude found that freshly cleaned crystal surfaces showed no traces of elliptical polarization, but that it appeared after the surfaces had been exposed to the air for some time, owing to the formation of surface films. Assuming that the dielectric constant is uniform throughout the film, he calculated the thickness of the film necessary to produce a given axis ratio $\bar{\rho}$. For glass of refractive index $n=1.5$ and $\bar{\rho}=.007$ he found that the thickness of the film necessary to account for the ellipticity is less than $\frac{1}{200}$ of the wavelength of the light.

Lummer and Sorge found, in repeating and extending the work of Lord Rayleigh, that the elliptical polarization of the light reflected from the surface of a glass prism was altered by polishing or rubbing one of the other faces, or even by subjecting the prism to pressure applied to the bases.

Their results were probably due to other factors than the alteration of surface films in the manner supposed, and they have never been confirmed.

Total Reflection. — We have seen in the Chapter on Refraction that, in the case of a ray of light passing from a medium of high to one of low refractive index, the refracted ray vanishes for incidence angles greater than a certain value, the energy being totally reflected. In this case the refractive index of the medium on which the light is incident (say air) with respect to the dense medium (say water) is less than unity. In the case of air and water the wave-velocities are as 100 : 70 which gives for water with respect to air $n = {}^{100}\!/_{70} = 1.43$ and for air with respect to water $n = {}^{70}\!/_{100} = 0.7$. We have the formula $\sin \Phi/n = \sin \chi$ and if $\sin \Phi/n > 1$, which will occur for all values of Φ greater than 44° 30′ (for which the sine is 0.7009), χ is no longer real, *i.e.* there is no refracted ray. We can study the nature of the reflected light, however, by substituting $\sin \Phi/n$ for $\sin \chi$ in the equations (24).

$$\therefore \cos \chi = \sqrt{1 - \frac{\sin^2 \Phi}{n^2}},$$

which quantity is imaginary if $\sin \Phi > n$.

We can write this in the form

$$\cos \chi = -i\sqrt{\frac{\sin^2 \Phi}{n^2} - 1}. \tag{26}$$

Substitution of the above in equation (24) gives us reflected light with a complex amplitude, which can be interpreted as a change of phase amplitude, since if we write $Re^{id} \cdot e^{2\pi i(t/T - \chi/\lambda)} = Re^{i\{2\pi(t/T - \chi/\lambda) + \delta\}}$ which is equivalent to $R \cos 2\pi(t/T - \chi/\lambda) + \delta$ in which δ represents a change of phase.

This change of phase, if it is different for the two components of the incident vibration, will result in the transformation of a linear vibration into an elliptical one, and as total reflection is one of the methods commonly employed to produce elliptically and circularly polarized light, the subject is of some importance. To calculate this phase change we write as before $R_p e^{i\delta_p}$ and $R_s e^{i\delta_s}$ for the components of the reflected amplitude, and obtain, since $\sqrt{\epsilon_2}/\sqrt{\epsilon_1} = n$, by substitution in (24)

$$E_s\left(\frac{i \cos \Phi}{\sqrt{\sin^2 \Phi - n^2}} - 1\right) = R_s e^{i\delta_s}\left(\frac{i \cos \Phi}{\sqrt{\sin^2 \Phi - n^2}} + 1\right),$$

$$E_v\left(\frac{i \cos \Phi \cdot n}{\sqrt{\sin^2 \Phi - n^2}} - \frac{1}{n}\right) = R_p e^{i\delta_p}\left(\frac{i \cos \Phi \cdot n}{\sqrt{\sin^2 \Phi - n^2}} + \frac{1}{n}\right).$$

If we multiply these equations by their complex conjugates, obtained by writing in them $-i$ for i, we find that $E_s^2 = R_s^2$ and

$E_p^2 = R_p^2$, *i.e.* the intensities of the reflected components are equal to those of the incident.

Suppose now that our incident light is plane-polarized vibrating in azimuth 45°. In this case $E_s = E_p$ and $R_s = R_p$, and if we substitute these values in the above equations and divide, we get

$$\frac{i\cos\Phi - \sqrt{\sin^2\Phi - n^2}}{i\cos\Phi \cdot n - \frac{1}{n}\sqrt{\sin^2\Phi - n^2}} = e^{i(\delta_s - \delta_p)} \frac{i\cos\Phi + \sqrt{\sin^2\Phi - n^2}}{i\cos\Phi \cdot n + \frac{1}{n}\sqrt{\sin^2\Phi - n^2}};$$

or

$$e^{i(\delta_p - \delta_s)} = e^{i\Delta} = \frac{\sin^2\Phi + i\cos\Phi\sqrt{\sin^2\Phi - n^2}}{\sin^2\Phi - i\cos\Phi\sqrt{\sin^2\Phi - n^2}};$$

and

$$\frac{1 - e^{i\Delta}}{1 + e^{i\Delta}} = \frac{-i\cos\Phi\sqrt{\sin^2\Phi - n^2}}{\sin^2\Phi}.$$

Multiplying this by its complex conjugate gives

$$\frac{1 - \cos\Delta}{1 + \cos\Delta} = \left\{ \frac{\cos\Phi\sqrt{\sin^2\Phi - n^2}}{\sin^2\Phi} \right\}^2,$$

since

$$e^{i\Delta} + e^{-i\Delta} = 2\cos\Delta;$$

therefore

$$\tan\tfrac{1}{2}\Delta = \frac{\cos\Phi\sqrt{\sin^2\Phi - n^2}}{\sin^2\Phi}. \qquad (27)$$

This expression shows us that the relative phase-difference Δ is zero for grazing incidence ($\Phi = \pi/2$), and also at the critical angle ($\sin\Phi = n$), in which n is the relative refractive index. If the denser medium has a refractive index 1.51 and the reflection occurs at an air surface, n in our equations will be 1/1.51.

To find the value of Φ which will give Δ its maximum value, we differentiate the last equation with respect to Φ, and obtain

$$\frac{1}{2\cos^2\frac{1}{2}\Delta}\frac{\partial\Delta}{\partial\Phi} = \frac{2n^2 - \sin^2\Phi(1+n^2)}{\sin^3\Phi\sqrt{\sin^2\Phi - n^2}}$$

and the maximum value of Δ is obtained at an incidence angle Φ' determined by

$$\sin^2\Phi' = \frac{2n^2}{1+n^2}.$$

The corresponding value of Δ is given by

$$\tan\tfrac{1}{2}\Delta' = \frac{1-n^2}{2n}.$$

For glass of refractive index 1.51, $\Phi' = 51°\ 20'$ and $\Delta' = 45°\ 36'$. A value of 45° occurs at incidence angles 48° 37′ and 54° 37′. Two

reflections at this angle will give us $\Delta = 90°$, and circularly polarized light will result if the incident light was polarized in azimuth 45°. This is accomplished by the Fresnel rhomb described in the section on Elliptical Polarization.

Penetration of the Disturbance into the Second Medium. — If we apply the equations (24), which express the relation between the incident and the refracted amplitudes, to the case of total reflection, we reach the somewhat astonishing conclusion that the refracted amplitude is not zero, which appears to be inconsistent with the *total* reflection of the energy. The case is a peculiar one, for although D may have a large value close to the boundary, it becomes zero at a distance of a few wave-lengths, the energy being entirely thrown back into the first medium. This decrease in the amplitude, as we advance from the boundary in the direction of the z axis, can be seen from equations (16) (17), which show that electric and magnetic forces in the second medium,

$$e^{i\frac{2\pi}{T}\left(t-\frac{x\sin\chi - z\cos\chi}{V_2}\right)},$$

are proportional to the real part of the complex quantity which, if we substitute for $\cos\chi$ the value given by equation (26), takes the form

$$\underline{e^{-\frac{2\pi}{TV_2}\sqrt{\frac{\sin^2\Phi}{n^2}-1}\cdot z}} \cdot e^{i\frac{2\pi}{T}\left(t-\frac{x\sin\Phi}{nV_2}\right)}. \qquad (28)$$

This formula represents a wave disturbance moving parallel to the x axis, which implies that the energy stream is along the boundary, and not down into the second medium. The amplitude, which is represented by the underscored part of (28), decreases as z increases, becoming sensibly zero when z is large in comparison to the wave-length $\lambda_2 = TV_2$. These boundary waves possess another peculiarity, in that they are not transverse, for in a transverse disturbance moving along the x axis in the second medium, X_2 must equal zero, which is not the case.

The existence of these waves can be shown experimentally by bringing a convex surface of glass of large radius of curvature into contact with the surface at which total reflection is taking place. The light will be found to enter the lens in an annular region surrounding the point of contact. This is due to the fact that the two glass surfaces come into optical contact, *i.e.* we may regard the air film as completely squeezed out. This circular patch over which total reflection fails, and complete transmission obtains, is surrounded by a ring which *transmits* light of a reddish, and *reflects* light of a bluish, tinge. The glass surfaces are not in contact here,

but the air film is too thin for total reflection to take place. Transmission will obviously occur for red light first, since the thickness of the film necessary to reflect light is measured in comparison with the wave-length. This experiment dates back to the time of Newton and Fresnel, and shows us that if the rarer medium is of extreme thinness, total reflection no longer occurs. A method employed by the author, which however is open to the objection that matter is brought into actual contact with the glass surface is to scatter minute particles on the reflecting surface; for example, smoke the hypotenuse surface of a right-angled prism *very lightly* with a flame. On sending a strong beam of light into the prism the smoked patch will be illuminated, and if viewed under a powerful microscope, each individual carbon particle will be seen to scatter light in all directions. This method was used by the author in establishing the granular nature of certain metallic films, which will be described in the Chapter on The Scattering of Light. The method was subsequently and independently originated by Cotton as a means of rendering visible ultra-microscopic particles, and was employed still later by Frölich [1] in his investigation of total reflection problems.

One of the most convincing demonstrations of the presence of the energy stream was that of Schäfer and Gross, who employed 15-cm. electric waves and a paraffine prism. They found evidence of the wave disturbance when the detector was brought up close to the surface.

The penetration of the disturbance into the rarer medium when total reflection occurs, can be shown to a fairly large audience in the following way.

One face of a large flint-glass 60° prism, four or five inches on a side, is covered with a square of plate glass with a thin layer of a solution of fluorescein between the faces. The light of an arc is concentrated by a large short-focus lens on the fluorescent film as shown in Fig. 259 at such an angle that half of the converging cone is totally reflected and the other half transmitted. Fluorescein molecules which are within a wave-length or two of the prism face are excited to luminosity by the radiation propagated in the solution parallel to the boundary. The transmitted rays excite a very powerful fluorescence, as all of the molecules contribute in this region, and we see an elliptical green field, one-half feebly luminous, the other shining with great intensity, as shown in the figure. The experiment can also be made with a small prism if a large one is not available. That only the molecules very close to the boundary are excited can be shown by varying the thickness of the fluorescent

[1] *Ann. der Phys.*, *63*, 900.

film, which remains of a constant low intensity in the region excited by the boundary wave.

An interesting relation was shown by Selenyi between the disturbance in the rarer medium in the case of total reflection, and the ability of a source of light in the rarer medium to send rays into the denser medium outside of the cone of refraction, which is the equivalent of the cone of total reflection. In Fig. 260 let A be a luminous point in air very close to the surface of the prism. The rays from A which enter the prism are confined within the cone BAC, the limiting rays AB and AC are those striking the prism at grazing incidence. If now a ray of light enters the prism along DA it will be totally reflected, but there will be a flow of luminous energy in the air along and very close to the surface, and if A is within a wave-length or two of the surface it will be illuminated. But if A can be illuminated by a ray along DA, it can, by the principle of reversibility, send light to D when it is itself luminous. We thus come to the conclusion that a small amount of light from the source A can penetrate the region outside of the cone of refraction *if it is sufficiently close to the surface.* Selenyi showed this by placing a small drop of a solution of fluorescein on the face of the prism, illuminating it with sun or arc light and viewing it from the direction DA. The author has repeated and confirmed this experiment, but the effect is easily overlooked if care is not taken, as the light is very faint. If a magnifying-glass is used to view the drop from a direction coinciding with the edge of the cone, one can find a position in which one sees half of the green disk brilliantly luminous and the other half barely visible, but distinctly green. This portion shines only by the fluorescent molecules that are within a wave-length of the glass surface. Precautions must be taken to have the background very black, by suitably placed screens.

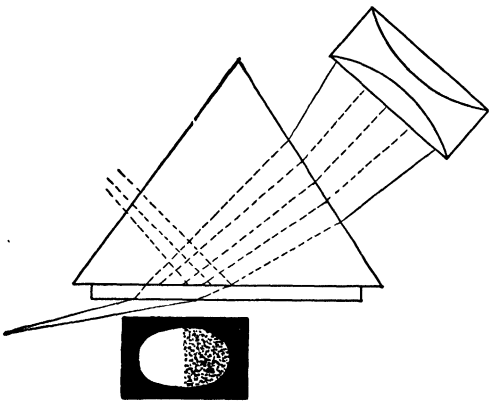


FIG. 259

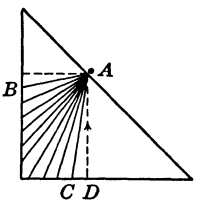
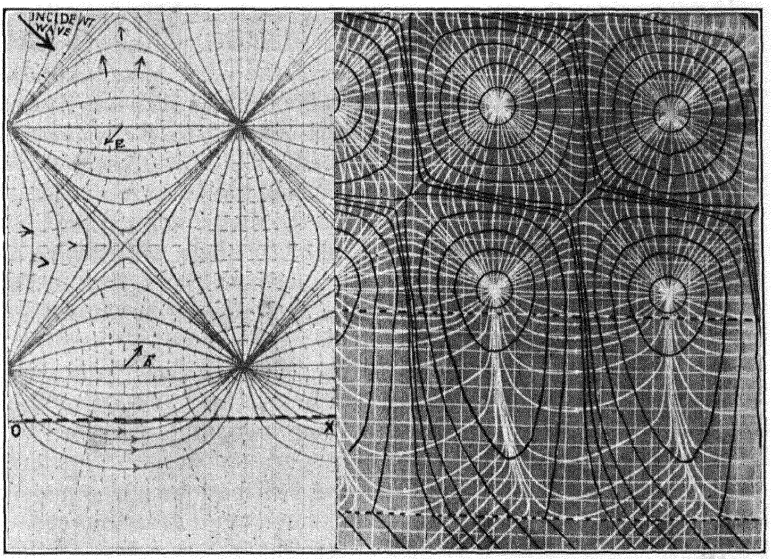


FIG. 260

Eichenwald has given a very instructive diagram illustrating the nature of the energy flow in the rarer medium when total reflection occurs.

In Fig. 261*a*, we have the momentary state of a system of plane-waves incident at 45° on the "glass-to-air" surface *ox*. The waves are not shown, only the magnetic lines of force (dotted) and the lines of energy flow (solid), which show the direction, at all parts of the interference field formed by the superposition of the incident and reflected light, of Poynting's vector, which is perpendicular to



a FIG. 261 *b*

the magnetic and electric lines of force, the latter being perpendicular to the paper. The waves are incident in the direction indicated by an arrow, the wave-length being equal to the diagonal of one of the squares which form the chess-board pattern. The dark spots where the lines representing energy flow converge are the points where the displacement currents have their maximum momentary value, while the directions of the maximum electric forces perpendicular to the paper are indicated by arrows *E*, ↙ indicating "out from" and ↗ indicating "down into" the paper. The directions of the lines of magnetic force, which surround the displacement currents are indicated also by arrow points. This diagram shows very clearly the nature of the disturbance in the rarer medium when total reflection occurs. Interference minima

equal to zero and separated by distance $(\lambda/2)\sqrt{2}$ parallel to ox, as we saw in the section on stationary light-waves in the Chapter on Interference. At the boundary ox, one sees that the energy stream lines pass out into air and then curve back through the boundary, so that as much energy flows back into the glass as flows out into the air. To follow the change with time one must consider the entire diagram as moving to the right with a velocity equal to $c/n \sin 45°$. This is equivalent to keeping the diagram fixed and letting the displacement currents and the electric forces perform their oscillations and reversals.

If we consider the diagram as moving to the right we see that, along the horizontal line $L=0$, the direction of the magnetic field will reverse, for every passage of one of the dark spots across a fixed point. This will produce a reversal of the electric field. Along the horizontal line $N=0$ the motion is parallel to the magnetic field and there is no induced electric force. This line is therefore a line of zero illumination. The magnetic field would appear to have its maximum strength halfway between the dark spots (on the line $L=0$). At any given point on the boundary there will thus be a periodic fluctuation of energy to-and-fro across the boundary, and a flow of energy in the lower medium in the y direction parallel to the boundary.

The intensity falls off very rapidly with increasing distance from the boundary, being equal to .01, .001 and .000001 of the incident intensity, at distances of one, two and three wavelengths.

A further contribution to the subject will be found in a paper by Schäfer and Gross [1] who calculated the diagram for the reflection when a second glass surface is brought up very close to the boundary. In this case a certain amount of the incident light enters the second glass surface without reflection, *i.e.* there is a small direct transmission. The momentary diagram for this condition is shown in Fig. 260*b*, the magnetic lines instead of being rounded off squares as in Fig. 260*a* are rhombs, and the outer ones of the upper tier of squares are continuous with those of the next lower tier. In this case we see that the energy flows into the second glass surface along definite lines, marked off by the interference field, this condition being for the momentary state only of course. The whole pattern moves to the right with time as in the previous case. As the distance between the glass surfaces diminishes the rhombs become more elongated, uniting to a line when the surfaces are in contact and reflection vanishes. This line is the magnetic sheet of the incident light.

[1] *Ann. der Phys.*, *32*, 648, 1910.

CHAPTER XIII

THE SCATTERING OF LIGHT

The absence of any lateral radiation from a beam of light traversing a vacuum was shown by Fresnel to depend upon the complete destructive interference of the secondary wavelets which can be considered as radiating from all points of space traversed by light. The same condition would obtain in a perfectly transparent and structureless medium but the presence of discontinuities of any sort violates the condition necessary for complete deductive interference, and a lateral scattering of the light takes place.

It was formerly supposed that such scattering occurred only as the result of small particles in suspension in a transparent medium, the blue color of the sky being referred to foreign matter in the atmosphere. Lord Rayleigh was the first to point out that a sufficient amount of scattering to fully account for the sky would be produced by the molecules of air alone, in the absence of any suspended matter, a theory that has been fully verified by experiment as we shall see presently.

We will take up first the scattering by small particles as this phenomenon was very fully investigated long before the discovery of the scattering by pure and transparent liquids and gases.

Scattering of Light by Small Particles. — If a beam of light is passed through a transparent medium containing in suspension small particles, the refractive index of which differs from that of the surrounding medium, light will be given off by the particles in all directions. If the particles are very small, the color of the scattered light is blue, and it is more or less completely plane-polarized, the direction of the electric vector being perpendicular to the direction of the incident beam and the direction of observation. If the incident beam is plane-polarized to start with, no light is scattered by the cloud of particles in directions parallel to that of the electric vector. In the case of particles of the order of magnitude of the light-waves, the amount of light scattered increases as the wave-length is decreased, which explains the preponderance of blue always observed in these cases. The subject was investigated experimentally by Tyndall with clouds precipitated by the chemical action of light upon the vapor of iodide of allyl. Whenever the particles obtained were sufficiently small, the laterally emitted

light was blue in color and polarized in a direction perpendicular to the incident beam.

With very small particles the intensity of the scattered light in directions nearly parallel to the incident beam is the same for directions "with" or "against" the beam, but for larger particles the scattered light is much more intense in the direction in which the incident light is travelling. The intensity of the scattered light increases very rapidly with the diameter of the particles, being proportional to the square of the volume or the 6th power of the diameter.

By making the particles sufficiently small we may obtain a violet of great intensity and purity. The fog formed by the condensation of sodium vapor in hydrogen gas at low pressure has been observed by the author to give a deeper color than any of the other media heretofore employed. The blue color is easily seen in tobacco smoke rising from the end of a lighted cigar. On standing, the smoke particles collect into larger aggregates and the blue color disappears. This is usually the case with smoke exhaled from the mouth. The blue color of the sky has its origin in a similar action exerted by the molecules of air.

The blue color of scattered light is easily shown by adding a few drops of dilute sulphuric acid to a very dilute solution of hyposulphite of soda in a glass jar or tank, concentrating sun or arc light at the centre with a lens. If the solution is too concentrated the precipitated sulphur scatters white light when its action is first noticeable. The proper concentrations are easily found by trial, the best results being obtained when the first trace of the cloud appears only after three or four minutes.

The deficiency of long-wave radiation in the light of the blue-sky is very strikingly shown in infra-red photographs of landscapes, made by the author in 1909. In these photographs the foliage appeared snow-white against a coal-black sky, and distant mountains, completely obscured by haze in ordinary photographs, came out clearly. Plate 7 is a reproduction of one of these early photographs, made through a ray filter consisting of a tank of dense cobalt glass filled with a solution of bichromate of potash, transmitting the spectrum region between 6900 and 7400. "Some of the photographs made with the infra-red filter showed the advantage gained in bringing out the detail of distant objects seen through atmospheric haze." [1] Aniline stained gelatine filters of similar transmission are now obtainable from the Eastman Kodak Co.

[1] R. W. Wood, "Recent Experiments with Invisible Light," *Roy. Inst. of Gt. Britain*, May 11, 1911.



PLATE 7. FOLIAGE AND BLUE-SKY PHOTOGRAPHED BY INFRA-RED RAYS.

Similar filters were extensively used during the war for airplane photography, as there is a "blue-sky" or haze between the earth's surface and a high-flying plane nearly as intense as the real sky. Quite recently mountains several hundred miles away have been photographed from airplanes by infra-red rays, and the method has been employed for "filming" night scenes in full sunlight in the moving-picture industry.

The intensity of the scattered light as a function of the wave-length, for obstacles of fixed size, small in comparison to λ, was calculated by Lord Rayleigh.[1]

Let i be the ratio of the amplitude of the incident to that of the scattered light, and V the volume of the disturbing particle. If r is the distance from the particle of a given point, the value of i at this point is shown to be proportional to $V/\lambda^2 r$, *i.e.* the amplitude varies inversely as the square, and the intensity inversely as the fourth power of the wave-length.

Observations were made of the distribution of energy in the spectrum of the light of the blue-sky by comparing its spectrum with the spectrum of direct sunlight, diffused by white paper. These values were compared with values calculated on the assumption that the intensity of the scattered light (*i.e.* blue-sky light) varied as $1/\lambda^4$. The two sets of values are given in the following table for four of the Fraunhofer lines:

C	D	b_3	F	
25	40	63	80	calculated.
25	41	71	91	observed.

Lord Rayleigh's formula for the intensity of the scattered light in a direction making an angle β with the incident ray is, if the incident light is unpolarized,

$$A^2\frac{(D'-D)^2}{D^2}(1+\cos^2\beta)\frac{m\pi V^2}{\lambda^4 r^2},$$

in which A^2 is the intensity of the incident light, D' and D the optical densities of the particles and the medium in which they are immersed, m the number of particles and λ the wave-length.

The formula shows that the intensity is twice as great in the direction from which the light comes originally as in a direction perpendicular to it.

This formula was derived for the case of small particles and we will now consider in more detail the case of a gas.

Before considering other cases of scattering, such as that exhibited by metal particles, colloidal solutions and granular films it will be well to develop the theory of scattering more fully

[1] *Phil. Mag.*, xli, 107–120, 274–279, 1871.

commencing with the simpler cases of scattering by gases and vapors.

The Scattering of Light by Gases: Blue-Sky. — The blue color of the sky was formerly thought to be due to a scattering of light by very small dust particles suspended in the air of the upper atmosphere, but Lord Rayleigh's calculations indicated that the scattering in this case was due to the actual molecules of air. Many years elapsed before this scattering of light by pure dust-free gases was detected in the laboratory by Cabannes[1] and independently and at about the same time by Lord Rayleigh's son (the present Lord Rayleigh).[2] Sunlight was focussed by a lens at the centre of a black tube, through a window in the wall, and a very blue cone of luminosity was observed through a glass window in the end of the tube, which was filled with air filtered through cotton-wool to remove dust particles (Fig. 262).

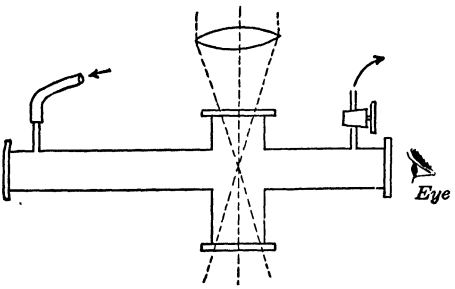


FIG. 262

Similar results were obtained with other pure gases, even with hydrogen, the scattering power of which was many times less than that of air, while that of ether vapor was much greater. Some experiments were made by the author at East Hampton,[3] New York, with a view of comparing the intensity of the light of the blue-sky with the intensity scattered in the laboratory by dust-free air. Sunlight was concentrated by a "reading-glass" lens six inches in diameter, at the centre of a black tube, similar to that of Lord Rayleigh, and the blue cone seen in the tube was matched against a bit of blue-sky reflected in a small mirror of black glass (which reflects 4%) through a very narrow slit in the rim of a large disk of black cardboard, rotating at high velocity in a dark room. In this way the intensity of the sky light was reduced to that shown in the tube. The total amount of air directly above us is the equivalent of an ocean of air of uniform density at atmospheric pressure five miles deep. This thickness of air, illuminated by normal sunlight, forms the blue-sky. In the laboratory the layer illuminated was less than one-eighth of an inch in thickness but the illumination at the focus of the lens was about 1400 times that of normal sunlight. The two could be matched by making the slit in the disk sufficiently narrow, and

[1] *C. R. 160*, 62. [2] *Proc. R. S., 94*, 453; *95*, 155, 1918. [3] *Phil. Mag., 39*, 423.

calculations showed that the blue-sky could be fully accounted for by molecular scattering of light by the air.

It is possible to see the scattering of blue light from a layer of air only a few feet in thickness, illuminated by normal sunlight. An extremely black background is necessary, *e.g.* a hole six inches in diameter in a closed wooden box three or four feet square painted black on the inside. A square of black velvet should be tacked on the inner wall opposite the hole and the box oriented with the plane of the aperture parallel to the sun's rays, observations being made from a distance of say ten or twenty feet. We require an absolutely black background for comparison with the illuminated air between our eye and the box. This is most easily secured by covering one end of a pasteboard mailing tube a foot or so in length and 1 ½ inches in diameter with black paper perforated at the centre with a hole 3 mms. in diameter. The open end of the tube is pressed firmly against the eye-socket, and pointed towards the aperture in the box and the small circle of light in the black paper brought over the aperture. It will then appear filled with a pale blue light in marked contrast with the very black field formed by the inner surface of the paper. The experiment must be made on a very clear day when there is no dust in the air. Once having observed the effect the distance can be reduced to three or four feet. Incidentally it may be mentioned that a very surprising effect can be seen if the tube is pointed at the sky or grass, or any brightly illuminated surface and both eyes kept open. After a minute or two the window in the black paper appears many times brighter than the sky or grass as seen by the other eye, fully twenty times as bright, for the observation was repeated using a black glass reflector (4% reflection) in front of the tube, which reduced the intensity to very nearly that of the field observed by the unprotected eye. The effect depends upon the "dark adaptation" of the protected eye.

Scattering by Liquids, Solids and Crystals. — A similar scattering occurs in the case of transparent dust-free liquids. Ether shows it very well when a beam of sunlight is focussed at the centre of the bottle. A Nicol prism properly oriented shows that the light is polarized. It can also be seen with distilled water, which however is difficult to obtain free from motes. They can be removed, however, by distillation in vacuo *without ebullition*, as was shown by W. H. Martin. A double bulb of the form shown in Fig. 263 is employed, the air being driven out by boiling the liquid in the two flasks, condensing the vapor in a beaker containing a small amount of the fluid. The tube is then sealed at the constriction, the boiling being continued gently until the glass is

on the point of collapsing, when the flames are removed and the seal quickly finished. The water is now all brought into the larger bulb which is immersed in a beaker of water at about 65° C., the other bulb being at room temperature. At the end of half an hour the liquid which has collected in the small bulb is shaken about and poured back into the large bulb, carrying with it the motes and dust particles which have adhered to the wall.

After two or three repetitions of this process, nearly the whole of the liquid is distilled over, and it will now be found to be "optically clean"; that is, it will show no bright specks when sunlight is focussed at the centre, but only the pale blue cone of light scattered by the water molecules. It must be viewed against a very black background, such as a hole in a box painted black on the inside.

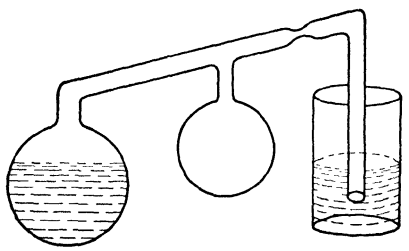


FIG. 263

Now a very striking thing is shown in the case of scattering by a liquid and its vapor. We might very reasonably expect the scattering of the light to be proportional to the number of molecules present, and as there are roughly 1000 times as many molecules in a given volume of liquid as in an equal volume of its vapor at atmospheric pressure, we might look for an intensity 1000 times as great in the liquid as in the vapor. As a matter of fact we observe an intensity less than 50 times that of the vapor.

This is, however, precisely in accord with the theory, for in the liquid we have a nearer approach to a continuous medium and if the medium were perfectly continuous and structureless there could be no scattering, for in this case the secondary waves would destroy each other, as imagined by Fresnel to account for rectilinear propagation.

This means that, in the case of the liquid, "interference" is taking place between the scattered radiations, and the lateral diffusion of light by the molecules is partially prevented.

Transparent crystals, such as quartz, scatter still less light since the molecules are arranged in a regular manner on a so-called "lattice," and this condition approaches the condition of a continuous medium from a mathematical standpoint; at all events there are no variations in density from point to point such as occur in a liquid as a result of thermic agitation or the to-and-fro oscillation of the molecules which constitutes heat.

The absence of scattering in the case of a crystal lattice can be

beautifully shown with a thin flake of mica. Both surfaces should be freshly cleaved, which can be most easily accomplished by rubbing the edge of a thin plate with the side of a needle until it is flattened, and then inserting the point between the layers of the mica. If a flake prepared in this way is attached to the edge of a carefully cleaned cover-glass, such as is used for microscope slides, the strong scattering of the glass, as contrasted with the absence of scattering by the mica, can be shown by focussing sunlight first on one plate and then on the other, viewing them against the black background alluded to above.

The Residual Blue. — Tyndall found that as the particles in his precipitated clouds increased in size, the blue color disappeared, the scattered light appearing white. If, however, it was received through a Nicol prism held in the position in which it would ordinarily extinguish the scattered light, the blue color appeared again in increased splendor. This blue color he named the "residual blue."

Lord Rayleigh [1] considers this phenomenon in a subsequent paper, and shows that if the incident light is polarized with its vibrations parallel to the z axis, the intensity of the light scattered along the z axis varies as the inverse 8th power of the wave-length, so that the residual blue is purer than the blue seen under ordinary conditions. With smaller particles no light at all would be seen in this direction under the conditions specified. He derived an equation which showed that zero illumination was to be expected in a direction inclined backwards, *i.e.* towards the source of light, and this was found to be the case. The experiments were made with a precipitate of sulphur, obtained by adding a small quantity of dilute sulphuric acid to a weak solution of hyposulphite of soda. The more dilute the solution, the slower is the process of precipitation, and the slower the change in the size of the particles. Solutions of such strength that no precipitate appears for four or five minutes will be found to give the best results, and the process can be arrested at any stage by the addition of a few drops of ammonia. The experiment should be performed in a dark room, a beam of sunlight rendered convergent by means of a long focus lens being passed through a glass tank containing the solution. The scattered light should be examined by a Nicol prism. Quoting from the paper above referred to, "In the early stages of the precipitation polarization is complete in a perpendicular direction and incomplete in other directions. After an interval the polarization begins to be incomplete in a perpendicular direction, the light which reaches the eye when the Nicol is in the position of minimum transmission

[1] *Phil. Mag.*, xii, 81.

being of a beautiful blue, much richer than anything that can be seen in the earlier stages. This is the moment to examine whether there is a more complete polarization in a direction somewhat more oblique, and it is found that with θ positive (*i.e.* towards the source) there is in fact an oblique direction of more complete polarization, while with θ negative the polarization is less perfect than in the perpendicular direction itself."

Abnormal Polarization and Color of Light Scattered by Small Absorbing Particles. — In the course of some experiments by the author on the fluorescence of iodine vapor, the precipitation of what appeared to be an iodine fog in one of the glass bulbs was observed. This fog scattered powerfully light of a deep red color, and on examining it with a Nicol prism it was found to be plane-polarized in a direction at right angles to that which is usually observed in the case of light scattered by small particles. When a powerful beam of light was sent through the bulb in a horizontal direction, the scattered light came off at right angles, with its direction of vibration (electric vector) horizontal instead of vertical. If the light was polarized before it entered the bulb, the light was scattered laterally in the directions of the vibration in the incident light. The phenomenon was first observed in a bulb containing a small flake of iodine, and air at atmospheric pressure, cooled by immersion in a beaker of water to which ice had been added. The best method of producing the colored fog is to introduce a bit of sealing wax the size of a pin-head, and a similar speck of iodine, into a small glass bulb, and then heat them both with a very small flame, concentrating sun or arc light at the centre of the bulb.

Nitroso-dimethyl-aniline, a substance which has proved of interest in connection with many optical problems, also gives excellent results. A few grains are placed in the bottom of a large test-tube, and heated with a small flame. A yellow fog at once forms in the tube, which exhibits the same phenomena shown by the iodine fog. Examined with a Nicol with its short diagonal horizontal, the scattered light is yellowish-green and very bright; if the Nicol is turned through a right angle the scattered light is white and much fainter. The nitroso absorbs the blue end of the spectrum. The fog is best examined through the open mouth of the tube, the arc light being focussed at the centre. Under certain conditions, these nitroso fogs showed brilliant colors which were not due to absorption. If the test-tube is fastened at an angle of 45° in a clamp-stand, and the light focussed about an inch below the open mouth, the fog will be found to whirl about in a vortex; and if it is examined with a Nicol, the stream lines will often be found to show brilliant colors. One stream will be bright red and the adjacent stream deep

green, and the colors appear to change to their complementary tints when the Nicol is rotated through 90°; though it is difficult to be sure of this effect, owing to the rapid changes which take place in the distribution of the colors.

These changes remind one of the curious effects observed with thin films of collodion deposited on silver surfaces described in the Chapter on Interference. When a powerful beam of unpolarized white light is concentrated normally upon the film, and the light scattered at grazing emission is examined with a Nicol, it is found that one-half of the visible spectrum is polarized vertically, *i.e.* parallel to the surface, the other half horizontally.

In the case of the iodine fog it seems possible that we are dealing with small transparent red spheres, and that the polarization results from two refractions and one internal reflection, as the effect was duplicated with a small glass bulb filled with a red solution. It is perhaps questionable, however, as to whether such small particles can operate in this manner. The varied colors shown by the nitroso fog have not been explained, though one suspects something in the nature of resonance.

Density Fluctuations and Light Scattering. — The destructive interference of the coherent radiations scattered in lateral directions by the molecules of a liquid will be greatly diminished if density fluctuations are present. The Brownian motion of small particles in a liquid prove that such fluctuations in pressure are present and as a result density fluctuations must also occur which will be proportional to the compressibility of the liquid.

Now the relative intensities of the scattered light for a number of liquids as measured by Martin and Lehrman are as follows:

Benzene	1.00	Xylene	1.26	Alcohol	.2
Toluene	1.14	Chlorbenzene	1.46	Water	.068

while bisulphide of carbon is certainly over 4. on this scale. We cannot explain these widely different values as due to a difference of compressibility, for the compressibilities of the various liquids are not very different in value. Nor can they be referred to differences in refractive index, for the $(\mu-1)^2$ term for water is .18 and for bisulphide of carbon .5 while the scattering power of the latter substance is over 50 times that of water. We shall see presently that the scattering power is associated with the optical anisotropy of the fluid and is determined largely by its Kerr constant. The theory of molecular scattering in its relation to density fluctuations was first worked out by Smoluchowski [1] and subsequently by Einstein.[2]

[1] *Ann. der Phys.*, *25*, 205, 1908.
[2] *Ann. der Phys.*, *33*, 1275, 1910.

The scattering power of saturated vapors increases enormously as the critical temperature is approached giving rise to the phenomenon of "Opalescence," which was formerly supposed to result from the presence of minute droplets suspended in the dense vapor, but which Smoluchowski explained as due to density fluctuations. These will increase with the temperature owing to the more violent agitation of the molecules.

Ramanathan measured the scattering power of saturated ether vapor above the liquid in sealed tubes at various temperatures. The following values are taken from his table:

Temp.	Scattering Intensity	Density
33°	1.	0.92
108	11.8	7.1
170	103	28.
186	553	44.
194	9950	71.
200	3316	71.
202	1680	71.
215	710	71.
217	553	71.

The increase in scattering is, of course, in part due to the increase in density, that is to the increased number of scattering molecules but it increases much more rapidly than the density. For example at 33° both values are practically unity, while at 186° the scattered intensity has increased 553 times while the density increment is only 47. The critical temperature is 194 and in the vicinity of this point the density fluctuations are abnormally great, giving rise to the phenomenon known as "Opalescence."

Measurement of Polarization. — The polarization of the scattered light is most conveniently measured with the following apparatus. The scattered light from the tank containing the liquid, and traversed by the beam of light, passes through a double image Wollaston prism oriented to give an upper and lower image with the electric vectors vertical and horizontal respectively. The prism is placed at such a distance in front of the tank as to cause the two images to appear in contact, or employing a biprism of small angle behind the Wollaston, as was done by Martens [1] we obtain two fields divided by a sharp line which vanishes when the fields are made equal. Next comes a Nicol oriented to extinguish the upper image. By turning this Nicol to the left or right through angle θ the images can be made of equal intensity, the double angle 2θ being read to ensure greater accuracy. (See Chapter on Polarization, page 340.)

[1] *Phys. Zeit.*, *1*, 299.

Depolarization of Scattered Light. — In the cases of gases and liquids illuminated by polarized light, the scattered light is frequently found to be only partially polarized, *i.e.* there is a depolarizing factor. Depolarization occurs when the molecule is not isotropic, that is when the "moment" of the induced dipole is not in the direction of the electric vector of the light-wave. Such a molecule is said to be anisotropic. The molecules are arranged in various orientations, and the dipole vibrations which produce the scattered radiations are not all parallel to the vibrations of the light.

The imperfection of the polarization or degree of depolarization is the quantity $r = \tan^2 \theta$ given on page 340 in the Chapter on Polarization. The method usually employed is the one there given which is due to Cornu. The Wollaston prism is placed at such a distance in front of the tank containing the liquid that the upper image (polarized with electric vector vertical) is in contact with the lower image. The Nicol is set to extinguish the upper or brighter image and then rotated through angle θ to give equality of the two images. Then $r = \tan^2 \theta$.

Nearly isotropic are the inert gases with a depolarizing factor .0055 and carbon tetrachloride .0088 observed by Cabannes; hydrogen comes next .02, oxygen .06 and carbon dioxide .08.

Liquids show a much greater depolarization, the ring compounds (benzene for example) being more anisotropic than the chain compounds (alcohols for example).

DEPOLARIZATION FOR LIQUIDS [1]

Benzene	.45	Chloroform	.24
Chlorbenzene	.50	Ether	.09
Nitrobenzene	.73	Carbon Tetrachlor.	.045
Bi-Sulp. Carb.	.64	Methyl Alcohol	.062

Ramanathan [2] found for vapor and liquid benzene the following values for r at different temperatures:

BENZENE

VAPOR		LIQUID	
Tempt.	r.	Tempt.	r.
35°	.072	35°	.43
100	.062	125	.32
182	.035	182	.19
228	.025	205	.14
267	.015	228	.095
		268	.03

[1] From table by R. Ganz, *Ziet. für Phys.*, *17*, 384.
[2] *Phys. Rev.*, *21*, 566, 1923.

This shows that the depolarization for the liquid is much greater than for the vapor, and that in both cases it decreases with rising temperatures. There appears to be no satisfactory explanation of the greater depolarization shown by the liquid.

He assumes that there is an extra scattering due to the anisotropy of the molecules, the scattered light consisting of three parts, a polarized part P_d due to density fluctuations, an unpolarized part $4N_a$ and a polarized part P_a, both due to the anisotropy of the molecules the value of r being

$$r=\frac{2N_a}{P_d+P_a+2N_a}.$$

The fluctuations of density and consequently P_d increase with rising temperature, consequently r decreases.

The presence of ultra-microscopic particles in the liquid (colloidal solutions) as Cabannes points out diminishes the value of r, since a liquid which gives *complete* polarization of the scattered light is a colloidal solution or a suspension of spherical isotropic particles. For pure water $r=0.1$ while colloidal silica has the value $r=.0025$. The presence of colloidal particles can be detected by observing the scattered light in a direction inclined at a few degrees to the incident beam in which case the molecular scattering becomes small in comparison to that caused by the particles which stand out as shining specks on a darker background. For a complete treatment of molecular scattering in all of its aspects the reader is referred to Cabannes's book *La diffusion moleculaire de la lumière.*

Molecular Structure and Depolarization. — This depolarization of the scattered light, which, when first observed might perhaps have been considered as of little interest or importance, has in recent years furnished much information about the arrangement of the atoms in molecules.

R. Gans [1] developed a quantitative theory of the depolarization, considering the polarizability of the molecule as varying with the direction of the field. The deformation produced by a field of given direction is represented by postulating a "deformation ellipsoid, of axes b_1, b_2, b_3, each representing the polarizability of the molecule when the electric force is parallel to an axis." The induced moment is then $p_1=b_1D$, $p_2=b_2E$, etc.[2]

Suppose a molecule oriented with respect to the field of force E as shown in Fig. 264. The vector E of the polarized exciting light is resolved along the axes b_1b_2 of the ellipsoid into components E_1 and E_2, and the corresponding induced moments are $\mu_1=b_1E_1$

[1] *Ann. der Phys.*, *64*, 481; *65*, 97, 1920–21.
[2] See Chapter on Dispersion Theory. Electric Dipoles.

and $\mu_2 = b_2 E_2$, the resultant moment μ being turned through angle Φ with respect to the electric vector of the light. The electric vector of the scattered light will be parallel to μ, *i.e.* a single molecule will emit completely polarized light but a collection of molecules, partially polarized light owing to the different orientations of μ.

Silberstein [1] regards the atoms as isotropic, since the degree of depolarization found for monatomic gases, or molecules having spherical symmetry such as carbon tetrachloride, is of the order of experimental error and can be regarded as zero. In the case of molecules the anisotropy results from a reciprocal electrical polarization resulting from the electrical moments of the single atoms induced by the light.

Fig. 265 represents two chlorine molecules the upper under the influence of an electric force parallel to the line joining the atoms, the lower with the force perpendicular to the line. The electric polarization is represented by the + − signs, and in the upper molecule we have increased polarization due to the

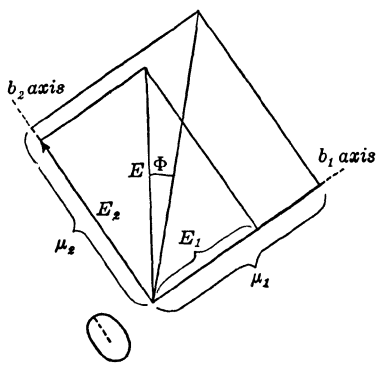


FIG. 264

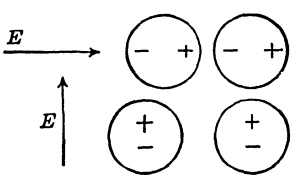


FIG. 265

action of each dipole on the other, while in the lower the polarizing effect of the field is diminished by this reciprocal action since the field of each induced dipole diminishes the polarization of the other. For a single isolated atom the distance between the + and -- sign would be intermediate between the two cases figured.

As illustrations of the application of the study of depolarization to molecular structure we may take the case of two similar molecular structures, carbon tetrachloride and methane, of which the degrees of depolarization are

CCl_4	.005
CH_4	.015

The former has the chlorine atoms at the corners of a symmetrical tetrahedron with carbon at the centre, as any other arrangement

[1] *Phil. Mag.*, *33*, 92, 215, 521, 1927.

would give a much larger depolarization. The higher value for methane was explained by Hund as due to the deformation of the tetrahedron by the interaction of the hydrogen atoms, owing to which they lie on the opposite corners of a quadratic column instead of a cube (Fig. 266).

Hund's model has no electrical moment, explains the anisotropy and is in good agreement with the infra-red absorption spectrum. A very full account of the theory of depolarization in its relation to molecular structure will be found in Vol. 10 of the *Ergebnisse der Exakten Naturwissenschaften* by H. A. Stuart, who has done a large amount of experimental work on the subject.

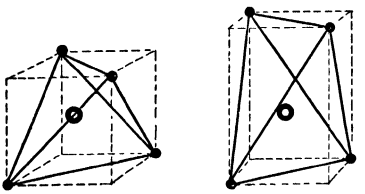


FIG. 266

Scattering and Transmission Colors of Granular Metal Films.—A series of experiments were carried out by the author [1] in 1902–03 on some remarkable colors exhibited by films of the alkali metals deposited by evaporation on the inner walls of exhausted glass bulbs. It seemed impossible to explain these colors as interference or diffraction phenomena, and as the microscope showed clearly that the films were granular, the hypothesis was made that we were dealing with optical resonance for light-waves. The films are very easily prepared in the following manner:

A number of small bulbs are blown of the form shown in the figure, and a piece of sodium or potassium is cut up under ligroin into blocks about 2 mms. on each edge. These are introduced into the bulbs as quickly as possible after wiping off the fluid, and the stems of the bulbs drawn down to a small bore for subsequent sealing.

They are quickly fastened to the branched tube and exhausted. It is a good plan to heat the metal until it fuses, while the bulb is still on the pump. The bulbs are now sealed off from the pump, and may be put away for future use, or experimented with at once. A burner should be made by drawing out a glass tube, which will give a pointed flame about half a centimetre high, and the tip of the flame allowed to play against the spot on the bulb where the metal lies (Fig. 267, *a*). Sometimes the whole bulb will suddenly flash a deep violet or blue, and sometimes the film will develop more slowly. A chain of three or four bulbs may be made, the sodium heated in one, and the clean molten metal shaken into the others, drops of various sizes sticking to the bulbs. Colored films can then be formed by heating these clean drops in the bulbs. This shows

[1] *Phil. Mag.*, April, 1902; Oct., 1902; Aug., 1903.

that the oxide present in the first lump has nothing to do with the production of the color.

If the metal is heated at one end of a rather long tapering bulb, the color is most intense near the metal and gradually fades away to nothing at the other end of the bulb. If the bulb is placed in strong sunlight with a black background, it is seen that in some places where the deposit is too slight to show much color by transmission, the light is scattered or diffused, and this diffused light is colored. The claret-colored or purple film, where the deposit is slight, scatters a green light, the surface appearing as if fluorescent. Now the spectrum of the transmitted light in these purple films has a heavy absorption band in the yellow-green, consequently the scattered light is the complementary color of the transmitted. If the film is greenish-blue, the scattered light is reddish. The microscope shows that in these deposits, which have the power of scattering light, the individual particles are rather widely separated, that is, the distance between them is large in comparison to their diameters. The appearance of a bulb in strong light is very much as if certain portions of its interior surface had been painted over with a solution of fluorescein. No trace of regular reflection is shown by these films, except of course the reflection due to the glass. The particles are so far apart that they apparently act as independent sources, the interference necessary for rectilinear propagation not being present. If the incident light is polarized, the scattered light is also polarized, which is not the case for ordinary diffuse reflection. Passing now to a part of the film where the color of the transmitted light is deeper, we find that there is no longer any trace of this scattered light. The color absent in the transmitted light is now regularly reflected, the particles being so close together that interference, as imagined by Fresnel, takes place.

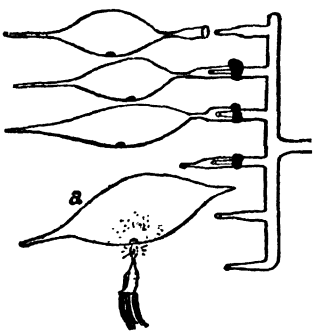


FIG. 267

One of the most surprising effects observed in these early experiments was that the color of the films changed when a spot on the bulb was cooled by the application of ice, or even by the evaporation of a drop of water. This was presently found to be due to the condensation of ligroin vapor (introduced with the metal) by which the metal particles were immersed in a medium of high dielectric constant. This was considered as analogous to the resonators of Aschkinass and Schaefer which responded to longer wave-

lengths when immersed in such a medium. It is a good plan to moisten the bit of metal with a little ligroin before its introduction into the bulb. The color changes are most remarkable. Pink and purple films become blue, while pale apple-green films change to a deep blue-violet, as deep as dense cobalt glass. Blue films often became perfectly transparent, the absorption band, originally in the red, moving out of the visible spectrum entirely. Spectroscopic examination showed that the immersion of the resonators in the liquid dielectric caused the absorption band to move towards the region of longer wave-lengths, as it should do according to theory.

Colors of Granular Films of Gold and Silver. — To ascertain whether the color effects were common to all metals in a state of fine subdivision, experiments were made with gold and silver, the former obtained by electrical discharges in high vacua from a gold cathode, the latter by employing Carey Lea's solution of allotropic silver.

The color of the gold deposit varies with the conditions under which the deposition takes place. Gold cathodes of two forms were employed, a flat plate about 3 cms. square, and a thick wire, screening off the radiation from all but the tip with a mica screen. The most interesting deposits were obtained from the small source. In one instance the film showed a brilliant green surface color, resembling fuchsine, the transmitted light having a purple tint. Owing to the transparency of the film a good deal of white light is mixed with the selectively reflected light; this can be cut off with a Nicol, if the reflection takes place at the polarizing angle for glass, and the colored light from the film, which is unpolarized, then appears in great purity. One plate showed patches of brilliant carmine red, deep blue and green, of a surprising intensity and saturation. The color of the selectively reflected light depended somewhat on the angle of incidence, a phenomenon observed also in the case of the sodium and potassium films. Increasing the angle of incidence changed the color from green to blue.

If the glass plate is placed near the tip of the gold wire, the green deposit, similar to gold leaf in its optical properties, is deposited at the centre. The film at this point is not granular, the metal vapor not having condensed to a fog before reaching the glass. Surrounding this is a film appearing light yellow by transmitted light, and bluish by reflected light. This seems to be what we should expect, for the smallest particles, which will resonate for blue light, will be deposited when the distance from the cathode is a little greater than that at which the molecular deposit occurs. Increasing the distance, we get larger particles, and the point of maximum reso-

nance moves up into the green, giving us a purple film with green surface color. At a still greater distance we get particles large enough to resonate for red, and the film appears deep blue by transmitted light. All of these variously colored films can be changed into the green structureless film by heating. We may regard the change as due to the fusing together of the resonators.

A very easy way to prepare a purple gold film, scattering green light where the deposit is very light, is to heat a piece of gold the size of a pin-head in a small quartz tube with an oxy-coal-gas flame or even with an ordinary blast-lamp. Near the bead we have a continuous deposit, showing the ordinary gold reflection: beyond this a purple film (by transmitted light) and still further along a very light deposit which scatters green light when held in sunlight and viewed against a black background. Silver films showing brilliant colors can be prepared by employing a solution of so-called allotropic silver described by Lea. Three solutions are prepared: a 30% one of ferrous sulphate, a 40% one of sodium citrate, and a 10% one of silver nitrate. Fourteen c.c. of the citrate solution are mixed with 10 c.c. of the ferrous sulphate solution, to which are then added 10 c.c. of the silver nitrate solution. A dense black precipitate at once forms, and the whole is at once poured into a filter. As soon as the liquid has entirely run through, the precipitate is washed with 10 c.c. (*not more*) of distilled water. This removes the salts which make the precipitate insoluble. After the water has entirely passed through the filter, about 25 c.c. of distilled water are poured into the filter, and the blood-red solution which runs through collected. As it does not keep very well, it is best to prepare it on the day on which it is to be used. It was found by accident that the colors could be obtained only when the glass was covered with a trace of gelatine.

A sheet of glass is washed clean, rinsed with fresh water, and the wet surface rubbed over with some shreds of gelatine. It is then drained for a few seconds and dried on a hot plate. A little of the silver solution is flowed over it, the surplus being drained off. If too much gelatine has been used, precipitation is apt to take place, the deposit taking the form of floating shreds of a reddish membrane. If no considerable precipitation occurs, the plate, which should have been quite warm when flowed, is placed once more on the hot plate until dry. The films formed in this way are usually deep red in color, though sometimes patches of deep violet form, with sharply defined edges. Violet patches may be easily formed in the following way: When the plate is about half dry, with a steaming film and a few small pools of the hot solution, it is removed from the hot plate, held at an angle, and treated with a few drops of

alcohol, which are allowed to run down across the still damp portion of the plate. These portions speedily dry into a most gorgeous mosaic of red, purple and violet patches, the experiment being especially striking in the lantern, as it occupies but a few seconds, and the color-display spreads over the plate like the blaze of a sunset.

These films are transparent to the entire ultra-violet region, even down to the cadmium lines, of wave-length 22 or thereabouts.

The Scattering of Light by Sodium and Potassium Fogs. — Allusion has been made to the deep violet light scattered by a condensing cloud of sodium vapor. The author has frequently observed that the color of the light transmitted through the sodium tubes, in the experiments upon the optical properties of the vapor, was colored a deep yellow instead of blue, as is usually the case. It was difficult to understand this at first, since the vapor is perfectly transparent to blue light, and somewhat less so to yellow-green light. The cause was finally found to be a scattering of the violet and blue rays by the fog of condensing vapor, which was so powerful that none of these rays was transmitted. The phenomenon was investigated further with the large tube used in the experiments on fluorescence, some potassium being introduced into the retort. The light from the arc was focussed about 20 cms. in front of the retort and the tube rapidly heated by a blast-lamp. Under this condition the vapor is puffed out in clouds from the aperture of the retort, and the condensing clouds are most wonderfully colored, red and orange predominating. It was found that if the heating was carefully regulated, a steady state could be maintained in which the fog scattered red light at the outer boundary, where the cone of rays entered it, yellow a little farther in and green at the point where it was first forming and where the cone of rays passed out into the vapor in which it was of course invisible. The cone of rays seen from the side, which can be accomplished by looking in at the edge of the glass window, resembled a spectrum, the blue end of which was wanting. A potassium fog thus scatters longer waves than a sodium fog, and the scattering power is confined to a comparatively narrow region of the spectrum, which apparently varies with the size of the particles in the fog. It is probable that we are dealing with something analogous to the phenomenon exhibited by the granular deposits of these metals already described. Further investigations along these lines should be made, especially with polarized light. A large brass tube without a retort, or even a glass tube, could be used.

Scattering by Metal Colloids. — The laws which govern the scattering of light by minute metal particles in suspension in a

liquid (colloidal solutions) are quite different from those which obtain in the case of small particles of a dielectric. A very extensive study of this subject has been made by G. Mie.[1]

Colloidal particles of gold scatter green light but the color is modified by the absorption of the solution, which appears red by transmitted light. Mie avoided this difficulty by observing the scattering of single particles with the ultra-microscope in solutions so dilute that absorption played no part. He found symmetrical scattering for the smallest particles, but as their size increased the scattering was greater in the direction in which the exciting waves were travelling. Illumination with unpolarized light gave scattered light completely polarized at 90°, the polarization becoming less as the size of the particles increased.

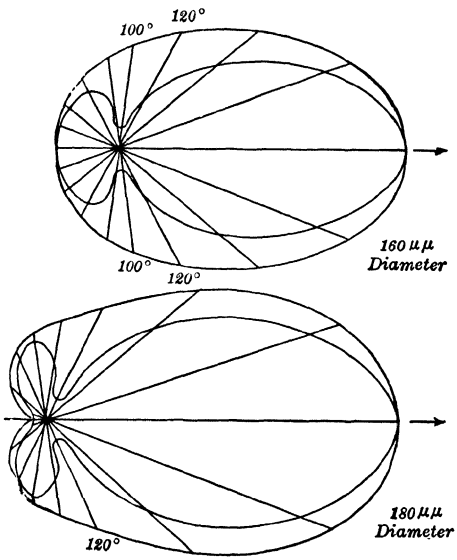


FIG. 268

Mie recognized two types of absorption "consumptive" in which the energy was transformed into Joule heat in the metallically conducting particles, and "conservative" in which it was scattered laterally, the former having a sharp maximum in the green, the latter in the red-yellow. If the particles are very small, the consumptive absorption predominates, and the solution is purple by transmitted light. Larger particles diffuse the red-yellow light and the solution appears blue. In studying the intensity distribution and polarization of the scattered light Mie obtained some very striking results, which are best understood by reference to his scattering diagrams, two of which are reproduced in Fig. 268. The direction of the light is indicated by arrows, and the scattering intensity is represented by the length of the lines radiating from the particles. It is very strong in the direction of the propagation of the light. The intersection of the lines with the outer and inner curves represent total intensity and unpolarized intensity respectively, consequently the polarized portion is the

[1] *Ann. der Phys.*, *25*, 377, 1908.

portion of each line included between the two curves. Maximum polarization occurs at 110° and 120° in the two cases figured. He calculated that for a perfectly conducting particle the maximum scattering would be in the direction opposite to that in which the light is travelling, and some twelve times as intense as in the opposite direction.

Colors in Metal Glasses and Metallic Films. — An extremely interesting treatment of the colors exhibited by glasses which are stained with metallic oxides, and by the films of the alkali metals deposited in vacuo, has been given by Garnett.[1] Some colored glasses are supposed to owe their color to the presence of minute globules of the metal. Colloidal solutions of the metals act in a similar manner, and recently the presence of the particles has been detected in both cases by Siedentopf and Szigmondy, by means of their so-called ultra-microscopic method, which is merely oblique illumination pushed to the limit. A powerful beam of light issuing from a horizontal slit is brought to a focus by means of a microscope objective, within the glass or liquid under investigation. The small particles scatter some of this light and appear as minute diffraction disks of light, when a microscope is focussed on the illuminated plane (Fig. 269). The phenomenon is analogous probably to the scattering of light by the air molecules or small particles suspended in the atmosphere, the action of which has been exhaustively treated by Lord Rayleigh. A simpler device, due to Cotton, is shown in the lower part of the figure, in which the illuminating beam is prevented from entering the microscope by total reflection.

Garnett shows that the colors of the glasses can be accounted for by the presence of the small metal spheres, and explains a number of very curious effects observed by Siedentopf and Szigmondy, some of which we will now consider.

Polarization Effects of Ultra-Microscopic Particles. — Some very curious and interesting effects were observed by Siedentopf and Szigmondy in the case of gold particles when the illuminating beam was plane-polarized. If the plane of the vibration was perpendicular to the plane containing the illuminating ray and the microscope, the little diffraction disks appeared everywhere in the field, and were of uniform illumination. The scattered light was polarized in the same plane as the incident. This condition is shown at *a* (Fig. 270). The field of the microscope is represented by the large circle, and the direction of vibration of the scattered light by the arrows. The appearance of the diffraction disks is shown below. If, however, the incident light vibrated parallel to

[1] *Phil. Trans. of Roy. Soc. Lond.*, Series A, vol. 203. p. 385.

the above defined plane, the particles scattered no light in the vertical direction, and the diffraction disks were formed by the oblique rays gathered in by the objective. In this case each diffraction disk appeared with a black spot at its centre as shown. Moreover, since the light scattered in any given direction by the particles comes to a focus at a given point in the second focal plane of the microscope, a black spot will be found at the centre of this plane.

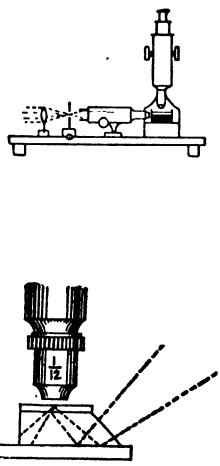


FIG. 269

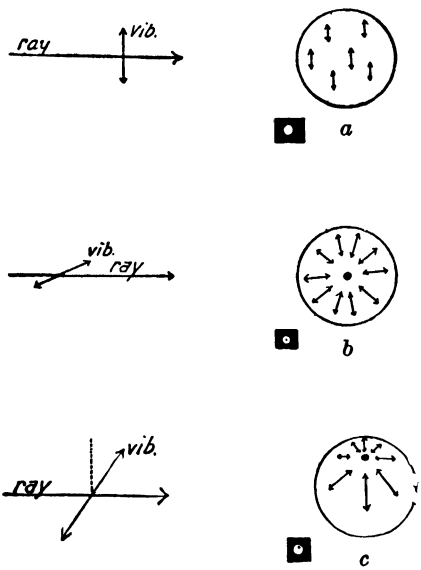


FIG. 270

This condition is shown at *b*. If the incident vibration is inclined at 45°, the spot appears on the sides of the diffraction disks, and there is a dark region on one side of the second focal plane as in *c*. The direction of the vibration at the second focal plane is indicated in each case by the double-headed arrows in the large circles.

CHAPTER XIV

THE RAMAN EFFECT ✳✳✳

Smekal's Prediction. — A paper was published by A. Smekal [1] in 1923 on the quantum theory of dispersion, showing that, in the case of the scattering of light by transparent media, frequencies other than those present in the original light might be found in the scattered radiations, a phenomenon now known as the Raman effect. He considered that a molecule of mass m in an energy state E_p (the sum of its electronic, vibrational and rotational energy) moving with a velocity v, on collision with a light quantum of energy $h\nu$ might pass to another energy state E_q and experience as well a change in velocity. From the conservation of energy we have

$$\frac{mv^2}{2}+E_p+h\nu=\frac{mv'^2}{2}+E_q+h\nu'.$$

He showed that the change in velocity was negligible so that

$$E_p+h\nu=E_q+h\nu' \quad \text{or} \quad \nu'=\nu-\frac{E_q-E_p}{h}$$

which means that the quantum gives up some of its energy to the molecule and is scattered with a lessened energy. If the molecule is in a higher energy state before the collision than after, the light quantum may receive energy and be deflected at a higher frequency giving a line on the short-wave-length side of the exciting line.

The frequency difference between the original and deflected radiation is designated by $\Delta\nu=(E_q-E_p)/h$ and if equal to zero, the quantum is scattered with unchanged frequency, the collision being elastic, and we have the Tyndall or Rayleigh scattering.

The dispersion theory of Kramers and Heisenberg, published in 1925, contains, as is now well known, a complete theory of the Raman effect, the essential features of which are that *two* transitions must be considered instead of one, as in the case of ordinary absorption. If there are three energy levels a, b and c (the molecular energies being in part electronic, and in part atomic vibration and molecular rotation), and the level c can be reached in an

[1] *Naturwiss.*, *11*, 873

ordinary transition from both the levels a and b, a line of altered frequency may be scattered. This will be more fully explained when we come to the theory of the effect.

Raman's Experiment. — The discovery of the predicted effect was announced in 1928 by Sir C. V. Raman and a few weeks later by Landsberg and Mandelstamm, the former working with liquids, and the two latter with crystalline quartz.

Raman focussed the light of a mercury arc at the centre of a large glass flask filled with the liquid under investigation and pointed the spectroscope at the illuminated region. Exposures of many hours were necessary to secure a photograph of the spectrum, owing to the faintness of the light. Numerous bright lines were found in the case of every liquid examined. In the majority of cases the new lines were of longer wave-length than that of the

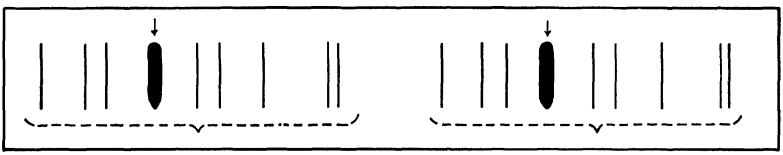


FIG. 271

exciting light, but in the case of some liquids such as carbon tetrachloride, chloroform and benzene, lines of shorter wave-length were found as well.

At first sight it might appear as if this was merely a case of fluorescence, but Raman found a remarkable relation between the wave-lengths of the new lines and that of the exciting radiation, which is quite different from anything observed in the case of the phenomenon of fluorescence.

In the first place he found that the arrangement of the spectrum lines in the groups excited by the various mercury radiations was exactly similar. For example, carbon tetrachloride gives a spectrum shown by diagram in Fig. 271, two mercury lines strongly scattered by the liquid being indicated by arrows. The lines to the left of the mercury lines are of shorter wave-length than that of the exciting line and are called "anti-Stokes lines," being exceptions to a supposed law formulated by Sir George Stokes, that, in the case of fluorescence, only radiations of wave-length longer than that of the exciting light could be emitted. The relative spacing is the same to the right and left of the exciting mercury line except that it is reversed. Secondly, he found a very remarkable relation between these new lines and the absorption bands of the substances in the infra-red region of the spectrum. If the frequencies of vibration of the Raman lines are subtracted from the frequency of the

exciting light the numbers obtained represent the frequencies of infra-red absorption bands. The matter turned out to be somewhat more complicated than was at first supposed, for not all Raman lines have infra-red absorption bands associated with them, and some infra-red bands are not represented by Raman lines.

The original and simpler theory, as first given by Raman, explaining the formation of anti-Stokes lines, is as follows: The quantum theory of light regards a beam of monochromatic light, say the green light of the mercury arc, as made up of discreet "packets" of energy, all of the same magnitude, called "light quanta." As the beam diverges in space these quanta merely get farther apart, without suffering any loss of energy. The energy content of a light quantum increases progressively as we pass from the red to the violet and ultra-violet.

When now one of these quanta enters the benzene, one of three things may happen. It may be diverted from its original direction without loss of energy, giving the line of unmodified wave-length in the Raman spectrum. It may give up some of its energy to a benzene molecule, and emerge from the medium with diminished energy, giving a line on the red side of the exciting line. Or it may abstract a certain definite amount of energy from a molecule already in an excited state — that is, in vibration as a result of temperature — and emerge with increased energy, giving an anti-Stokes line. It is found that the intensity of these anti-Stokes lines increases as the temperature of the medium is raised, as more molecules are then in excited states.

The ideal source of light for the excitation of Raman spectra would be a lamp (combined with a filter) which emitted powerful monochromatic radiations of a single frequency only. The nearest approach to this ideal is the helium vacuum tube or the hot cathode helium arc, surrounded by a tube of nickel oxide glass as employed by the author.[1] In this case the excitation is by a single frequency ($\lambda = 3888$) and a single Raman spectrum is formed, instead of a number of superposed spectra, as in the case of the mercury arc excitation. A number of these, reproduced on Plate 8, Fig. 3, give an excellent idea of single Raman spectra. (λ values at top.)

The helium method is not recommended, however, as the mercury arc gives much less trouble, though the study of the spectrograms involves more labor.

As originally employed by Raman the illumination is not very efficient, as only a small percentage of the light of the lamp is utilized. The arrangements now generally used are based on those described by the author shortly after Raman's first publication,

[1] *Phil. Mag.*, 7, 858, 1929.

the liquid under investigation being contained in a glass tube provided at one end with a window of glass fused to the tube, or with a flattened bulb, the other end being drawn off obliquely and painted black. This tube is surrounded with a second tube in which water circulates to prevent overheating, and the arc brought up as close to the tube as possible.

If only very small quantities of material are available for study (10 c.c. or less) a form of tube shown one-half natural size in Fig. 272 may be employed. The Raman tube should be made of thin-walled soft glass tubing (not Pyrex) about 13 mms. in diameter, one end blown round and then flattened. A water-jacket tube about 23 mms. internal diameter surrounds this tube, which is held in place by a rubber gasket at the bottom. After slipping the gasket over the tube the upper portion is drawn down as shown and painted with black "Duco." The upper portion of the large tube is wrapped in tinfoil wound with a copper tube 3 or 4 mms. in diameter, the space between being filled with water or a filter solution. Cooling is effected by water circulating in the copper spiral. A concave ¾-inch cylinder of polished sheet aluminum is slipped around the inner tube on the side opposite the lamp, unless a filter solution which attacks aluminum is being used. The tube is illuminated with a capillary quartz-mercury arc, almost in contact with the outer tube.

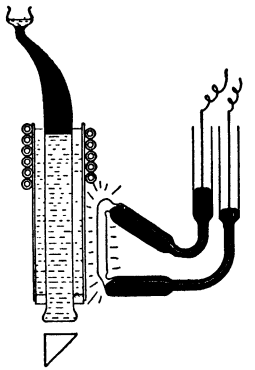


FIG. 272

Other forms of early tubes designed by the author will be found described in the *Physical Review*, 1929–32. The most efficient arrangement of all was found to be the combination of a Hanovia 220-volt quartz-mercury arc in the metal housing turning on trunnions as supplied by the manufacturers, with filter solutions contained in large glass tubes, which act as cylindrical lenses. A solution of quinine sulphate dissolved in water acidulated with a few drops of sulphuric acid removes the ultra-violet and the 4046 group, giving an excitation by the 4358 line. Its chief disadvantage is that it is rapidly decomposed by the light, turning yellow. This action was found to be due chiefly to the lines of shorter wavelength than 4046, which are absorbed by "noviol" glass, which however transmits 4046. A thin sheet of pale noviol combined with quinine is the most efficient filter, for complete absorption of 4046 and perfect transmission for 4358. The most suitable thickness is noviol "O," 2.5 mms. obtainable ground and polished from the

Corning Glass Company, Corning, N. Y. A plate six inches square should be ordered, cut into strips 2×6. The burner is operated in a horizontal position just below a narrow aperture in the roof of the iron box. The noviol plate is supported about 1 cm. above the aperture as shown in Fig. 273, and a bent sheet of thin aluminum directs the air blast of an electric fan across the aperture under the noviol plate. The fluid filter is contained in a glass tube 5 cms. in diameter and 25 cms. in length mounted just above the aperture.

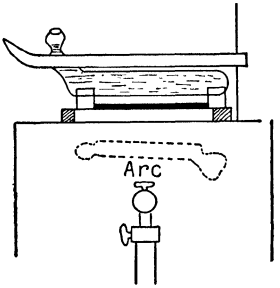


FIG. 273

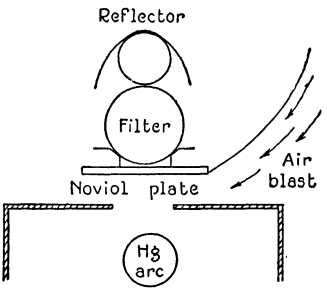


FIG. 274

This acts as a cylindrical lens as well, forming an image of the arc along the axis of the tube containing the liquid under investigation, constructed as just described. Over this is placed a reflector of thin, highly polished sheet aluminum now obtainable in the market. (Fig. 274.) A more convenient and equally efficient filter is a saturated solution of sodium nitrite, as suggested by A. H. Pfund. This requires no plate of noviol glass.

The Raman tube is of a form similar to the one already described. The apparatus should be mounted about two metres from the spectrograph, and an image of the end of the tube focussed on the slit with a lens of about 25 cms. focus.

A few words regarding the best method of pointing the collimator of the spectrograph down the axis of the illuminated tube and forming the image on the slit may be found helpful.

In general, when the source of light is a long cylinder, seen end-on, it is important to mount it at such a distance from the spectrograph that both ends of it can be fairly well focussed on the slit simultaneously. This means that the distance from the source to the lens should be much greater than the distance from the lens to the slit.

First determine the position for the slit by bringing the eye to the proper point for looking along the axis of the tube, marking the point with a pin held in a clamp-stand, and bringing the slit up to the pin-head.

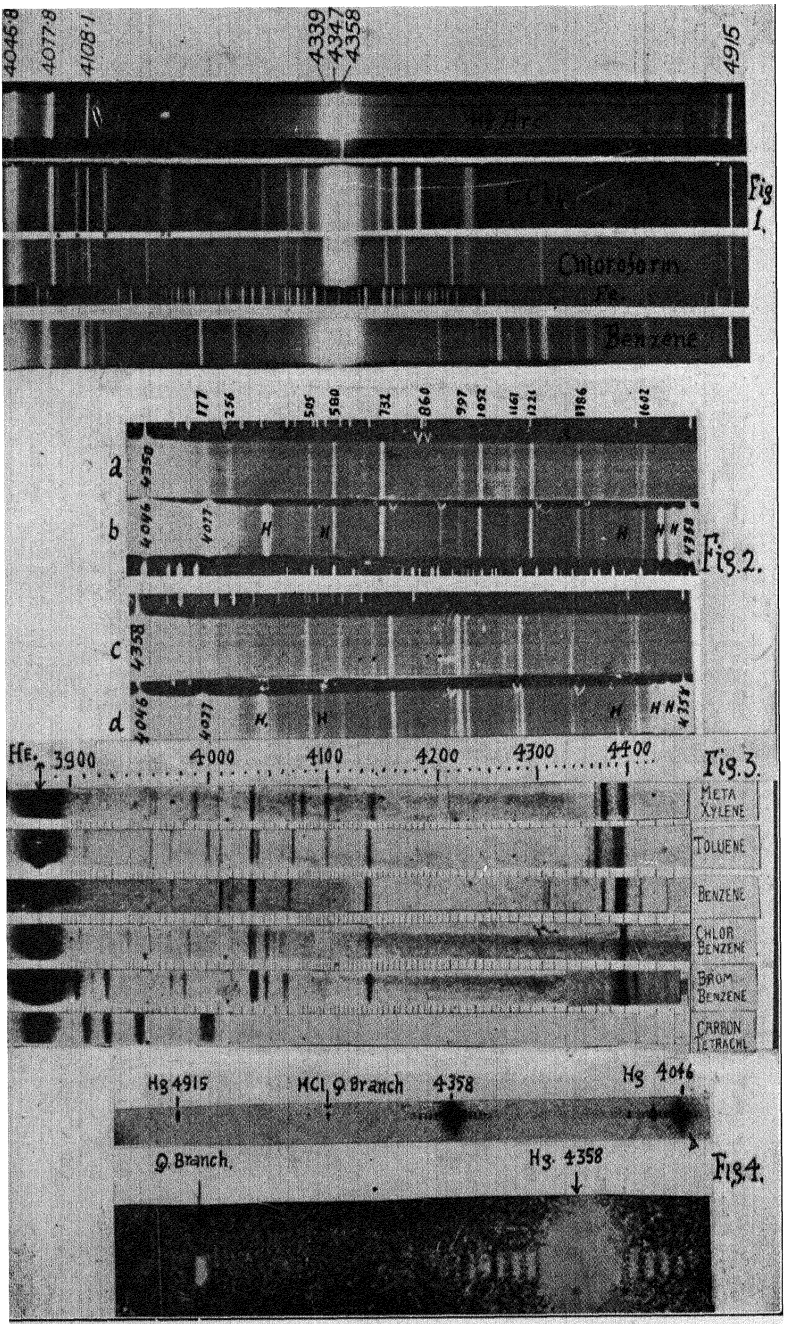


PLATE 8. RAMAN SPECTRA.

Next look into the spectrograph from the position of the plate (*i.e.* from the focal plane of the camera) and if the aim has not been very bad, the end of the tube will appear as a small circle of light somewhere in the field of the camera lens. Next shift the spectrograph until the circle of light is in the centre of the lens, keeping the slit position fixed.

To form an image of the end of the tube on the slit two methods are available. If a large spectrograph is being used, a small circle of paper should be mounted over the end of the tube and strongly illuminated with a lamp. This can then be focussed at once on the slit. With small spectrographs the following method is very convenient. The lens is first mounted practically in contact with the slit and the spot of light in the camera lens watched. As the lens is moved towards the tube and away from the slit the spot of light gradually enlarges, finally filling the entire field of the lens. The lens should be moved so as to keep the spot centrally placed in the field, and when the whole field is uniformly illuminated, the lens is in the correct position. This method enables us to utilize the brightest part of the source, if the illumination is not uniform, for, as the spot enlarges the attention can be focussed on the bright portion and this is then made to fill the field. Raman spectra of carbon tetrachloride, chloroform and benzene, excited by the total radiation of the mercury arc are reproduced on Plate 8, Fig. 1.

For excitation by 4046 a tube filled with a solution of iodine in carbon tetrachloride may be used. A saturated solution should first be made, and diluted for use. A 1:10 dilution will reduce 4358 by about the right amount. It can be abolished entirely with a 1:5 dilution, but 4046 will then be considerably reduced as well. With a dilution of 1:35 the excitation is by both 4046 and 4358, but the continuous background is considerably reduced between these two lines and faint Raman lines come out which cannot be seen without the iodine filter. The best method to follow, and the one most likely to give results free from error is to take two photographs of each substance, one with the noviol quinine filter, and one with noviol iodine, the noviol suppressing 3650, and the iodine 4358.

In general the spurious lines are more readily picked out by mere inspection in the case of 4358 excitation than with 4046, as they form triplets with the real lines which can usually be recognized at a glance. The greater separation of 4046 and 4077, makes it difficult or impossible to pick out the lines due to 4077 without measuring their distance from the 4046 line.

It is very helpful to make enlargements of each spectrogram on

such scales that the Raman lines can be brought into coincidence, as shown by Plate 8, Fig. 2, *a* and *b*, on which the numerals above the spectrum refer to the frequency differences between the exciting line and the Raman line. In other words enlarge the 4358 excitation more than the 4046, which is more dispersed by the prism spectrograph, and mark on the 4046 plate or enlargement all lines excited by 4077. This can be done by simply making two marks on a plate or slip of paper separated by the distance between 4046 and 4077 and then fitting the left-hand mark to all of the strongest lines; if the right-hand mark coincides with a much fainter line the chances are that this was excited by 4077. The distance between the two marks must be slightly reduced as we move up the spectrum, of course. Examining spectrum *a*, we see at once that line 177 obtained with 4358 excitation would not be found with 4046 excitation as it is in coincidence with 4077. In general we can be fairly sure that lines which appear in both spectra are real. For example, in spectrum *a* one might easily come to the conclusion that lines 1161 and 997 were merely the 4348 excitation of lines 1221 and 1052 (excited by 4358) but they are found also in the lower spectrum, though much narrower, as in spectrum *a* they are fused with the lines excited by 4348. In addition to the lines shown there is a group at frequency difference 3050 in the case of many organic compounds. For the study of this group (shown at λ 4400, Fig. 3), excitation by 4046 is all that is necessary, as the band then occurs in a region free from mercury lines and of high photographic sensitivity. With 4358 excitation, however, the band falls at a point of very low sensitivity just above the two mercury lines at 4915.

The spectra shown on Plate 8, Fig. 2, are as follows:

a. Raman spectrum of ortho-xylene excited by Hg 4358. Filter of noviol glass and sulphate of quinine. Line $\Delta\nu = 860$ not previously recorded marked with white *w*.

b. Ortho-xylene spectrum excited by 4046–4077. Filter of noviol glass and iodine in carbon tetrachloride. Lines excited by 4077 marked with a white *v*. $\Delta\nu = 860$ appears here also. The 4358 mercury group is at extreme right. Lines marked *H* are mercury.

c. Spectrum of toluene excited by 4358. Note doublet at right recorded as single line by previous observers. This doublet also appears in *a* and *b*. Quinine and noviol filter.

d. Toluene spectrum excited by 4046–4077. Lines due to 4077 marked *v*, faint mercury lines marked *H*. Iodine and noviol filter.

The frequencies or rather wave-numbers of the chief mercury lines concerned in the excitation of Raman spectra are given in

the following table, the lines being starred in proportion to their intensity.

λ	Wave-Numbers	λ	Wave-Numbers
3650	27388***	4339	23039*
3655	27353**	4348	22995**
3663	27293*	4358	22938***
3906	25592*	4916	20336
3984	25098		
4047	24705***	5461	18308***
4078	24516**	5770	17328**
4108	24335	5791	17265**

Relation of Raman Lines to Infra-Red Absorption. — If the Raman lines are simply lines formed by the combination of the natural frequencies of the molecular resonators with the frequency of the light we should be able to determine the resonator frequencies, *i.e.* the infra-red absorption lines or bands from the observed Raman frequencies. This can be done to a certain extent only, for we may have Raman lines which are not associated with any absorption lines and *vice versa*, and there seems to be no simple rule governing relative intensities. In the following table are given the early λ observations on benzene and quartz, the Raman lines expressed in wave-length values (μ) of the infra-red absorption lines with which they are related.

Benzene				Quartz			
Int.	λ Ra.	λ Abs.	Int.	Int.	λ Ra.	λ Abs.	Int.
5	16.5 μ	12.95	3	0	118		
3	11.7	11.8	4	0	94		
10	10.1	10.3	4	5	80	78	0
	——	9.75	10				
2	8.5	8.67	8	6	48		
0	7.3	——		2	38.2	38	1
	——	6.75	10	1	28.5		
4	6.28	——				26	4
3	6.23	6.2	0	1	24.7		
	——	5.5	2	10	21.5	21	10
	——	4.8	2	2	14.2	14.8	1
	——	4.2	3				
5	3.27	3.25	10	0	12.5	12.5	4
	——	2.49	6	1	8.6	8.5–9	9

Theory of the Raman Effect: Diatomic Gases. — Polar Molecules: In considering the theory of the Raman effect it will be best, as in most optical effects, to begin with the case of gases. The Raman spectrum of HCl, reported by the author in *Nature*,

February 2, 1925, and the announcement of the spectrum of CO and CO_2 by F. Rasetti in the following number of the same journal were the first observations made with gases. Hydrochloric acid is of especial interest as its infra-red absorption spectrum is well known and as it is a polar molecule (+ −) with a single vibration frequency of its two atoms. Its spectrum at atmospheric pressure was obtained with a glass tube nearly two metres in length illuminated by a long Cooper-Hewitt mercury arc placed in contact with it, both being surrounded with reflectors of thin sheet aluminum.[1]

The three prime requisites for obtaining the effect with gases at atmospheric pressure are: very intense illumination of a long column of gas, an absolutely black background, and complete absence of scattered light from the front window. The last condition is the most important and the most difficult to fulfil.

The apparatus employed consisted of a tube of soft glass 5 cms. in diameter and about 150 cms. in length, drawn off into a tapering cone at the rear, and constricted and expanded at the front in contact with a very long Cooper-Hewitt mercury arc lamp of glass.

The reflectors were made by wrapping rectangular sheets of the metal around a glass tube of much smaller diameter than that desired, a size being chosen such that the edges of the sheet remained in contact when it expanded by its elasticity.

The cylinders are slipped over the tubes by springing them open with the fingers, and on releasing them they clamp themselves firmly around the two tubes.

The Raman spectrum of the gas obtained with this apparatus was found to consist of a single line, the frequency difference between which and the unmodified line agreed very exactly with the "missing line" of the band absorption spectrum at 3 μ, while the unmodified line was accompanied by faint companions, for which the frequency difference was equal very nearly to the frequencies of alternate lines of Czerny's [2] pure rotation spectrum in the remote infra-red. The strong line was identified by Kemble and by Dieke independently as a Q branch ($\Delta j = 0$) of superposed or unresolved lines. These results were confirmed by a subsequent study with a more powerful spectrograph by Wood and Dieke.[3]

An enlargement of the entire spectrum is reproduced on Plate 8, Fig. 4, and a highly enlarged portion of the part showing the Q branch excited by 4046 and the rotation bands to the right and left of the 4358 line.

We will now examine in some detail the relation between the

[1] R. W. Wood, *Phil. Mag.*, Apr., 1929.
[2] *Phil. Mag.*, March, 1929.
[3] *Phys. Rev.*, *35*, 1355, 1930.

absorption and Raman lines of HCl and study the energy transitions involved, for by so doing we obtain a very clear notion of the fundamental principles involved in the effect, and see clearly why certain strong Raman lines (Q branch for example) do not correspond in their $\Delta\nu$ values with absorption lines, while the frequencies of many absorption lines are not represented as frequency differences $\Delta\nu$ in the Raman spectrum (alternate Czerny lines). The temperature effect will also be brought out.

The levels for HCl are well known from the experimental study of its absorption spectrum. They are shown in Fig. 275.

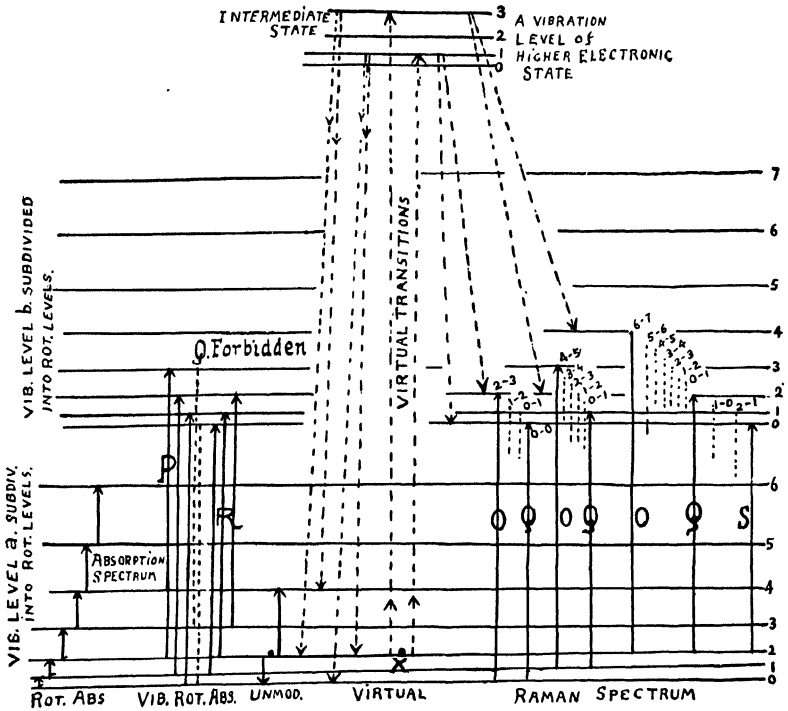


FIG. 275 (P and R should be transposed)

Two vibrational levels "a" and "b," each subdivided into rotational levels, spaced by assigning integral values to m in the expression $B(m+\frac{1}{2})^2$. We know from experiment that absorption transitions involving a change of both vibrational and rotational energy, and giving the band at 3.4 μ (see page 149) involve a change of rotation $\Delta j = \pm 1$, *i.e.* a molecule on rotation level 2 can absorb radiation of either one of two frequencies and be raised thereby to either level 1 or 3 of vibration level b (P and

R branch arrows in Fig. 275). The transition $\Delta j = 0$ is not allowed, as we know from the absence of the Q branch (missing line). There are also absorption bands in the remote infra-red, discovered by Czerny, involving rotational transitions only, shown by the arrows at the left.

We now require the selection rule governing the Raman effect, *i.e.* we must determine the Δj values from our energy diagram.

Selection Rule and Virtual Transitions. — Theory and experiment both show that Raman transitions are possible between two given energy levels, only when there is a third level with which both can combine, *i.e.* between which transitions are allowed. This third or intermediate state is represented at the top of Fig. 275 as a higher electronic state with vibration and rotation similar to those of the lower state. To determine the Raman transitions we plot $\Delta j = \pm 1$ transitions from given lower states of "a" to the upper electronic state, which, for molecules with 2 quanta of rotational energy are represented by dotted arrows from point X in Fig. 275. From the upper levels thus reached we plot downward transitions (following the same rule $\Delta j = \pm 1$) to the rotation levels of vibration levels a and b, thus determining the rotational levels of "a" or "b" to which molecules on the levels of "a" can be raised in the Raman effect. The left-hand set of dotted arrows represents virtual transitions for the case in which the Raman transitions for rotational energy only are involved (unmodified line and companions), the right-hand set transitions involving a change of vibrational energy as well (Q branch line and companions). For our molecule in rotational state 2, we find that these allowed transitions are to the rotational states 0, 2 and 4, therefore the selection rule for the Raman effect (determined by the dotted arrows, which are called virtual transitions) is that the rotational energy j changes by 0 or ± 2 instead of ± 1 only. We see at once that the Raman transition $\Delta j = 0$ is allowed, giving us a Raman line corresponding to the absent Q branch of the absorption spectrum. It is practically a line *since* the Q branch lines are nearly superposed, the transitions $0-0$, $1-1$, $2-2$, being very nearly equal, as a result of the circumstance that the upper vibrational level b is nearly a replica of a. The slight difference causes a wing on the Raman line, observed in the photograph. It must be understood, however, at the start that these virtual transitions are not to be taken literally in the case of the Raman effect, where a single process only is involved, but the structure of the mathematical theory suggests the introduction of the third states as a convenient way of deriving the selection rules.

We will consider first the line of unmodified wave-length with its

companions, as found experimentally. The virtual transition arrows, governed by the usual selection rules $\Delta j = \pm 1$ show that a molecule in lower rotational state 2 may make Raman transitions to $j=4$, $j=0$ or remain in state $j=2$, the first and second giving companions of higher and lower frequencies respectively, the latter the line of unmodified wave-length. These transitions, it must be noticed, occur within the lower vibration levels, *i.e.* only changes of rotation energy are involved.

The lengths of the two short arrows in the figure, one pointing up, the other down, represent the frequencies added to and subtracted from the exciting frequency (unmodified line) to give the companion lines, that is the Raman frequencies. Now these $\Delta j = \pm 2$ transitions do not occur in the absorption spectrum, but it is obvious that there are real absorption transitions that have a corresponding frequency, for example the Raman Δj transitions $2-4$, and $2-0$ have the same frequency differences as absorption transition $6-7$ and $2-3$. This follows from the law of spacing for the rotation levels, the distances above the ground level being given by $B(j+\frac{1}{2})^2$, j taking integral values. The distances are thus (if we write $B=1$)

1.	2.25	3.	12.25	5.	30.25	7.	56.25
2.	6.25	4.	20.25	6.	42.25		

and the differences between 6 and 7, and 2 and 4 are equal (14). The Raman and infra-red transitions are shown in Fig. 276 which

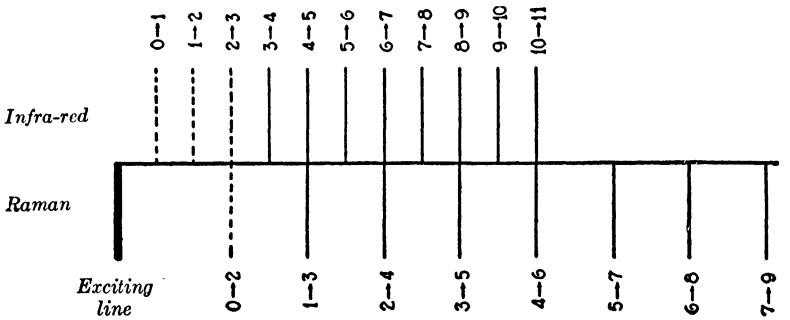


FIG. 276

shows that the companions of the line of unmodified wave-length have $\Delta\nu$ values equal to the frequencies of the alternate lines of Czerny's infra-red rotation spectrum. We will now consider the case of the Raman Q branch line found experimentally. This corresponds very nearly in position to the missing line of the 3 μ band, the transitions responsible for which band occur between

the rotational levels of "a" and "b," the lower and next higher vibrational levels, and are shown by the upward arrows of the P and R branches of the band, midway between which the $\Delta\nu$ value of the Raman line was found. It thus corresponded to the Q branch, for which the transitions are $\Delta j = 0-0,\ 1-1,\ 2-2$, etc., and as these are very nearly equal the corresponding lines are superposed, giving the strong line observed experimentally. The selection rule is shown by the dotted arrows at the right, from the intermediate electronic level to the rotation levels of the upper vibration level "b," and the Raman $\Delta\nu$ transitions are given for molecules with 0, 1 and 2 quanta of rotational energy. For those on the ground level we have the transition $0-0$, giving the Q branch line, and $0-2$ a companion on the short wave-length side. This is marked $2-3$ in the figure, this being the absorption transition of equal frequency value. Between these we see that there are two lines of the R branch of the 3 μ absorption band $0-1$ and $1-2$. For molecules on level 1, *i.e.* having one quantum of rotational energy we have Q line $1-1$ and companion $1-3$, with 4 lines of the R branch between, while molecules on level 2 yield, in addition to the Q line a high frequency companion $2-4$, and a low frequency one $2-0$ with six infra-red lines between in the first case and two in the second. These companion lines (resembling P and R branches) are called O and S branches instead, the transitions being $\Delta j = \pm 2$ instead of ± 1.

We see by measuring that the $\Delta j = +2$ Raman transition $2-4$ is equal to the $\Delta j = +1$ absorption transition, as in the case of the unmodified line, and that Fig. 275 is applicable to the present case also. It is also clear that molecules on the O level give a Raman line corresponding to the third absorption line of the P branch, those on levels 1 and 2 giving lines corresponding to the fifth and seventh absorption lines. With the gas at a higher temperature, we should thus have the lines which are far removed from the Q line of increased intensity as more molecules would be initially in higher rotational states.

Summing up we find that the intensity of a Raman line which corresponds with a transition of the molecule from the level a to the level b, has no connection at all with the transition probability between these two levels, which would determine the intensity of the corresponding absorption line. In order to find the intensity of a Raman line we must consider virtual transitions with a certain third level c. If this level c can be reached in an ordinary transition from both the levels a and b, the Raman line is allowed. The Raman line frequency difference can therefore be considered as the difference between the frequency of two lines of the molecule, one

of which must be an absorption line. The theory gives further the exact intensities of the lines expressed in terms of the transition probabilities between the different levels of the molecule, but we need not go here into details of the theory.

Non-Polar Molecules. — The molecules of hydrogen, nitrogen and oxygen are non-polar, and their vibrational energy is therefore uninfluenced by the alternating electric fields of light. On this account they show no vibration-rotation absorption bands in the infra-red, such as are found with polar molecules such as HCl, CO and CO_2. Their Raman spectra were obtained by Rasetti [1] who employed a quartz tube containing the gas under a pressure of 10–15 atmospheres and illuminated by the light of a low pressure, water-cooled mercury arc, with the current deflected by a magnet to give a very intense radiation of the 2536 line. His spectrograms showed the unmodified line accompanied by a most beautiful series of rotation bands (Figs. 277, 278). The same transition rela-

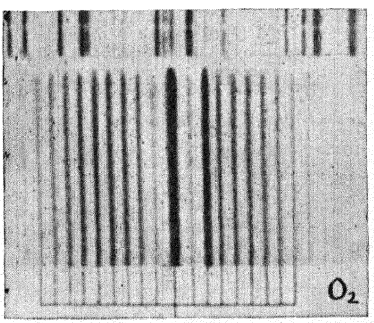


FIG. 277

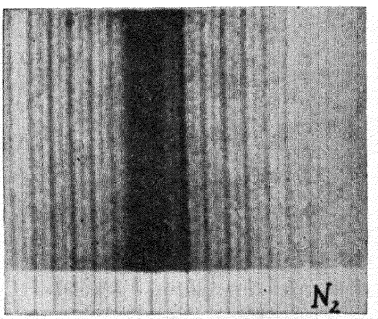


FIG. 278

tions, given for HCl, hold in this case also, the lines corresponding to rotational changes of 0−2, 1−3, etc. The fact that the alternate lines are faint in the case of nitrogen, shows that in the initial state of the gas, molecules with rotational energies corresponding to odd number levels are more numerous than those corresponding to even number levels, consequently transitions 1−3, 3−5, etc., occur less frequently than 0−2, 2−4, etc. This is related to nuclear spin employed in explanation of the iodine resonance band with alternate missing lines, Plate 13. In the case of oxygen Rasetti placed a basin of mercury in the spectrograph, the vapor of which absorbed the light of the unmodified line and prevented the halation due to overexposure which otherwise obscured the first two lines of the series. The strong line in coincidence with the second rotational line is mercury 2534.

[1] *Phys. Rev.*, *34*, 367, 1929.

McLennan and McLeod [1] obtained Raman lines from liquid oxygen, nitrogen and hydrogen, which are non-polar molecules and show no infra-red absorption, that is a to-and-fro vibration of the nuclei is not excited by radiation of any frequency. They each have, however, a natural frequency which, if it could be excited would give an infra-red band, and this frequency can be determined from their Raman spectrum. They have band groups in the ultra-violet due to transitions from an upper electronic level to the various vibration levels of the lower electronic level, as was found by Dieke and Hopfield. Now McLennan and McLeod found that the frequency differences for the Raman lines was the same as that for the ultra-violet bands. Their results are given in the following table:

	Excit. Line	Raman Line	$\Delta\nu$ Ra.	$\Delta\nu$ U.V.
	4047	4317	1552	1552
O_2			1555	
	4358	4674	1551	1552
	4358	5027	3049	3085
	4047	4469	2335	2331
N_2			2331	
	4358	4849	2322	2331
	4047	4980	4632	4633
	4358	4427	354	347
H_2	4358	4473	588	578
	4047	4864	4149	4159
			4162	

The underscored values were found by Rasetti [2] for the substances in the gaseous state. The smaller values in the last column correspond to vibration transitions 0 to 1 (fundamental), the larger to transitions 0 to 2 (first overtone). These are the first cases in which the first overtone is excited in a scattering process, involving the surrender of two quanta of vibrational energy to the scattering molecule.

Triatomic Molecules. — These may have a linear or triangular structure, the former being non-polar, the latter in general polar. Determinations of the dielectric constant distinguish between the two. We will begin with the simplest case, that of N_2O, which, according to present theory, has the structure N—N—O, but which, in order to simplify the treatment, we shall assume as N—O—N, the older model. Further along the Raman spectrum will be discussed in the light of fuller, more careful determinations

[1] *Nature, 123*, 60, 1929.
[2] *Phys. Rev., 34*, 367, 1929.

of the absorption lines, and we shall see why the assumed formula is incorrect. In this way will be brought out how our knowledge of molecular structure gradually improves as new experiments are made. The Raman spectrum of the gas was found by Dickinson-Dillon and Rasetti [1] to consist of a single line $\Delta\nu = 1282$; while McLennan-Smith-Wilhelm observed in addition a faint line 2224 in the liquefied gas. Its absorption and Raman spectra are shown at the bottom of Fig. 280. With a linear triatomic molecule, three types of vibration are possible, as shown at the left of Fig. 279.

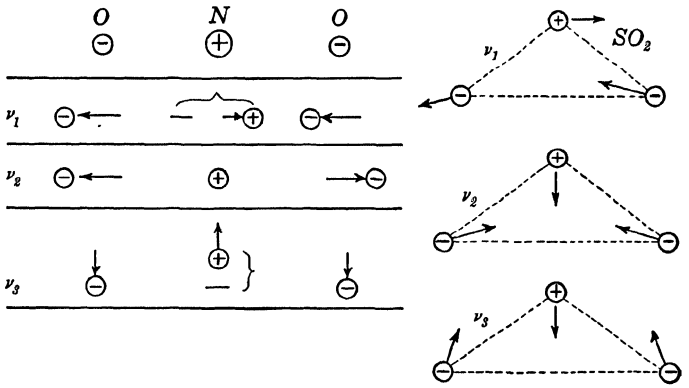


FIG. 279

One of these, with frequency ν_2 is a symmetrical expansion and contraction, the electrical centres of gravity remaining in coincidence. This frequency, therefore, does not appear in absorption or is optically inactive. The frequencies ν_1 and ν_3 appear in absorption, since the electric centres are displaced by the vibrations as shown by the figure in which the displaced centres are indicated by brackets: Vibrations can therefore be instigated by the alternating electric force of radiation. The optically inactive frequency may, however, give rise to absorption in combination with an optically active frequency, a weak absorption band occurring at frequency 3500, which is the summation tone of the inactive $\nu_2 = 1281$ plus the active $\nu_1 = 2223$. The energy levels corresponding to these three types of oscillation are shown in Fig. 280, absorption being designated by straight arrows, and the energy transition corresponding to the Raman line by a wavy arrow.

It will be noticed in the figure, in which the directions of motion for a given phase of the vibration is represented by arrows, that the + electrical centre is displaced to the right, and the −

[1] *Phys. Rev.*, *34*, 582.

to the left in the case of the ν_1 type, while in the case of the ν_3 type the + is above, the − below. The mathematical theory of the Raman effect shows that if we perform what is known as a symmetry operation, that is reverse the direction of the arrows, and find the polarity reversed, the frequency does not appear in the Raman effect. This is the case with ν_1 and ν_3 but not with ν_2,

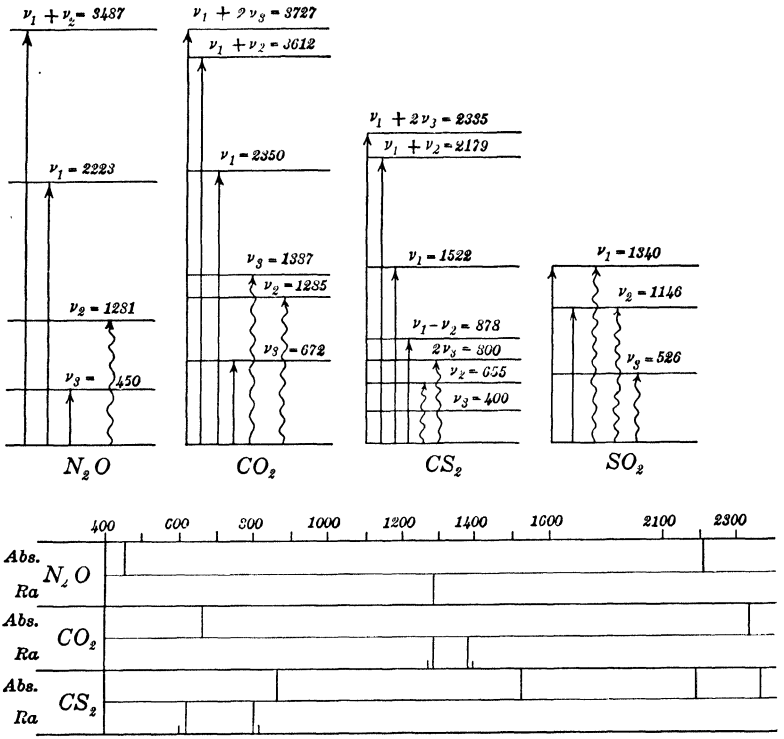


FIG. 280

which alone is permitted as a Raman line. In the case of the triangular molecule SO_2 reversal of the arrows does not reverse the direction of the polarity (except for amplitudes of impossible magnitude) and all three frequencies are consequently represented as Raman lines.

It is to be noticed that the Raman line 1281 is equal to the frequency difference between absorption lines 3500 and 2223, which shows us that the frequency $\nu_2 = 1281$, though optically inactive can be found by the Raman effect, which is thus seen to be capable of determining the frequencies of modes of vibration in the

molecule which do not respond directly to radiation of their own frequencies.

McLennan-Smith-Wilhelm [1] observed both 1292 and 2224 Raman lines in liquified N_2O. The presence of the latter which represents the ν_1 vibration is contrary to what is expected from the symmetry operation, and results from the circumstance that the real structure of the molecule is N—N—O as we shall see presently.

We will take next the case of CO_2 which has the structure O—C—O, and the same three types of vibration as N_2O. It gives a Raman spectrum of two lines 1387 and 1285 instead of the single line exhibited by N_2O. Its absorption spectrum consists of three bands $\nu_3=672$ or 14.9 μ, $\nu_1=2350$ or 4.28 μ and a double band 3727, $(\nu_1+2\nu_3)$ and 3612 $(\nu_1+\nu_2)$ at about 2.7 μ, neither Raman line being represented in absorption. Various attempts have been made to account for this spectrum. Eucken attributed 1387 to a transition from the ground level to the first ν_2 level (optically inactive) and 1285 to a transition from the latter level to the second ν_2 level, the difference in distance between the levels being the result of the anharmonic nature of the vibration. Fermi has shown, however, that this is incorrect, and accounts for the double line in the following way. The ν_2 vibration gives the 1285 line while the overtone has a frequency of 1387 and ordinarily would not appear in the Raman spectrum. Owing, however, to proximity of these two energy levels (ν_2 and $2\nu_3$), the vibrations are "degenerate" (Entartet), a term applied when we have the same or nearly the same frequency for two modes of vibration. As a result the transition forbidden by the symmetry operation is permitted and both lines appear. The frequencies of the two lines are the frequency differences between $\nu_1=2350$ and $\nu_1+\nu_2=3612$ and $\nu_1+2\nu_3=3727$, and since ν_1 is known from absorption, ν_2 the optically inactive frequency can be calculated as well as observed as a Raman line. The two Raman lines have faint companions as shown in the diagram below the energy levels.

Carbon disulphide, S—C—S, has a Raman spectrum similar to that of CO_2. The energy levels and transitions are shown in Fig. 280, the Raman lines being 655 and 800.

The absorption and Raman lines are given in the table on page 462, the former being determinations by Bailey and Cassie [2] while the assignment of the modes of vibration are by Placzek. The two Raman lines correspond to frequency differences between the infra-red lines 1522 and the lines 2179 and 2335, the former

[1] *Trans. Roy. Soc.*, Canada III, *24*, 197, 1930.
[2] *Nature, 126*, 350, 1930.

INFRA-RED AND RAMAN SPECTRUM OF CS_2

INFRA-RED	RAMAN	$\nu_3=400$
—	655	ν_2
—	800	$2\nu_3$
875		$\nu_1-\nu_2$
1522		ν_1
2179		$\nu_1+\nu_2$
2335		$\nu_1+2\nu_3$

giving the optically inactive frequency $\nu_2=657$, the latter the overtone 813 of the assumed frequency of $\nu_3=400$. Degeneracy in the case of 655 and 800 is considered as responsible for the appearance of the $2\nu_3$ line 800 which is much fainter than 655 as the coincidence between the two frequencies is less close than in the case of CO_2, its two Raman lines being of equal intensity. A faint companion has been given as accompanying the 655 line, and more recently 800 has been found to have also a companion,[1] making the analogy with CO_2 complete.

We will return now to the case of N_2O for which we gave an earlier and incorrect treatment for the sake of simplicity. The appearance of the $2\nu_3=800$ line of CS_2 is considered as resulting from degeneracy brought about by the $\nu_2=655$ vibration. There is very nearly as close a coincidence in the case of N_2O for which $2\nu_3=900$ and $\nu_2=1281$, and if N_2O has the same structure as CS_2 one would suppose that a Raman line of $\Delta\nu=900$ would appear, owing to degeneracy. The absorption spectrum has recently been more fully investigated by Plyler and Barker [2] who give the following assignment of frequencies:

		INTENSITY
ν_3	589	40
ν_2	1285	70
ν_1	2224	100
$2\nu_3$	1167	20
$2\nu_2$	2564	16
$2\nu_1$	4420	2
$\nu_2+\nu_3$	1868	0.5
$\nu_2+2\nu_3$	2462	6
$\nu_1+\nu_3$	2799	1.2
$\nu_1+2\nu_3$	3366	1.6
$\nu_2+\nu_1$	3482	10.
$2\nu_2+\nu_1$	4736	0.8

The fact that ν_2 appears in absorption is in itself sufficient to show that the structure previously assumed N—O—N cannot be the true one, as the vibration would be optically inactive. With

[1] R. W. Wood and G. Collins, *Phys. Rev.*, Nov. 1, 1932.
[2] *Phys. Rev.*, 1931.

a structure N—N—O a vibration of the ν_2 type would shift the electric centres, and give absorption. The $2\nu_3$ and ν_2 are now in closer agreement than on the previous assumption and it seems to be worth while to make a more careful search for a Raman line at 1167 ($2\nu_3$), though Dieke suggests that there may be no coupling between the frequencies in this case, or that the coupling may be different from that in the case of CS_2, ν_1 ought also to appear in the Raman spectrum.

We will now consider the strongly polar molecule SO_2 which is known to be triangular in shape with three modes of vibration as shown in Fig. 279. All three are optically active, the displacement of the electric centres resulting from vibrations being least for ν_2. It will be observed that the modes of vibration are analogous to those of the linear molecule, but if we perform the symmetry operation for each, we find that the direction of the polarity is not reversed as is the case with the ν_1 and ν_3 vibrations of the former type. All three frequencies consequently appear in the Raman spectrum, the ν_2 frequency being much the most intense, as in the linear molecule it is the only frequency that can appear. The absorption and Raman transitions are shown in Fig. 280.

Tetrachlorides. — We will consider briefly one more case, the Raman spectra of carbon, silicon, titanium and tin tetrachlorides. In this case we have a carbon or metal atom at the centre of a tetrahedron with the chlorine atoms at the corners, and there are two optically inactive vibrations, one a symmetrical expansion and contraction of the tetrahedron, and the other a to-and-fro or rotatory movement of the chlorine atoms on the surface of a sphere: in neither case is there a displacement of the electric centres.

Carbon tetrachloride, a liquid which has been very extensively investigated, shows the following Raman lines: 219, 314, 459, a doublet 758–789, and a very faint line, missed by many observers, 1539. The first point to be noticed is that the mean of the doublet 773 is equal to the sum of 459 and 314, an exact accidental degeneracy. It is on this account that 773 splits into a doublet, as called for by theory. Placzek and van Wijk [1] showed that the components are of equal intensity, which proves exact coincidence of the frequencies. Furthermore the 1539 line has the double frequency of 773 and is therefore an overtone of that vibration. Dennison considers that we have four ground vibrations, two of them optically inactive, $\nu_2 = 219$ and $\nu_1 = 459$, the latter corresponding to the expansion of the molecule, the former to the rotatory movement referred to above. The other two types of motion in

[1] *Zeit. für Phys.*, *67*, 582.

which all five atoms are probably concerned have not been determined.

Anti-Stokes lines are especially strong in carbon tetrachloride, three being usually recorded. The double line has recently been found by the author.

The frequencies of six molecules of this type are given in the following table:

	ν_2	ν_4	ν_1	ν_3
CCl_4	214	311	450	773
	0.75	0.8	0.1	0.75
$SiCl_4$	148	220	422	608
	1.	0.9	0.14	1.
$TiCl_4$	119	139	386	491
	0.8	0.8	0.14	0.5
$SnCl_4$	104	136	367	401
	1.	1.	0.38	0.8
CBr_4	123	183	265	667
$SnBr_4$	64	88	220	279

The depolarizing factors are given below each ν value.

Relation between Raman and Resonance Spectra. — Though it is perhaps doubtful whether anyone has yet obtained evidence of a gradual transition from a Raman spectrum to resonance radiation, or a resonance spectrum by lowering the frequency of the exciting radiation until it coincides with an absorption frequency, it is nevertheless instructive to compare the two types of emission in their relation to energy levels. In Fig. 281 we have the three energy levels of thallium discussed in the Chapter on Resonance Radiation. The unexcited or normal atom is in state E_k and is able to absorb the frequency ν_{kx} as the transition $E_k - E_x$ is allowed. From the upper level E_x the electron may return to E_k yielding resonance radiation, or to E_l, leaving the atom in an excited state with its energy increased by an amount corresponding to frequency ν_{kl}, and emitting the thal-

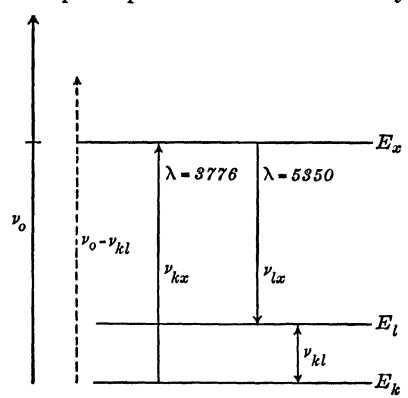


FIG. 281

lium green line 5350 (resonance spectrum or fluorescence). If we compare these transitions with those obtaining in the Raman effect we shall see that this green line is the counterpart of a Raman line, and that the frequency ν_{kl} represents the "forbidden" absorption transition $E_k - E_l$ ($7790 cm^{-1}$ or 0.97 μ).

Suppose now that the exciting radiation ν_0 is of higher frequency corresponding to the arrow at the left of Fig. 281 and we consider real transitions, just discussed, as *virtual transitions*, as understood in the theory of the Raman effect. Energy corresponding to ν_{kx} is abstracted from the impinging light quantum (virtual transition $E_k - E_x$) while energy ν_{lx} is given back, leaving the atom in an excited state and the light quantum with a reduction of energy corresponding to ν_{kl}, giving the Raman line of frequency $\nu_0 - \nu_{kl}$ represented by the dotted arrow at the left, a line of somewhat shorter wave-length than the green thallium line.

If now ν_0 is gradually diminished, the Raman line moves with the exciting line towards the green, coinciding with the thallium line when $\nu_0 = \nu_{kx}$, *i.e.* $\nu_{kx} - \nu_{kl} = \nu_{lx}$ or the green line. The absorption line ν_{kl} corresponding to this Raman line is, however, absent as the transition $E_k - E_l$ is not allowed. To obtain anti-Stokes Raman lines we must suppose the atom initially in state E_l, the virtual transitions occurring in reversed order, the smaller energy ν_{lx} abstracted from and the larger ν_{kx} given to the light quantum, yielding a Raman line of frequency $\nu_0 + \nu_{kl}$. The complete theory of the Raman effect indicates that as ν_0 approaches either of the transition frequencies mentioned the intensity of the Raman line will increase reaching its maximum value for coincidence.

To show this experimentally with thallium is probably impossible for the vapor would have to be at a very low pressure, involving a tube of enormous length, illuminated by a group of lines of wave-length slightly less than 3776. Rasetti [1] describes a series of nearly equidistant doublets between 2000 and 4000 A.U. in the spectrum of oxygen excited by the 1849 line of a quartz-mercury arc. These doublets seemed analogous to the doublets of the resonance spectrum of iodine described by the author, which are obtained only with the vapor at very low pressure, and with an exciting radiation (Hg 5461) coinciding with an absorption line. From the circumstance that Hg 1849 does not coincide exactly with an oxygen absorption line, and that the oxygen was at atmospheric pressure, Rasetti concluded that he was observing the transition from a Raman to a resonance spectrum.

This conclusion is criticized by P. Pringsheim on the ground that 1849 is a very broad line in the mercury arc as used by Rasetti, and

[1] *Proc. Nat. Acad.*, *15*, 411, 1929.

may easily have encroached on an absorption line, and that sulphur vapor at 250 mms. gives a normal resonance spectrum. The author has photographed oxygen absorption lines in the "wing" of the 1849 line of a quartz-mercury arc operated at a moderate temperature (30 seconds after starting).

It seems useless to look for the effect with any of the media commonly investigated for Raman effect as they have no sharp resonance frequencies. It is important to note this difference in the two cases just considered. In the case of resonance radiation and resonance spectra, we regard the atom or molecule as the source of the radiation of altered frequency, and there may be, and usually is, a finite and measurable time between absorption and emission. In the Raman effect we are dealing with "scattering" and not with absorption. The Raman line frequencies are not identical with the natural frequencies, but only to frequency differences, which we regard as abstracted from, or added to, the frequency of the exciting light. As was said before the virtual transitions are not to be interpreted literally, but only as indicating what values of $\Delta\nu$ can be compounded with the primary frequency.

Frequency Changes in the Unmodified Line. — That the line of unmodified wave-length is less homogeneous than the exciting radiation was noticed very early in the study of the Raman effect. It is especially noticeable in the case of benzene in which case the wings on the lines are broad enough to be detected with spectroscopes of low power. In the case of excitation by Hg 4358 the presence of the two strong mercury lines on the short-wave-length side mask the effect, but it is very conspicuous on the other side. It can be seen in the spectrograms made with helium excitation on Plate 8.

Cabannes and Daure [1] working with a Fabry-Perot interferometer found a displacement towards the red of the line of unmodified wave-length of from .01 A.U. for butane vapor, to .035 A.U. for prophyl alcohol. They found also that the rings disappeared when only that part of the scattered radiation polarized parallel to the exciting beam was utilized. (The polarization of the Raman lines will be considered presently.)

This is the portion that produces the depolarization of the Tyndall or Rayleigh scattering, and results from fluctuations in the orientation of anisotropic molecules, as was shown in the Chapter on Scattering. It can, therefore, be regarded as the factor preventing interference in the Fabry-Perot instrument, or in other words the factor responsible for the broadening of the line. The portion polarized perpendicular to the beam shows much less broadening.

[1] *Comptes Rendus, 186,* 1533.

Vacher [1] repeated and confirmed these observations finding in addition that the broadening increased with the angle of scattering, .001 for 30°, .0087 for 90° and .013 for 150°. He concluded that the unmodified line consisted of two parts, the component of short wave-length being unchanged, while the other was shifted towards the red.

Influence of Density on Intensity Ratio of Modified and Unmodified Scattering. — If the scattered radiation is incoherent, *i.e.* if no definite and permanent phase relation exists between the disturbances sent out from the centres of scattering, then the scattered intensity in lateral directions will be the sum of the intensities due to the individual centres, and the more closely these are packed together, that is, the greater the density of the medium, the greater will be the intensity of the scattered radiation. Now the modified scattering is always incoherent.

The modified frequencies (Raman lines, wings bordering the exciting line, etc.) therefore appear with intensities proportional to the density of the medium. This is not the case with the radiations of unmodified frequency, which is coherent, and since interference can take place in lateral directions, as imagined by Fresnel to account for the rectilinear propagation of light, the total intensity will not be the sum of the intensities of the individual centres.

The destructive interference in lateral directions will become more and more effective as the density of the medium is increased, *i.e.* the more nearly it becomes homogeneous and continuous, therefore as we increase the density, the increase in the intensity of the scattered radiation due to the increasing number of emitting centres is partially offset by destructive interference, the energy thus lost laterally going along with the exciting waves, as waves of the same type, as we shall see in the Chapter on Refraction Theory.

In the case of a gas at low pressure, there is little destructive interference, and we have the unmodified lines strong in comparison to the modified. With liquids and solids, especially when they are at a low temperature, the modified radiation is relatively strong, for by increasing the density we have introduced the element of destructive interference for the unmodified radiations.

In the case of crystals, where the molecules are not only densely packed, but also arranged in regular patterns on a lattice, the destructive interference reaches its maximum value, and the unmodified radiation may almost disappear. In the absence of thermic agitation it would vanish entirely.

Polarization of Raman Radiations. — The discovery that many of the lines of modified frequency were strongly polarized was

[1] *Comptes Rendus, 191*, 1121.

announced by Raman [1] in one of his earliest publications. The degree of polarization of the Raman lines varies in any given spectrum, being strongest for the optically inactive frequencies.

The depolarizing factor p (see Chapter on Scattering of Light) for a number of compounds from determinations by Bhagavantam are given in the following table. Complete polarization is indicated by $p=0$, no trace of polarization by $p=1$.

Benzene	605	849	992	1178	1584	1605	3046	3061		
	1	1	.05	1	1	1	1	.35		
Cyclohexane	800	992	1028	1268	1445	2853	2890	2937		
	0.1	0.1	0.3	0.4	0.5	0.13	0.5	0.1		
Pentane	402	764	843	867	1154	2857	2873	2915	2936	2962
	0.5	0.5	0.5	0.5	0.3	0.7	.22	0.7	.5	.75

The values of the depolarizing factor for the tetrachlorides are given in the table in the preceding section.

The polarization measurements are made by the method described in the Chapter on Scattering of Light. In some cases the liquids have been illuminated with polarized light, and in others with natural light.

The line 992 which is common to benzene and cyclohexane is very strongly polarized and is absent in the absorption spectrum which indicates that it is due to a symmetrical vibration which does not change the electrical moment.

Cabannes [2] has reported very interesting polarization experiments with quartz, certain Raman lines showing reversed polarization. Lines 465 and 308 (21.5 μ and 48 μ) were polarized in the usual manner, while lines 262 and 800 (38 μ and 12.5 μ) were polarized at right angles to this direction, the effect being independent of the orientation of the crystal. In the case of the line 125 (80 μ) he found a dependence on orientation, the polarization being reversed (*i.e.* with electric vector parallel to the exciting beam), when the optic axis was perpendicular to the beam and to the observation direction, while with the axis parallel to the beam there was complete depolarization. This reversal of polarization is analagous to the remarkable discovery made by Hanle [3] and Bar [4] that when certain liquids were illuminated with circularly polarized light and observations made parallel to the light, certain lines were circularly polarized in the same sense as the exciting radiation, while other lines (those normally showing strong depolarization) were circularly polarized in the reversed sense.

[1] *Ind. J. of Phys.*, *2*, 387, 1928.
[2] *Comptes Rendus*, *185*, 249, 1929.
[3] *Naturwiss.*, *19*, 375; *Phys. Zeit.*, *32*, 556.
[4] *Naturwiss.*, *19*, 463.

CHAPTER XV

THEORY OF DISPERSION AND SELECTIVE REFLECTION

Theory of Refraction and Dispersion. — The first attempt to formulate a theoretical explanation of the circumstance that the retardation of light by a transparent medium increases with decreasing wave-length was made by Cauchy in 1836, employing the elastic solid theory of light. The velocity of propagation of transverse waves in an elastic medium is proportional to the square root of the elasticity divided by the density of the medium for all cases in which the wave-length is large in proportion to the distance between the particles constituting the vibrating system, or in the case of light, the distance between the ether particles. As the wave-length of the light is less in matter than in free ether Cauchy based his treatment on this shortening of the light-waves and deduced the dispersion formula which goes by his name, in which we have the refractive index $n = A + B/\lambda^2 + C/\lambda^4$.

If the constants A, B and C were determined from observations of n for three widely separated wave-lengths, the formula was found to represent the dispersion of most substances with considerable accuracy. Biot, however, pointed out that the change of wave-length was too small to be the factor contributing to dispersion, and that dispersion should occur in free ether if Cauchy's theory were correct.

Neuman in 1841 introduced the idea of a reciprocal action between the ether and material particles and O'Brien in 1844 made use of absorption, and missed by a narrow margin a prediction of anomalous dispersion, which was discovered by Le Roux in 1862, and promptly forgotten, until some time after its rediscovery by Christiansen and the investigations of Kundt in 1870 and 1871.

The discovery of anomalous dispersion, and the relation existing between absorption and dispersion, put the matter in a new light. The refractive index of a medium, which exercised strong selective absorption, was found to increase rapidly as the absorption band was approached from the region of longer wave-lengths. This made it seem extremely probable that the dispersion of so-called transparent media was due to absorption bands in the ultra-violet; in other words, that there was no essential difference between normal

and anomalous dispersion, the former being only a special case of the latter, the observations being restricted to a range of wave-lengths too narrow to show any anomalies.

As was pointed out by Lord Rayleigh, the foundation of the modern theory of dispersion was in reality laid by Maxwell in the form of a question propounded in an examination paper.[1] The same idea subsequently occurred to Sellmeier, who has always been regarded as the founder of the theory. Sellmeier sought for the cause of dispersion in the vibrations of the atoms of the molecule caused by the repeated impacts of the light-waves. These atoms would naturally have free-periods of vibrations of their own, and would be set in motion by the light-waves exactly as a tuning-fork is set in vibration by waves of sound.

Sellmeier deduced a formula which is practically identical with a special case of the more recent electromagnetic dispersion formula, and which represents the dispersion for those wave-lengths for which the medium is comparatively transparent, *i.e.* on both sides of the absorption band. Within the region of absorption it breaks down, for reasons which will appear presently. Sellmeier's formula was as follows:

$$n^2 = 1 + \frac{D\lambda^2}{\lambda^2 - \lambda_m{}^2},$$

in which n is the refractive index for wave-length λ, and λ_m the wave-length in ether of light of the same frequency as that of the absorbing atom. If more than one absorption band exists, the formula takes the form

$$n^2 = 1 + \Sigma\frac{D\lambda^2}{\lambda^2 - \lambda_m{}^2},$$

the summation being taken for as many terms as there are atoms of different periods.

The trend of a dispersion curve traced by this formula is shown in Fig. 282 for a medium having two absorption bands. The portion ABC is the portion observed in the case of transparent media in the visible and ultra-violet regions, the part AB being alone represented by the Cauchy formula. The portion AX can be observed in the infra-red, while Cauchy's formula would predict Ab for the region.

For a medium in which the refractive index is governed by a single absorption band (very nearly true for sodium vapor as we shall see). In the case of very long waves the denominator of the

[1] *Camb. Calendar*, 1869, Math. Tripos Exam.

Sellmeier formula becomes approximately equal to λ^2, and we have

$$n=\sqrt{1+D},$$

which, if D is small, as in the case of a gas, does not differ much from unity. As λ decreases the value of the fraction $\lambda^2/\lambda^2-\lambda_m^2$

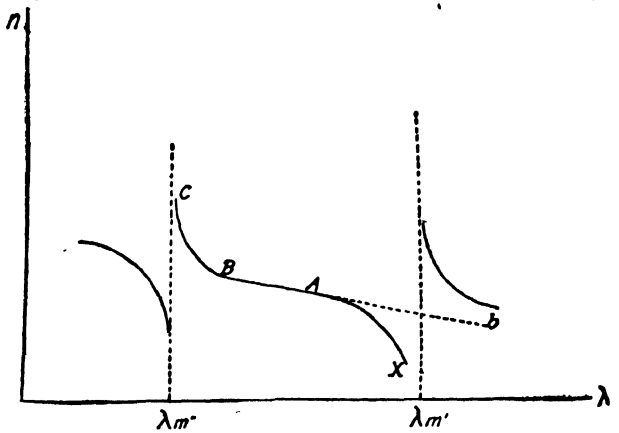


FIG. 282

increases, becoming infinite when $\lambda=\lambda_m$. For values of λ less than λ_m the sign of the term changes, and we have values of n, which are less than unity, the lowest values being for wave-lengths close to the absorption band; as λ decreases n increases, becoming unity for infinitely short waves. The form of the curve is shown in Fig. 283. As we shall see later on, the Sellmeier formula represents most perfectly the dispersion of sodium vapor, in which the refraction and dispersion are due almost entirely to the influence of a single absorption band (in reality a close double band).

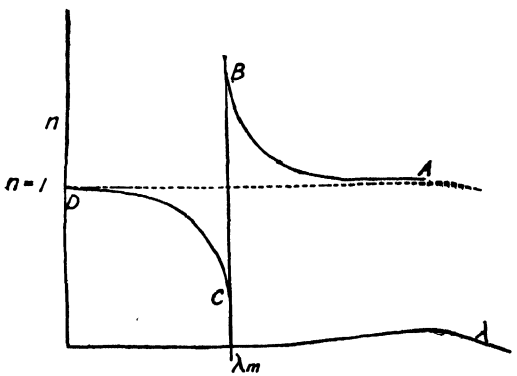


FIG. 283

If we confine our attention to the region AB of the curve, we see that the dispersion here is normal, the refractive index increasing with decreasing λ, and the curve convex towards the axis of

abscissae, as is the case with all ordinary transparent media for visible radiations. The decrease in refractive index occurs when we pass across the absorption band, in the above case the drop being very sudden. A further decrease in the value of λ causes an increase in n, the curve now being *concave* towards the axis of abscissae. The dispersion here is normal in that n increases as λ decreases, but it differs from the dispersion along the branch AB in that the rate of change of n with λ becomes less as λ becomes less, while along AB the rate of change of n with λ increases as λ decreases. This is due to the fact that in the former case we are receding from the absorption band, while in the latter we are approaching it.

The dispersion of glass or water, for example, is represented by a curve of form similar to AB, consequently we may infer that an absorption band in the ultra-violet is responsible for the dispersion. Curves of the form CD are never found for ordinary transparent media in the visible spectrum. If, however, the investigations are

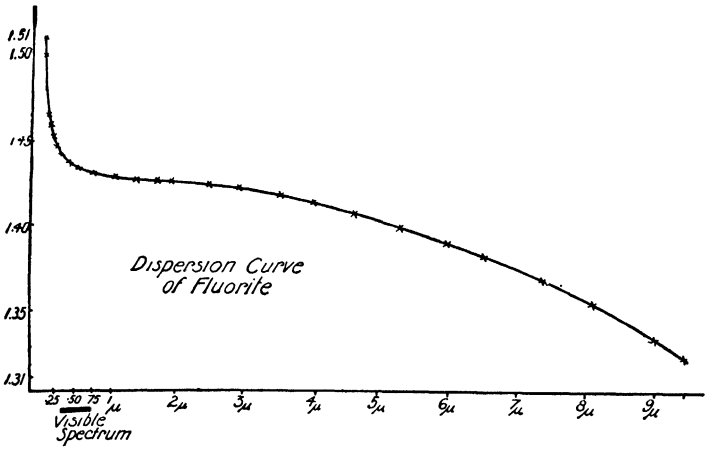


FIG. 284

carried into the infra-red region, we find frequently that the curve, which in the visible region of the spectrum was convex towards the axis of abscissae, eventually becomes concave in this direction, indicating that an absorption band is being approached. The curve for fluorite (Fig. 284) is an example.

The Sellmeier formula represents the refractive index of an absorbing medium to a high degree of approximation if not applied to the frequencies near the centre of the absorption band. If $\lambda = \lambda_0$ we obtain $n = \infty$.

Helmholtz was the first to employ two fundamental equations,

one representing the propagation of the light, the other the vibration of the resonators or absorbing atoms or molecules and showed that if a damping factor (or friction term as it was called) was inserted in the equation representing the oscillation of the resonator, the refractive index curve on the long-wave-length side of the absorption band rose to a maximum value and then descended joining the branch on the short-wave-length side, as shown in Fig. 285, the refractive index being unity for $\lambda = \lambda_o$.

Helmholtz assumed the medium made up of molecules between which the ether, made up of much smaller particles, freely penetrated. The atoms of the molecule were regarded as capable of vibration about a position of equilibrium, and were coupled to the ether particles by forces similar to those existing between the ether particles. Any displacement of the ether particles, as by a light-wave, causes a displacement of the atoms between the ether particles themselves. When a wave enters the medium, it is considered as propagated through the agency of the ether alone; that is, there is no direct propagation of a disturbance from molecule to molecule or from atom to atom. Helmholtz considered that the molecule remained at rest, but that the atom could be displaced from its position of equilibrium by the vibration of the ether, and when so displaced was drawn back by a force of restitution proportional to the displacement. Though the atoms are independent of each other, and each is free to vibrate by itself, they will, when disturbed in succession by a passing wave, have displacements which collectively form a wave-curve, just as chips floating on water waves, though not transmitting the waves, will be arranged in the form of the wave-curve though in this case there is no force of restitution tending to draw back the displaced chips.

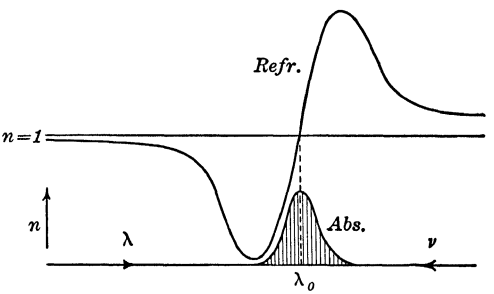


FIG. 285

To explain absorption Helmholtz considered that the vibration of the atom was accompanied by something analogous to friction, otherwise the energy taken out of the light would be given back to the ether and there would be no absorption.

The velocity with which waves are propagated through a medium can be determined in terms of the elasticity and density of the medium. The measure of the elasticity is the force of restitu-

tion exerted upon a displaced particle by the neighboring particle. In the present case, the displaced ether particle is urged back into its position of equilibrium not only by the forces exerted by neighboring ether particles, but also by the forces existing between the ether and the atoms. The atoms are, however, not stationary, but are set in motion to a greater or less degree by the waves.

The mathematical treatment showed that the phase relation between the vibrations of the atoms and the ether is a function of the frequency of the light in relation to the frequency of the atoms. On one side of the absorption band the phases are more or less in agreement, on the other side opposed. In one case the force of restitution of the displaced ether particle was considered as increased, and in the other diminished, and as the velocity of a wave is proportional to the square root of the elasticity divided by the density, the retardation of the light on the red side of the absorption band and the acceleration on the violet side was accounted for.

The modern treatments as applied to the electromagnetic theory bear a strong resemblance to these earlier treatments, but define more precisely the nature of the oscillators, the physical basis of the friction terms, and the manner in which the velocity of the wave is altered.

Free and Forced Vibrations. — Inasmuch as the more recent, as well as the earlier, dispersion theories involve the phase relationships between the vibration of the light and that of the atoms, ions or electrons, which, having natural free periods of their own, are thrown into forced vibration of the same frequency as that of the light, it will be well at the outset to devote a little space to the discussion of vibrations of this type.

The older theories of dispersion, based on the elastic solid theory of light, considered that the atoms in the molecule were capable of vibration in natural frequencies of their own, and when excited to forced vibration, by the periodic forces of the light, reacted on the vibration of the ether and thus modified the velocity of the wave. This reaction must obviously depend upon the phase relationship between the vibrations of the light and those of the atoms. The phenomenon can be studied to advantage by mechanical models. Pohl, in his recent book on mechanics and acoustics, describes a vibrator of the type of the balance wheel of a watch, excited to torsional vibrations by properly timed impulses from a motor of controllable speed, and subjected to a damping force of easily adjustable magnitude by a magnetic field which excites Foucault currents in the copper balance wheel. A simpler apparatus is a pair of pendulums, one of large, the other

of small mass, suspended from an elastic rod clamped in a vise. The large pendulum excites the smaller by impulses fed to it through the rod.

The object of making one of small mass is to eliminate its effect on the other. Two steel balls, of very different masses suspended by threads answer every purpose. The smaller can be damped by a wire attached to its base which dips to a variable depth in a glass of water. If the damping is small or equal to zero, time is required for the establishment of a steady state. Fig. 286 from Pohl's book shows the variation of the amplitude of the forced vibration with the frequency (as fraction of natural frequency ν_0) of the exciting force of constant amplitude.

The free vibrations with various values of the damping factor are shown in the upper insert at right. The ratio of two successive amplitudes of the free vibration is K the damping ratio, the natural logarithm of which is Λ the logarithmic decrement. These curves show that the amplitude of the forced vibration increases rapidly as the frequency of the exciting impulse approaches the natural period of the vibrator, passing through a maximum when the two periods agree. This maximum has its largest value for the condition of smallest damping (curve A corresponding to curve A in the inset). The phase relation between the two vibrations is shown by the lower part of Fig. 286, abscissae being the frequency ratio ν/ν_0, and ordinates the phase displacement in degrees.

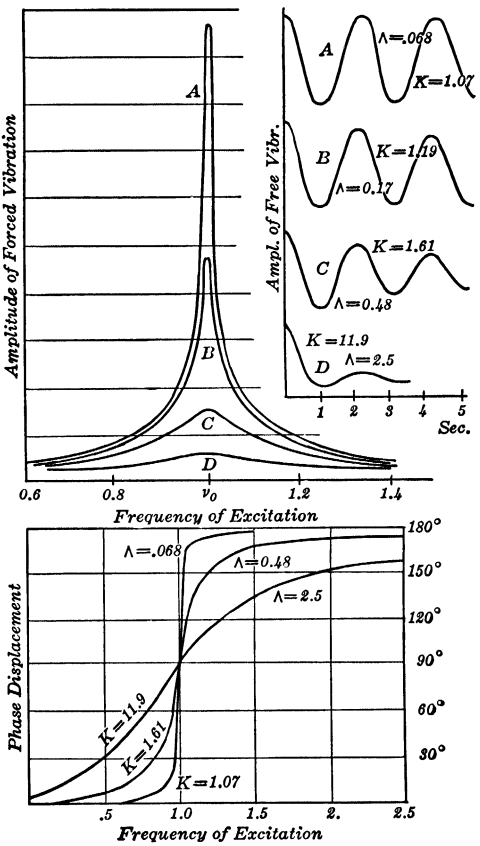


FIG. 286

We see that for small values of ν (ν_0 being the natural frequency) the phases agree, the two pendulums swinging together. As ν increases the vibrator lags behind the exciting force, the phase displacement being 90° when $\nu = \nu_0$ after which it increases with further increase of ν approaching (asymtotically) the value of 180°.

With a small damping factor the change of phase occurs most rapidly in the immediate vicinity of the value ν_0, while with heavy damping it is spread out over a greater range of ν values as shown by the three curves which were made by Pohl from observations on his damped balance wheel. With no damping at all there will be an abrupt change of phase from 0 to 180° at the point ν_0 that is when we pass through the point of exact resonance, or in optics the centre of the absorption band. This phenomenon can be observed by adding a second light pendulum to our elastic bar, making one slightly longer, the other slightly shorter than the heavy pendulum. After the system has settled down to a steady state the lighter pendulums will be seen to be swinging in opposite directions. Or we may show the phenomenon with a single pendulum, holding the cord in the hand. If we move the hand to-and-fro with a slower frequency than that of the pendulum, the pendulum will swing in phase with the hand while if the motion of the hand is at a higher frequency the pendulum swings in a direction opposite to that of the hand.

We can apply this result to the phenomenon of optical dispersion at once. Inasmuch as the modification of the velocity of light depends upon the phase relationship between the vibrations of the light and those of the atoms or electrons, it is clear that with small damping the modification of velocity will be appreciable only for frequencies in the immediate vicinity of the natural frequency (dispersion of sodium vapor at low pressure, for example) which will be discussed presently. Damping may result from radiation (scattered by the oscillators) or from collisions with other molecules.

In our optical problems the oscillators may be dipoles, that is molecules or atoms in which the "centres of gravity" of the + and − electrical charges (or more properly electrical centres) are not in coincidence. Vibrations will be excited in such a molecule by the alternating electric force of the light, and the molecule will reëmit or scatter the light. We will now consider this scattering.

Molecular Scattering. — In the Chapter on Scattering of Light we confined our study to the phenomena exhibited by the scattered radiations and we will now consider molecular scattering with regard to its effect on the propagation of the light-waves.[1]

[1] Relation of Scattering to ref. index first treated by Lord Rayleigh, *Phil. Mag.*, 5. 47, 1899.

Suppose the circle in Fig. 287 represents the molecule or dipole and that polarized light with its electric vector in the direction of the double-headed arrow is incident from the right. Scattered polarized waves will be emitted as shown with maximum amplitude in a plane at 90° to the electric vector of the incident light ($\theta = 90°$) and of zero amplitude in the two directions indicated by the dotted arrows for which $\theta = 0$. The amplitude in other directions is proportional to sin θ, θ being the angle between the scattered ray and the electric vector of the incident light at the origin (the dotted arrow).

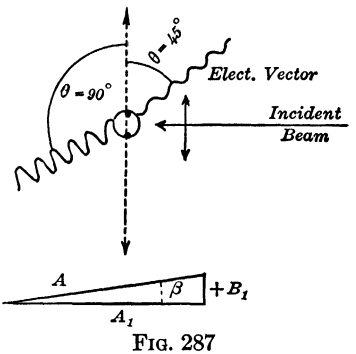


FIG. 287

Schuster's treatment of this scattering and its effect in modifying the velocity of the light, *i.e.* causing refraction, is as follows: Let the original light be defined by some vector $R \cos(\omega t - lx)$ in which x is measured from the molecule. The corresponding vector in the scattered light at a distance r may be written in the form $AR \sin\theta/r) \cos(\omega t - lr - \beta)$ in which β represents a possible change of phase at the dipole.

Writing α for $\omega t - lr$,

since $\cos(\alpha - \beta) = \cos\alpha \cos\beta + \sin\alpha \sin\beta$

$A \cos(\omega t - lr) \cos\beta + A \sin(\omega t - lr) \sin\beta$

and writing A_1 for $A \cos\beta$ and $+B_1$ for $A \sin\beta$ (see diagram in Fig. 287) we have for the scattered vector

$$(1) \qquad [A_1 \cos(\omega t - lr) + B_1 \sin(\omega t - lr)] R\frac{\sin\theta}{r}$$

the factor B_1 indicating the phase change.

Taking the square of the amplitude as the measure of intensity or energy, and noting that the average value of $\sin^2\theta$ over the whole sphere is $\frac{2}{3}$, we have, as the total energy

$$(2) \qquad \tfrac{2}{3}4\pi(A_1^2 + B_1^2)R^2$$

which is derived from the incident beam.

If parallel rays traverse 1 sq. cm. of thickness dx of a medium containing N molecules per c.c. and R^2 be the energy passing into this area per sec., the coefficient k, which measures the gradual weakening of the beam due to scattering, is defined by

$$(3) \qquad k = \frac{1}{R^2}\frac{dR^2}{dx} \quad \text{or} \quad -dR^2 = kdxR^2.$$

A layer of thickness dx and unit cross section contains Ndx molecules and scatters energy

$$\tfrac{8}{3}\pi(A^2+B^2)NR^2dx = -dR^2$$

(4) hence $$k = \tfrac{8}{3}\pi(A^2+B^2)N.$$

In the direction of the incident light, and no other, the scattered radiations will coöperate in the formation of a plane-wave propagated along with the incident waves, since a displacement of one of the dipoles will have no influence on the ultimate phase at a given point if the scattered ray is in the direction in which the wave is propagated. The effect of the radiations from the layer dx at a point O at distance p from dx is, by the Fresnel-Huygens principle, equal in magnitude to half that of the first zone of area $\pi p\lambda$ and the phase of the resulting vibration is that corresponding to a distance $p+\lambda/4$, the effect of the scattered light at O thus being a plane-wave, which we will call the synthetic wave.

(5) $$[A_1 \sin(\omega t - lp) - B_1 \cos(\omega t - lp)]R\lambda N dx$$
of amplitude $AR\lambda Ndx$ and phase $\beta+\pi/2$

which is to be combined with the incident wave, which at the point O is $R\cos(\omega t - lp)$.

The term in B_1 indicates a diminution in amplitude while the term in A_1 indicates a change of phase.

In Fig. 288 AB represents the amplitude R, and phase of the incident wave, BC the amplitude and phase of the synthetic wave, which we resolve into two components (dotted) one of amplitude $A_1R\lambda Ndx$ which alters the *phase* of the primary wave, the other of amplitude $-B_1R\lambda Ndx$ which decreases the amplitude of the primary wave, the resultant vibration AC being retarded in phase by the small angle $\Delta\Phi$ which can be represented by its sine, or approximately

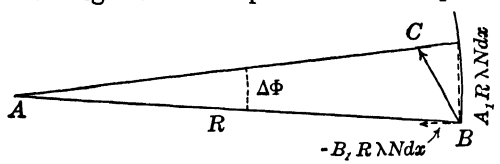


FIG. 288

$$\Delta\Phi = \frac{A_1R\lambda Ndx}{R} = A_1\lambda Ndx.$$

If the layer dx were uniform in its properties and had a refractive index n, it would cause a retardation of the incident light of $(n-1)dx$ the equivalent of a phase alteration of $2\pi(n-1)dx/\lambda$.

Equating: $$2\pi(n-1)\frac{dx}{\lambda} = A_1\lambda Ndx$$

shows that the medium behaves as one having refractive index n if

$$2\pi(n-1)=A_1\lambda^2N. \tag{6}$$

The proportional diminution in amplitude is $B_1\lambda N dx$,

$$i.e.\ R(1-\epsilon) \text{ if } \epsilon=B_1\lambda N dx.$$

And the diminution of energy or intensity is

$$R^2(1-\epsilon)^2=R(1-2\epsilon+\epsilon^2)$$

in which ϵ^2 can be neglected as a small quantity. The diminution of energy is thus twice the diminution of amplitude or $2B_1\lambda N dx$.

In all cases of more immediate interest B_1 is small compared with A_1 so that B_1^2 can be neglected in (4) which we now combine with (6)

$$k=\frac{8}{3}\pi A_1^2=\frac{8\pi N}{3}\left[\frac{2\pi(n-1)}{\lambda^2N}\right]^2=\frac{32\pi^3(n-1)^2}{3\lambda^4N}$$

nearly the equivalent of the expression first deduced by Lord Rayleigh [1] for material particles.

It is of interest to enquire what happens when the dipoles are so close together, and so uniformly distributed in space that lateral scattering is absent or nearly so, as a result of destructive interference. In this case we have as the vector representing the synthetic wave (into which all of the energy radiated by the dipoles goes), in Fig. 288 the vertical dotted arrow instead of BC. The phase retardation of the primary wave is practically the same as before, but as there is no component of the synthetic wave parallel to AB there is no loss of amplitude. From this it appears that the effect of lateral scattering is to alter slightly the phase of the synthetic wave, rotating the dotted vertical vector to the position BC. In other words the refractive index is the same regardless of the arrangement of the dipoles in space. Lord Rayleigh laid the foundations of this manner of treating refraction and dispersion in his paper on the effect of irregularly arranged obstacles, small in comparison to the wave-length, on the propagation of light. Planck, a few years later published a series of papers dealing with scattering in its relation to absorption, showing that damping of the vibration could be fully accounted for by scattering. The further developments are due to Schuster, Ewald, Esmarch, Natanson and Herzfeld. We will now consider the physical processes involved from a non-mathematical point of view. The secondary or scattered waves are all that we need to account for refraction and dispersion. These secondary waves constitute the light scattered by all transparent media when traversed by a

[1] *Phil. Mag.*, *5*, 47, 1899.

primary beam, which is usually referred to as the Rayleigh scattering, the blue color of the sky being the most conspicuous example of the phenomenon. We are familiar with this lateral diffuse scattering, which one might suppose played no direct part in modifying the velocity of the light, but there is, in addition to this, and depending upon it, a coöperative effect of a curious kind which is the factor lying at the bottom of all refraction and dispersion effects. The secondary waves, *in the direction of the primary beam,* combine to produce a wave of the same type as that of the primary wave, that is, if a system of plane-waves is traversing the medium, the secondary waves from the oscillators coöperate in the production of another set of plane-waves which run along parallel to the original waves, "synthetic waves" we shall call them, to differentiate them from the diffuse radiation which goes off in lateral directions. These waves, by their interference with the primary waves, are the cause of refraction, dispersion and attenuation, by which we mean the weakening of the light by lateral scattering, as distinguished from absorption in which the energy is transformed into heat.

We have just seen that the phase of a mechanical oscillator lags behind the phase of the exciting vibration by an amount increasing from zero for low frequencies to the value of 90° when $\nu = \nu_0$ (the condition for resonance) and then increases to 180° for high frequencies.

The phase of the synthetic waves, built up from the radiations of the oscillators, lags 90° behind that of the oscillators (in the same way that the phase of the wave, built up from the secondary waves in the manner imagined by Huygens and Fresnel, is a quarter of a period behind that calculated as a direct propagation along a ray), consequently we obtain the phase of the synthetic wave by adding 90° to the phase lag of the oscillator with respect to the primary wave.

We will now consider a medium, say a gas, made up of dipoles, traversed by plane-waves of gradually changing frequency. The dipoles will be thrown into forced vibrations, and emit scattered radiation, the amplitude of which will increase as the frequency of the light approaches that of the dipoles. This scattering of radiation reduces the amplitude of the light traversing the medium, and the synthetic waves, the amplitude and phase displacement of which with respect to the primary waves varies according to the frequency, continuously retard or advance the phase of the primary waves, which is the equivalent of a change of velocity. The loss of energy by scattering can be termed "attenuation" to distinguish it from absorption, in which the energy abstracted

from the light is changed into heat. The curve in Fig. 289 represents the phase lag of the dipoles with respect to the light, the angular values for frequencies ν_1, ν_2, ν_0 (ν_0 = natural frequency of dipoles), ν_3, ν_4 being given at the right, while the amplitudes and phases of the synthetic wave for the frequencies ν_1, ν_2, ν_0 (where largest amplitude occurs) ν_3 and ν_4 are indicated by the dotted arrows on the phase lag curve, their phase retardations being given

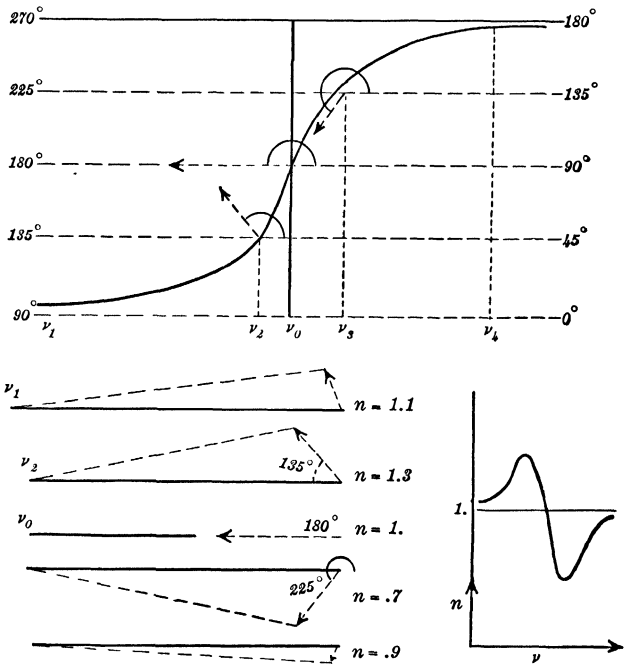


FIG. 289

at the left, obtained by adding 90° to the values at the right. The amplitude values are found as follows. (Treatment by Herzfeld.)

The amplitude of a forced vibration is proportional to $1/\nu_0^2-\nu^2$, ν_0 being the dipole frequency and ν the frequency of the light.

On the low frequency side the amplitude of the dipole approaches the small finite value $1/\nu_0^2$ with decreasing ν, while on the high frequency side it approaches 0, as shown by curve a Fig. 290.

If the amplitude of the scattered wave were proportional to the dipole amplitude we should have a condition contrary to Rayleigh's law, according to which short waves are more powerfully scattered than long. The amplitude of an electromagnetic wave is, however, proportional to the acceleration and not the amplitude,

i.e. to $\nu^2/\nu_0^2-\nu^2$ which gives curve *b*, the Rayleigh scattering being 0 for low frequencies, and increasing with increasing frequency. For the amplitude of the synthetic waves we have curve *c*, obtained by dividing $\nu^2/\nu_0^2-\nu^2$ by ν, since the combination of long waves to form a synthetic wave is more efficient than the combination of short waves.

If we are to represent the amplitude of the synthetic wave by a vector which is to be combined with the vector of the primary light to account for refractive index, we must consider the thickness of the layer of dipoles to be proportional to the wave-length of the light, *i.e.* we must consider a different thickness for each of the five frequencies. We therefore multiply the value $\nu/\nu_0^2-\nu^2$ by λ, which gives $\nu\lambda/\nu_0^2-\nu^2 = c/\nu_0^2-\nu^2$ which gives a curve similar to curve "*a*," *i.e.* our vectors are to be taken proportional to the amplitudes of the dipoles. These vectors for the five frequencies in question are shown by the arrows on the curve in Fig. 289, the one at ν_1 having a small finite value and the one at ν_4 having 0 value. Their phases are indicated at the left, and the maximum amplitude is at ν_0 of course. These we compound with a vector representing the wave which is traversing the medium, the amplitude and phase of which is represented by the horizontal lines in the lower left-hand figures.

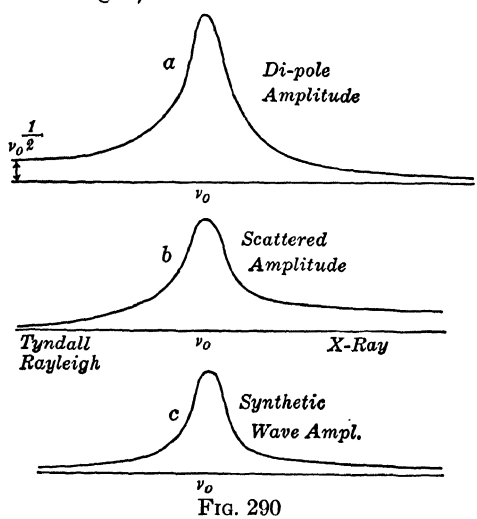


FIG. 290

For the frequency ν_1 the phase lag of the synthetic wave is slightly greater than 90° as given in the figure. This can be seen from the circumstance that if we combine a 90° vector with the vector representing the light, we find the amplitude increased instead of diminished. The lag is a little greater than 90° due to the radiation damping of the dipoles, and we accordingly compound the vector representing the synthetic wave with that of the primary wave as shown by the lower diagram at the left for ν_1. The resultant is the dotted line representing the phase and amplitude of the light as modified by the synthetic wave, for it must be remembered that a progressive retardation of the phase of

a wave-train such as occurs here, is the equivalent of a slower propagation. We will assume the refractive index corresponding to this retardation of the wave to be $n=1.1$. We can resolve the small vector into two components as in Fig. 288, the horizontal one causing the attenuation of the light. This makes it appear as though the decrease of amplitude were caused by interference of the synthetic wave with the primary wave, and this is a legitimate way of looking at the matter. We must remember, however, that when there is a decrease of intensity due to interference, we must find out where the lost energy goes. In the present case it goes into the diffusely scattered radiation.

Advancing the frequency to ν_2 we find a greater reduction of amplitude and retardation of phase, indicated by the increased length and counterclockwise rotation of the dotted arrow which causes moderate attenuation and a higher value of n, say 1.3. Coming now to the exciting frequency n_0 (the natural frequency of the dipole) we find the vectors opposed (180° phase-difference). Here we have very powerful scattering and attenuation of the light, but no effect on its phase which means $n=1$. For the next higher frequency ν_3 the phase lag is 225°, which is the equivalent of an advance of 135° on the next following wave, consequently the phase of the wave-train is advanced, corresponding to a value of n less than one, say $n=.7$. The attenuation is moderate and of the same value as at the frequency ν_2. Lastly at ν_4 the phase lag is 270° but the amplitude is 0 and $n=1$. The general trend of the curve representing n as a function of ν is given at the right of the lower figure. No attempt has been made to make the figures exact in this presentation which is to be regarded merely as an assistance in visualizing the mechanism by which attenuation and dispersion are brought about.

Refraction and Dispersion. — Consider now the case of an atom consisting of a positively charged core surrounded by a system of negatively charged electrons, or a molecule made up of a positively charged atom and say two atoms with negative charges, such as CO_2. If in the normal state the "centre of gravity" of the negative charges (or more properly the electrical centre) does not coincide with that of the positive charge, the atom or molecule is termed a natural dipole. If the electrical centres coincide, they will be displaced in an electric field and we then have an *induced* dipole. If the magnitude of the displacement is independent of the orientation of the molecule in the electric field it is said to be isotropic; if dependent on orientation, anisotropic. The moment μ of a natural dipole, in which we will assume the electric centre of q electrons each of charge e separated by a distance l from the +

charge is $\mu = qel$. The presence of natural dipoles can be shown experimentally by measuring the dielectric constant at different temperatures, since such molecules orient themselves in the field. This orientation may be upset wholly or in part by molecular impacts, the effects of which increase with rise of temperature.

In the case of induced dipoles, if E is the strength of the electric field in the medium, $\pm qe$ the displaced charge, f the force acting between the displaced charges for unit displacement and l the displacement for which the forces of separation and restitution are in equilibrium, we can write

$$qeE = fl$$

and the induced moment

$$p = qel = \frac{q^2e^2E}{f} = aE$$

in which a is the Polarizability. This induced polarization makes the field-strength E in the medium less than the field-strength D in a vacuum, since in the case of, say, a plate condenser with the medium between the plates, the + charges of the molecules are all turned towards the plate with a negative charge, and the negative plate thus has an induced + charge on its lower side while the + plate has a negative charge (Fig. 291). The ratio $D/E = \epsilon$ is called the dielectric constant.

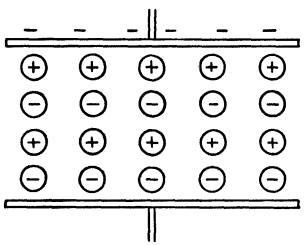


FIG. 291

In the case of a plate condenser in vacuum with surface density of charge s, the field strength between the plates is $D = 4\pi s$. If the plates are immersed in a medium, it becomes polarized, and polarization charges of surface density S_p develop on the plate, so that we have an opposing field $4\pi S_p$ directed against the original field $4\pi s$. The resultant field strength

$$E = D - 4\pi S_p$$

or the expression for the dielectric constant becomes

$$\epsilon = \frac{D}{E} = 1 + \frac{4\pi S_p}{E}.$$

The displacement in a medium due to force E is

$$D = E + 4\pi P$$

in which P is the polarization in unit volume. If N is the number

of molecules in unit volume and p the polarization of a single molecule

$$D=E+4\pi Np$$

or

$$\epsilon=1+4\pi N\frac{p}{E}=n^2. \tag{1}$$

Thus far we have considered a dielectric between the plates of a condenser subject to a steady field which produces polarization of the medium. If the field is suddenly removed the displaced + and − electric centres of the molecules return to their normal positions, executing vibrations in the process, so that the medium exhibits polarizations of decreasing magnitude. This is the circumstance which modifies the frequency of the oscillatory discharge of a condenser when a dielectric is introduced between the plates.

The equation of motion for the case in which the field has been removed is

$$M\frac{d^2x}{dt^2}+fx=0 \tag{2}$$

in which M is the combined mass of the vibrating electrons in the molecule.

If now the medium is subjected to the alternating electric force

$$E=E_0\cos\frac{2\pi t}{T}$$

in which E_0 is the "amplitude" or maximum value of the force

$$M\frac{d^2x}{dt^2}+fx=qeE \tag{3}$$

there will be a periodic polarization represented by

$$x=A\cos\frac{2\pi t}{T}$$

and differentiating and substituting in (3)

$$\underset{\text{Reaction of inertia}}{-M\frac{4\pi^2}{T}A\cos\frac{2\pi t}{T}}+\underset{\substack{\text{Reaction of}\\ \text{elastic force}}}{fA\cos\frac{2\pi t}{T}}=qeE_0\cos\frac{2\pi t}{T}. \tag{4}$$

For fast vibrations the inertia term is larger than the elastic, while for slow vibrations the reverse is true. The inertia reaction is retarded 180° on the reaction of the elastic force, as shown by the − and + signs.

Dividing (4) by $\cos 2\pi t/T$ gives

$$A\left(f-M\frac{4\pi^2}{T^2}\right)=qeE_0$$

$$A=\frac{qeE_0}{f-M\dfrac{4\pi^2}{T^2}}=\frac{qeE_0}{4\pi^2M}\left(\frac{1}{\dfrac{f}{4\pi^2M}-\dfrac{1}{T^2}}\right).$$

Now $f/4\pi^2M$ is the reciprocal of the square of the periodic time of the free vibration of the medium, *i.e.* of the oscillators, since $(2\pi/T_0)^2=f/M$;

hence
$$A=\frac{\dfrac{qeE_0}{4\pi^2M}}{\dfrac{1}{T_0^2}-\dfrac{1}{T^2}}=\frac{qeE_0}{4\pi^2M(\nu_0^2-\nu^2)}. \tag{5}$$

As we have seen from (1) $n^2-1=4\pi P_0/E_0$ and $P_0=qeNA$, in which A is the amplitude of the polarization, q the number of vibrating electrons in the molecule, and N the number of molecules in unit volume.

Substituting for P_0 and giving to A its value as defined by (5), we have now[1]

$$n^2-1=\frac{\dfrac{q^2e^2N}{\pi}}{M(\nu_0^2-\nu^2)}. \tag{6}$$

This is the dispersion formula as developed by Drude and Voigt in which polarization is represented by quasi-elastically bound electrons thrown into forced vibrations. The dielectric constant takes its correct value from the formula when we put ν very small in comparison to ν_0, *i.e.* operate with infinitely long waves

$$\epsilon-1=\frac{Nq^2e^2}{\pi M\nu_0^2}=\underset{(\lambda=\infty)}{n^2-1} \tag{7}$$

which shows that ϵ becomes larger as ν_0 diminishes, *i.e.* as the electrons are less firmly bound. Dispersion formula (6) holds, however, only for regions not too near the natural frequency: On the low frequency (long λ) side of the resonance frequency ν_0 we have $n>1$, rising in value as ν increases, and becoming infinite for $\nu=\nu_o$. Here there is a change of phase of 180° for the vibration with respect to the wave, and with further increase of ν, n comes up from $-\infty$ approaching the value 1 for very high frequencies.

[1] Or if we write, as is often done in dispersion treatments $\omega_0=2\pi\nu_0$ and $\omega=2\pi\nu=2\pi/T$, we have $n^2-1=(4\pi q^2eN)/[M(\omega_0^2-\omega^2)]$.

The form of the curve is shown in Fig. 292. This means that for very long waves the refractive index is equal to the square root of the dielectric constant, while for very short waves its value is unity. The physical explanation of this difference is that for slow frequencies there is a vibration of the dipole while for high frequencies the charges remain practically at rest, *i.e.* they are unable to respond. This accounts for the absence of refraction in the case of X-rays. It will now be shown that the difficulty with the infinite value of the refractive index which is approached as we apply the formula to values of ν nearer and nearer to ν_0 (from the low frequency side) and its imaginary value on the other side of ν_0 is taken care of by the introduction of a damping factor.

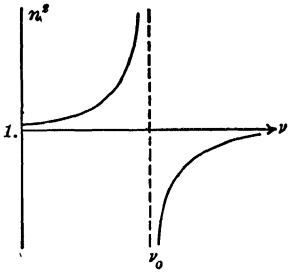


FIG. 292

The Damping Factor. — The dispersion formula as developed in the last section holds only for regions of the spectrum outside of the absorption band. The damping factor was omitted since a clearer conception of physical processes is obtained by this simplification. If a damping factor is introduced, as was first done by Helmholtz, we obtain a formula applicable to all frequencies. In this extension of the development the exponential method of representing harmonic motion, and the use of complex quantities will be employed.

We write in place of equation (3)

$$M\ddot{x}+4\pi^2M\nu_0{}^2x=qeE_0e^{2\pi i\nu t}$$

$$A=\frac{qeE_0}{4\pi^2M(\nu_0{}^2-\nu^2)} \qquad x=Ae^{2\pi i\nu t}.$$

The equation of motion, with a damping term introduced is

$$M\ddot{x}+4\pi^2M\nu_0{}^2x+2\pi\nu'\dot{x}=qeE_0e^{2\pi i\nu t} \tag{10}$$

$$A=\frac{qeE_0}{4\pi^2M(\nu_0{}^2-\nu^2+i\nu'\nu)}. \tag{11}$$

$2\pi\nu'$ is the damping factor, which is proportional to the velocity $\dot{x}$, ν' being the frequency corresponding to the "half width" of the absorption line, *i.e.* the frequency difference between the two frequencies corresponding to the points at which the intensity of the band has fallen to one-half of its value at the centre.

Now A/E_0 is the ratio of the dipole amplitude of the atom or

molecule, to the amplitude of the electric force, and qex corresponds to p in equation (1). Making this substitution in (1) we have

$$n^2=1+4\pi Nqe\frac{A}{E_0}=1+N\frac{q^2e^2}{\pi M(\nu_0^2-\nu^2+i\nu'\nu)} \qquad \text{by (11).}$$

Writing $$\rho=\frac{Nq^2e^2}{\pi M}$$

$n^2=\rho/\nu_0^2-\nu^2+i\nu'\nu$ which multiplied by its complex conjugate gives

$$n^2-1=\rho\frac{\nu_0^2-\nu^2-i\nu'\nu}{(\nu_0^2-\nu^2)^2+\nu'^2\nu^2}.$$

And separating the real from the imaginary

$$n^2-1=\rho\frac{\nu_0^2-\nu^2}{(\nu_0^2-\nu^2)^2+\nu'^2\nu^2}. \qquad (12)$$

The Complex Refractive Index. — To make the dispersion formula take care of the regions of the spectrum which are strongly absorbed it must be completed by the introduction of the absorption coefficient k. This is most conveniently done as follows:

Let $$y=Ae^{2\pi i\left(\frac{t}{\mathrm{T}}-\frac{nx}{\lambda}\right)}$$

represent the displacement in a medium of refractive index n, in the case of no absorption, A being the amplitude and e to the power indicated corresponding to $\cos 2\pi(t/T-nx/\lambda)$. The amplitude decreases however with the advance of the wave if there is absorption, and we accordingly write A as the function of x associated with absorption.

$$y=\underbrace{Ae^{-k'x}}_{\text{amplitude}}\cdot e^{2\pi i\left(\frac{t}{\mathrm{T}}-\frac{nx}{\lambda}\right)}=Ae^{-k'x+2\pi i\left(\frac{t}{\mathrm{T}}-\frac{nx}{\lambda}\right)}=$$

$$Ae^{2\pi i\left[\frac{t}{\mathrm{T}}-x\left(\frac{n}{\lambda}+\frac{k'}{2\pi i}\right)\right]}=Ae^{2\pi i\left[\frac{t}{\mathrm{T}}-\frac{nx}{\lambda}\left(1-i\frac{k'\lambda}{2\pi n}\right)\right]}$$

and writing

$$\kappa=\frac{k'\lambda}{2\pi n}$$

$$y=Ae^{2\pi i\left[\frac{t}{\mathrm{T}}-\frac{nx}{\lambda}(1-i\kappa)\right]}=Ae^{2\pi i\left(\frac{t}{\mathrm{T}}-\frac{\tilde{n}x}{\lambda}\right)}$$

in which we have written $\tilde{n}$ for $n(1-i\kappa)$; $\tilde{n}$ is called the complex refractive index, and we see that the final equation is the same as

(13) except that we have $\tilde{n}$ in place of n. Substitution of this in (12) gives $n^2-2n^2i\kappa-n^2\kappa^2-1$, in place of n^2-1. (Amplitude decreased in ratio of $1:e^{-2\pi n\kappa}$ in traversing distance λ.) The relation of κ to k is $k=2n\kappa$. Equating the real parts gives

$$n^2(1-\kappa^2)-1=\Sigma\rho\frac{\nu_0^2-\nu^2}{(\nu_0^2-\nu^2)^2+\nu'\nu}. \tag{13}$$

Equating the imaginary

$$2n^2\kappa=\Sigma\rho\frac{\nu'\nu}{(\nu_0^2-\nu^2)^2+\nu'^2\nu^2}. \tag{14}$$

The summation sign in the two equations makes them applicable to media having more than one absorption band, the summation being for as many terms as there are bands.

Experimental Proof of Dispersion Formula. — In the earlier treatments of dispersion experimental verification of the formulae was sought chiefly in the case of solids and liquids. At the time there were no reliable data on the dispersion of gases over a wide spectral range, such as we have at the present time.

As it is usually best to begin the study of any physical process with matter in its simplest condition, we will apply the formula to the dispersion of a monatomic gas, argon.

Of especial interest is the question of whether q as determined, from observations of the refractive index, gives us the number of electrons in the molecule which contribute to the dispersion, and whether or not this number agrees with the number of loosely bound or valence electrons. We can also calculate the value of ν_0, the frequency of the absorption band. In the case of gases n differs but little from unity and can be written $n^2-1=2(n-1)$. The ratio of the charge to the mass of an electron e/m has been very accurately determined by Millikan and the number N is known, so that these values can be inserted in our dispersion formula. Remembering that by M we represented the combined mass of the vibrating group of electrons in the atom, we write

$$M=qm$$

and have

$$n-1=\frac{qe^2N}{2\pi m(\nu_0^2-\nu^2)} \tag{8}$$

or writing

$$\rho=\frac{qe^2N}{m\pi}$$

$$n-1=\frac{\frac{\rho}{2}}{\nu_0^2-\nu^2}. \tag{9}$$

We require the values of the constants ρ and ν_0 the natural frequency of the dipoles. As we shall see presently we can calculate the wave-length of the absorption band (which gives us ν_0) from measurements of the dispersion. If we can determine it experimentally so much the better, but this is not always possible. The value of the constant ρ we determine in the same way, and since the value e/m is now very accurately known, we can calculate the value q, the number of electrons in each atom which coöperate in causing dispersion. These are called "dispersion electrons" and under certain circumstances q may be equal to the number of the valence electrons of the chemist. The dispersion of argon has been determined by Quarder [1] from 2441 to 5782 A.U., and the values are well represented by the formula

$$n-1=\frac{4.9981\cdot 10^{27}}{17953\cdot 10^{27}-\nu}$$

which gives for the wave-length corresponding to ν_0, the ultra-violet absorption band, 708 A.U. and $p=4.58$. Cuthbertson investigated the ultra-violet absorption of argon with a vacuum concave-grating spectrograph and found a broad region of absorption beginning at 900 A.U. with the gas at 5.5 mms. pressure, and 687 with .02 mm., with no return of transparency on the short-wave-length side. This is undoubtedly a continuous absorption band due to ionization of the atom, analogous to the continuous band below the head of the principal series of hydrogen, and appears to be identified with the band which influences dispersion. He experimented with all of the rare gases and came to the conclusion that in all the region of maximum absorption occurred at a higher frequency than that calculated from the dispersion formula.

The close agreement between the observed and calculated values of n is shown in the following table:

ARGON

λ IN A.U.	OBS. $(n-1)10^6$	CAL.	DIFF.
2441	303.78	303.97	+.19
2618	300.38	300.37	−.01
2961	295.50	295.28	−.22
3349	291.62	291.43	−.19
4275	286.34	286.25	−.09
5105	283.79		
5218	283.50	283.03	+.13
5700	282.55	282.77	+.22
5782	282.47	282.65	+.18

[1] *Ann. der Phys.*, *74*, 255, 1924.

If more than one type of resonator is present, *i.e.* if we have more than one absorption band, the refraction is represented as a summation of terms similar to the one already developed.

$$n^2-1=\sum\frac{qe^2N}{m\pi(\nu_0{}^2-\nu^2)},$$

in which the q and ν_0 values will differ in each term of the summation, as many terms being taken as there are absorption bands or lines.

If we include infra-red absorption bands, the q values determined for them from the e/m ratio, come out hopelessly small, from which Drude concluded that these bands are due to the vibration of entire atoms or ions. In this case the formula is written

$$n^2-1=\sum\frac{qe^2N}{m\pi(\nu_v{}^2-\nu^2)}+\sum\frac{qe^2N}{M\pi(\nu_r{}^2-\nu^2)}$$

ν_v and ν_r being ultra-violet and infra-red frequencies and M the mass of the ion.

Dispersion of Hydrogen. — As an example of the application of the formula to a diatomic gas, we may take the case of hydrogen, the dispersion of which has been determined by Kirn [1] from $\lambda=1854$ to 5462. Schuler and Wolf found his results well represented by the following formula

$$n-1=\frac{0.75379\cdot 10^{27}}{16681.3\cdot 10^{27}-\nu^2}+\frac{0.919974\cdot 10^{27}}{10130.5\cdot 10^{27}-\nu^2}$$

with $\lambda=734.5$ A.U. and 942.6 A.U. $q_1=.69$, $q_2=.84$

HYDROGEN

$(n-1)10^7$

λ	Obs.	Cal.	Diff.
1854	1759.96	1760.29	+.33
1862	1755.41	1755.74	+.33
1935	1718.24	1718.18	−.06
1990	1693.95	1693.73	−.22
2302	1594.18	1593.93	−.25
2379	1576.81	1576.68	−.13
2535	1546.90	1547.01	+.11
2753	1515.00	1515.15	+.15
2894	1498.59	1498.83	+.24
2968	1491.01	1491.33	+.32
3342	1461.33	1461.39	+.06
4047	1427.41	1427.45	+.04
4078	1426.32	1426.37	+.05
4359	1417.33	1417.72	−.01
5462	1396.50	1396.20	+.30

[1] *Ann. der Phys.*, *64*, 566, 1921.

We shall see presently that in the case of the vapors of the alkali metals, the absorption responsible for dispersion is practically restricted to that of the lines of the principal series, and that the greater part of the effect is due to the resonance line (*D* line doublet). The continuous absorption which begins at the head of the series in the ultra-violet, has little or no effect on the dispersion. In the case of argon the reverse is true. If a two-member formula is used in calculating the n values, assigning to one ν_0 the value of the resonance line of argon ($\lambda = 1029$ A.U.) and to the other the wave-length of the maximum of continuous absorption as calculated from the dispersion measurements ($\lambda = 703$ A.U.) the q value for the former is .025 and for the latter 4.59. This shows that the absorption at 1029 is extremely weak, for q can be regarded also as a measure of the intensity of absorption.

Dispersion of Metal Vapors. — Results obtained with the vapor of the metal sodium furnishes one of the most striking examples of the application of the formula, for q comes out equal to unity, corresponding to the single valence electron of the atom.

A study of the dispersion of the vapor of metallic sodium was made by the author, in 1904, with a view of testing the dispersion formula.[1]

As the phenomena exhibited by this vapor are among the most beautiful in physical optics, they will be considered in some detail.

The apparatus for showing the selective dispersion of the vapor is very easily prepared. Though glass tubes have sometimes been used in these experiments, their use is discouraged, as they are very liable to crack.

Procure a piece of thin steel tube 30 mms. in diameter and 40 cms. long, drill a small hole near one end and braze in a short piece of brass tubing. The wall of the steel tube should be as thin as possible, to prevent heat conduction, a millimetre or less if possible.

Close one end of the tube with a piece of thin plate glass cemented with sealing wax. To make an air-tight joint, the steel tube should be heated until the wax will melt when brought in contact with it. Spread a layer of the wax around the end of the tube, warm the glass plate, and press it firmly against the wax ring, first softening the latter with the Bunsen flame. Now go over the joint with a minute pointed gas flame burning at the tip of a glass tube drawn down to a point, until the wax runs into close contact with the glass. A joint properly made in this way will hold a cathode ray vacuum for hours.

A tube of pyrex, or other glass of low expansion can be used

[1] Wood, "A Quantitative Determination of the Anomalous Dispersion of Sodium Vapor," *Phil. Mag.*, Sept., 1904.

in which case a metal trough for holding the sodium should be made by sawing a longitudinal strip from a brass tube of such size as to slip easily into the glass tube. This prevents the molten metal from making contact with the glass. If it is not employed the tube usually cracks if allowed to cool and then reheated.

Mount the tube in a horizontal position in a clamp stand at the height of the slit of the spectroscope, which is to be used for viewing the phenomenon, and introduce eight or ten small pieces of clean dry sodium, pushing them down the tube one by one, so that they may lie side by side along the middle portion of the tube. Now close the other end with a piece of plate glass, in the same way. The tube should next be placed in communication with an air pump through a stopcock, which can be closed after the exhaustion. A U mercury manometer between the tube and the stopcock is useful as a means of getting the residual pressure of the hydrogen correct, and as an indicator of leaks. If a metal tube is used strips of wet cotton should be wrapped around it near the ends, to prevent the sealing wax from softening, and water poured on from time to time. A strip of wet cotton laid along the top of the tube improves the conditions when using a metal tube.

The under side of the tube is heated by a row of small flames, furnished by a burner, as shown in Fig. 293. The burner is made

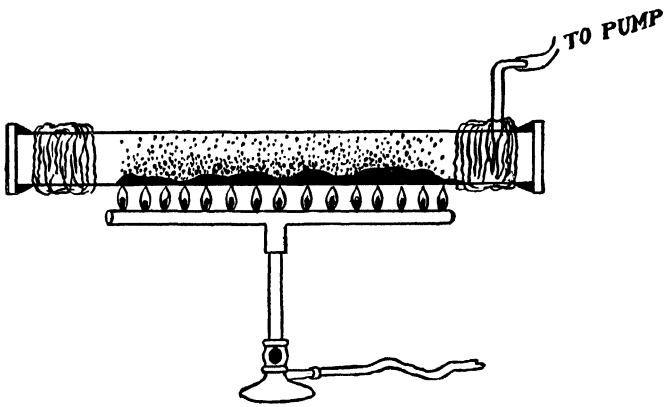


FIG. 293

by drilling a dozen or more minute holes in a piece of brass tubing, closing the ends, and mounting it on a Bunsen burner. On heating the tube the sodium vaporizes, and diffuses gradually to the cooler parts of the tube. The metal contains a large amount of hydrogen, most of which should be removed by the pump. The flames should be very low at the start, and raised gradually. A high vacuum

should not be used, for it is the presence of hydrogen at low pressure that makes the sodium vapor distribute itself in a non-homogeneous manner. A pressure of one or two centimetres can be used. The vapor is given off along the heated floor of the tube, but its diffusion towards the cooler top is interfered with by the residual gas in the tube. The vapor will be found to be very dense along the floor and highly attenuated along the top. This condition can obtain only when some other gas is present to keep the pressure balanced.

The sodium vapor is therefore the optical equivalent of a prism, for the amount of it traversed by the light increases as we pass from the top to the bottom of the tube; the equivalent of a prism with its refracting edge horizontal, resting upon its base. The deviation of the rays will be gradual, however, as in the non-homogeneous media described in the Chapter on Refraction. The density gradient is steeper near the bottom of the tube, and to secure good definition it is a good plan to place in front of the tube an opaque screen perforated with a horizontal slit, say 1 cm. in width, selecting the position of best definition.

The arrangement of the apparatus is shown in Fig. 294. A horizontal slit is illuminated by focussing the sun or arc upon it, and

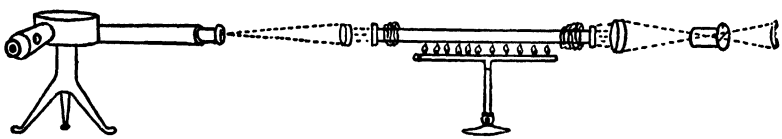


FIG. 294

the emergent rays, rendered parallel by a lens, passed down the sodium tube. A second lens forms an image of the horizontal slit across the vertical slit of the spectroscope. This image should be carefully focussed, so that the spectrum appears as a brilliant narrow band with sharp edges.

On heating the tube, the sodium prism deviates the rays of different wave-length up or down by different amounts, curving the spectrum into two oppositely directed branches. The spectrum on the green side of the *D* lines will be found to bend down in the spectroscope, which means that the rays are deviated upwards in passing through the sodium tube, since the spectroscope inverts the image of its slit. This means that this phase velocity is greater in the sodium vapor than in vacuo, or the prism acts for these rays like an air prism immersed in water. The red and orange region is deviated in the opposite direction; these rays are therefore retarded by the vapor.

A photograph of the anomalously dispersed spectrum is shown in Fig. 295, and colored drawing on the Frontispiece, Fig. 3.

If a metal tube is used it may be found that the performance of the tube is improved by laying a long pad of wet cotton along the top. This makes the temperature gradient steeper and facilitates the formation of the non-homogeneous vapor cylinder. A tube of the dimensions given should show the phenomenon on as great a scale as the photograph reproduced in Fig. 295. In this case the position of the *D* lines was recorded on the plate by holding a sodium flame in front of the slit, after the exposure was over. If the spectrograph is capable of resolving the *D* lines, the initial stage is as shown by Fig. 296 and Fig. 2 of the Frontispiece.

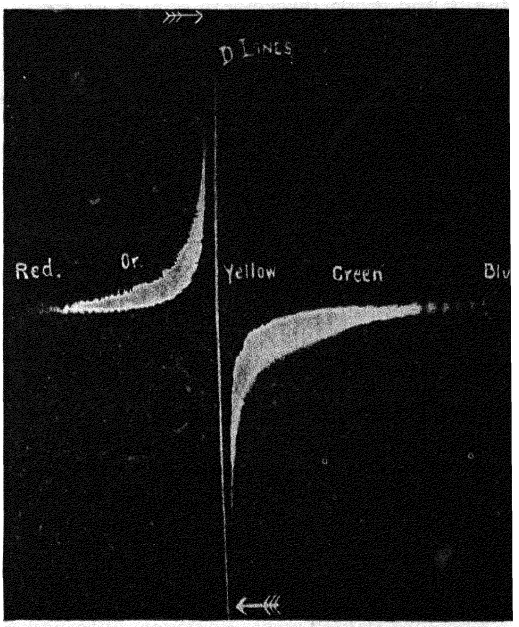


FIG. 295

If we remove the spectroscope and place an eyepiece in the position previously occupied by the slit, we shall see the anomalous spectrum produced by the sodium prism, as shown on the colored Frontispiece, Fig. 1.

The development of this spectrum from the horizontal image of the slit, seen when the tube is cold, is most instructive to watch, and shows us at once that the sodium prism deviates some of the rays up and others down.

If the electric arc is employed as the source of light, the extreme violet will be found to occupy the position of the undeviated image of the slit. Then comes the blue, sometimes in contact with the violet and sometimes slightly separated by a fine dark line, owing to the fact that the violet light comes from the fluted carbon band of the arc, which is separated from the blue by a comparatively dark region. Then comes a wide gap corresponding to light absorbed by the sodium vapor in the blue-green region (the chan-

nelled spectrum), and above this a beautiful flare of color ranging from blue-green through grass-green to yellow. The red and orange portion of the spectrum is on the other side of or below the undeviated image, forming another brilliant flare of color. It is separated from the violet by a wide dark band, due to the absorption in the vicinity of the D lines. If the density of the vapor is increased by heating the tube to a higher temperature, the red flare extends lower down, grows fainter and finally fades away, owing to the presence of the fluted absorption bands in red. The green and blue persist, however, becoming more widely separated, but finally the green disappears almost entirely. It is best to arrange the gas cock so that the height of the flames can be controlled without leaving the eye-piece, for it is surprising how slight a change is necessary to completely alter the general appearance of the spectrum.

To obtain this spectrum in its greatest splendor, it is best to use a tube of pyrex glass about 100 cms. long heated in the same way, containing fifteen or twenty small pieces of sodium. The temperature gradient is steeper in a glass tube and the dispersion much more powerful. The flames may be turned up and down during the experiment, but they should never be turned quite out, for the tube will almost invariably crack on reheating after it has once cooled down.

By employing a quartz spectrograph photographs were made of the anomalous dispersion at the ultra-violet lines 3303 and 2852. It is much less strong at 3303 and barely noticeable at 2852. These absorption lines all belong to the principal series of sodium, and they are spaced along the spectrum according to a definite law.

Though we can measure the relative indices of the vapor by this method, we have no means of determining the absolute values, for the prism angle is unknown.

Absolute values were, however, obtained by means of the interferometer, by comparing and measuring the shifts of the fringes obtained by introducing a given amount of sodium vapor into the path of one of the interfering beams. A full description of the methods and apparatus employed will be found in the paper referred to. In brief, it consisted in placing a sodium tube, electrically heated, between two of the mirrors of a Michelson interferometer. Uniform heating was necessary in this case, as no prismatic action was desired. The instrument was illuminated with two sources of light, one a helium spectrum tube which gives us a bright yellow light, D_3, very near the sodium absorption band, the other a spectroscope arranged to furnish a beam of approximately monochromatic light in any desired part of the spectrum.

Two sets of fringes were thus formed, and the drifts of the systems were recorded by two observers as sodium vapor was formed in the tube. To obtain the dispersion very near the absorption band, the helium tube was placed in a powerful magnetic field, which caused the line to become double (with suitable arrangement of the apparatus). The two components were of very nearly the same wave-length, the distance between them being about $\frac{1}{20}$ of the distance between the D lines, yet the dispersion of the vapor was so powerful that the two sets of fringes were displaced at rates so different that the fringes disappeared entirely at regular intervals, owing to their "out-of-step" superposition.

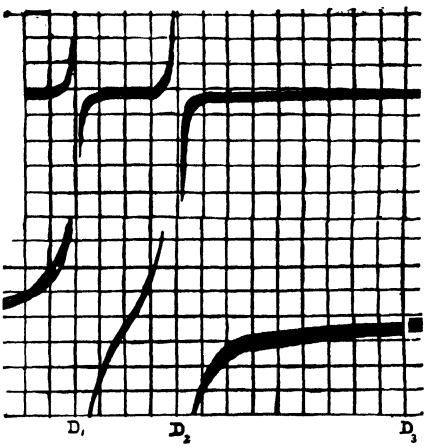


FIG. 296

By employing the method of "crossed prisms," relative determinations were made still closer to the D lines than the helium line. To see the effect close to and between the D lines, the tube should be only slightly heated, and a grating spectroscope employed. As the vapor prism forms, we see presently the portions of the spectrum adjacent to the absorption lines curve away in opposite directions, as shown in Fig. 296. As the vapor becomes denser, the light disappears between the D lines, and we have the stage previously described.

Absolute values of the refractive index were obtained by heating the tube to a known temperature, measuring the length of the vapor column, and counting the fringe shift produced when monochromatic light of known wave-length was used to illuminate the instrument.

The Refraction and Dispersion of Dense Sodium Vapor. — An attempt was made to make a table showing the refraction and dispersion of the vapor at the pressure corresponding to 644° C. At this pressure a column 8 cms. in length has a distinct blue color, and the absorption band at the D lines is broad enough to cut out a region over 30 A.U. in width, or five times the distance between the lines. The fringe shifts for about thirty different wave-lengths, caused by heating the tube to 644°, were counted. Observations could not, of course, be made within the absorption band, but the values of n within this region could be calculated from observations

made with less dense vapor. The helium fringes became very faint when the vapor had acquired a density sufficient to give a shift of about 100 fringes, owing to the finite width of the helium line. They could be made distinct again by interposing a thick plate of optical glass in the air path of the interferometer, and the count continued. Other interesting points in regard to the behavior of the vapor will be found in the original paper. Shifts of 400 fringe-widths were obtained in this way. Beyond this point it was impossible to go with helium light, the monochromatic constituents of the yellow line becoming too much spread out by the dispersion of the vapor to be brought together again by going out any farther in the system. The green mercury line was then used, with which shifts as high as 50 fringes were observed, corresponding to a shift of 2500 helium fringes. In this case the vapor column glowed with an orange-colored light.

A few of the results obtained are recorded in the following table, wave-lengths in the first column, relative shifts with respect to the helium line in the second, and total shifts which would be observed for a shift of 50 green mercury line fringes ($\lambda = 5460$) in the third. The values in the second column are shifts corresponding to a shift of 100 helium fringes. The shift for 5460 is seen to be 4, consequently at the high density giving the shift of 50 fringes for this line the helium fringes would be shifted by $100 \times {}^{50}/_{4} = 1250$ as given in the table. The refractive indices are given in the fourth column.

λ	Shift Corresponding to 50 Helium Fringe Shifts	Total Shift Corresponding to 50 Green Hg Line Shifts	Ref. Index n	
7500	2	25	1.00012	
6200	6	75	1.00029	.000285 (cal.)
6013	14	175	1.00066	
5960	25	313	1.00116	
5916	60	750	1.00297	
5875 He	100	1250	0.9954	.9958 (cal.)
4848	50	625	.9977	
5827	25	313	.9988	
5460 Hg	4	50	.999829	
4500	1.4	17.5	.999951	.999965 (cal.)

For wave-lengths very near the D lines the following values were obtained:

λ	n (Obs.)	n (Cal.)	λ	n
5882	.991	.989	5904	1.01
5886	.974	.975	5899	1.02
5888.4	.944	.945	5897	1.09
5889.6	.614	.697	5896.4	1.38

These observations were applied to the Sellmeier formula considering the D lines as a single absorption band, the constant m being calculated from the n values for two widely separated wave-lengths 5460 and 5850, the values found being .000056 and .000054, a surprisingly close agreement.

An inspection of the formula shows us that, according as we are on the red or blue side of the absorption band, the refractive index (squared) is given by adding to (or subtracting from) unity the value of the constant $m=0.000055$ multiplied by the fraction $\lambda^2/\lambda^2-\lambda_m{}^2$.

In the case of sodium vapor the value of the fraction may be several hundred or even thousand. In the case of $\lambda=5882$ the fraction is 367, and yet the observed and calculated values agree closely. For $\lambda=58884$ the fraction is 1940; and for $\lambda=5889.6$ we have a value as high as 3944. The product of these very large numbers and the small fraction 0.000055 give, however, values of the index which are in close agreement with the observed values.

Discrepancies occur in the immediate vicinity of the D lines which can be explained in the following way: To get values in any way consistent with the observed values it was necessary to assign to λ_m the value of the D_2 line, the mean value 5893 being too far removed from the wave-lengths in question to give the requisite steepness to the curve. The calculated values, therefore, apply to a medium with a single band at D_2 and with a constant $m=0.000055$. This gives us a pretty good approximation to the observed curve, but the latter is due to the combined effects of the bands D_1 and D_2, the presence of the D_1 band tending to make the observed curve flatter than the calculated. A more correct approximation could be obtained by assigning to λ_m a value intermediate between D_2 and 5893. The proper method of procedure would, of course, be to make use of two members in the dispersion formula, one for D_1 and the other for D_2, thus:

$$n^2=1+\frac{m\lambda^2}{\lambda^2-\lambda_m{}^2}+\frac{m'\lambda^2}{\lambda^2-\lambda_{m'}{}^2}.$$

A few of the calculated values are included in the table. This was the first attempt at a quantitative proof of the dispersion formula for a medium of this nature. Bevan [1] subsequently made similar observations with the other alkali metals, measuring the dispersion at the first four lines of the series and using a formula of 8 terms in the case of rubidium, on account of the very wide separation of the doublets.

More recently Goldhammer has compared the values found

[1] *Proc. Roy. Soc.*, *83*, 421, 1910; *85*, 58, 1911.

by the author with the calculated values using both a one-term formula in which the two D lines were considered as a single broad line

$$n-1=\frac{139.475\cdot 10^{24}}{259.440\cdot 10^{27}-\nu^2}\ (\nu=\text{wave-number})$$

and a two-term formula considering the two D lines separately. The formula

$$n-1=\frac{4.5391\cdot 10^{24}}{258.894\cdot 10^{27}-\nu^2}+\frac{9.0782\cdot 10^{24}}{259.418-\nu^2}.$$

His results show a much closer agreement between observed and calculated values than the earlier calculations, the superiority of the two-term formula as we near the D lines being apparent. A few of his values are given in the following table:

λ	$(n-1)10^5$ Obs. Wood	$(n-1)10^5$ Cal. from (1)	$(n-1)10^5$ Cal. from (2)
2260	−1	−.5	−.5
3270	2	1	1
3610	3	2	2
4500	5	4	4
5400	15	14	14
5700	40	40	39
5750	50	55	53
5807	91	94	89
5843	150	170	157
5850	180	200	183
5867	310	340	310
5877	460	530	462
5882	920	1010	830

The large discrepancies at 3270 and 3610 are doubtless due to the influence of the ultra-violet absorption line at 3300, while the wave-length 2260 is below the head of the series. The values for q_1 and q_2 (page 490) calculated from these values are .3 and .6, their sum being nearly 1 corresponding to one electron.

The "Hacken-Methode." — The method of studying the anomalous dispersion of gases and vapors was greatly improved in 1912 by Roschdestwenski.[1] He employed at first a method due to Puccianti, focussing an image of the horizontal interference fringes of a white light source on the slit of a spectroscope. The spectrum then appears traversed by horizontal dark bands, and if a dispersing medium is introduced into one of the optical paths, the bands will be bent up or down according to circumstances. At an absorption band, where the refractive index changes, for example,

[1] *Ann. der Phys.*, *39*, 307.

from values greater to values less than unity, the dark bands will be bent up on one side, and down on the other, curving in opposite directions to the right and left of the band as shown in Fig. 298*a*. Here we have a photograph of the dispersion curve at the *D* lines of sodium, which can be translated into numerical values if we can determine the shift of the dark bands at various points of the spectrum. (Photographs by Filippov and Prokofjew.)

To increase the accuracy of such measurements Roschdestwenski devised a most ingenious method, which he named the "Hacken-Methode." If a glass plate is placed in one of the optical paths, the horizontal bands in the spectrum are thrown into an inclined

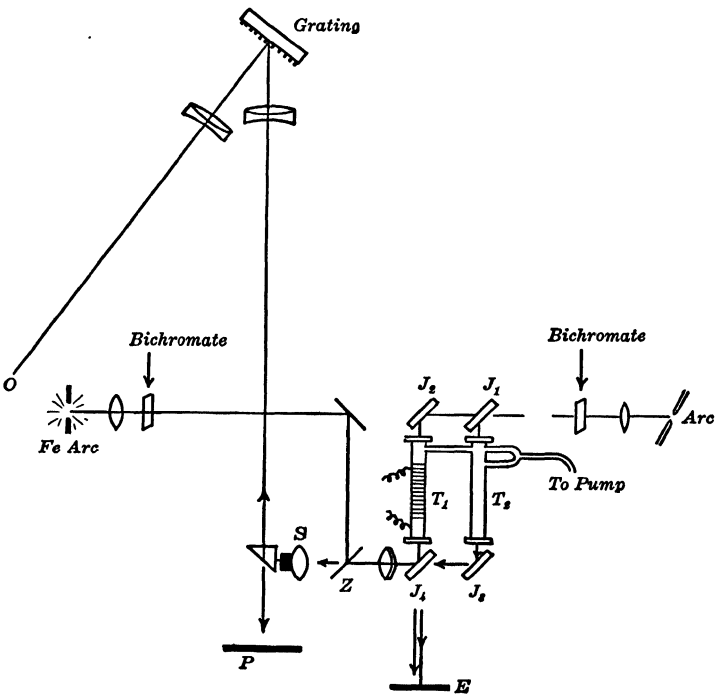


FIG. 297

position, due to the dispersion of the glass, the inclination increasing with the thickness of the plate. If now sodium vapor is introduced into the other path it will incline the fringes in the opposite direction, and there will be two points, to the right and left of the absorption band where compensation of inclinations occurs, the bands having a "hook" or v-shaped bend, which points *up* on one side and *down* on the other side of the absorption band.

Inasmuch as the inclination of the bands due to the vapor is much reduced in the vicinity of the hook, they can be seen in regions much nearer the centre of absorption than was the case before the introduction of the glass plate, Fig. 298*c*, showing the appearance of the fringes. To determine the dispersion it is necessary only to determine the value of λ at the points at which the hooks appear, and the initial slope of the fringes due to the glass plate since $dn/d\lambda$ is the tangent of the dispersion curve, and the slope of the sodium dispersion curve at the λ values where the hooks appear is identical with the slope due to the glass plate.

This method is an extremely valuable one and has been employed extensively in the investigation of the dispersion of gases, both in the normal and in the electrically excited state, in particular by

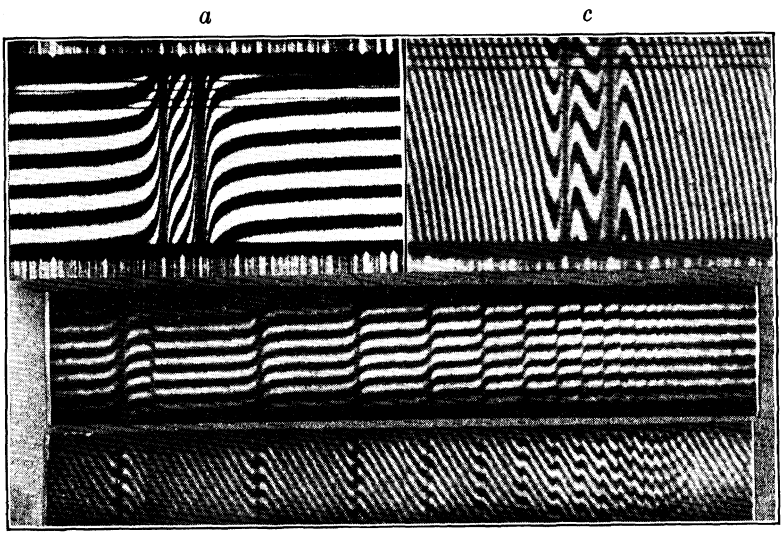


FIG. 298

Ladenburg and his collaborators. Roschdestwenski's disposition of apparatus is shown in Fig. 297. He used an interferometer consisting of 4 plates of which J_1 and J_4 were half silvered; J_2 and J_3 heavily silvered. This is Mach's modification of the Jamin instrument, giving a wide separation of the beams. The light from an arc focussed through a cell of bichromate of potash solution is divided at the plate J_1, the two beams traversing the paths indicated. The lens L_2 projects an image of the horizontal fringes on the vertical slit s of the grating spectrograph. The other pair of rays project a fringe system on the screen E. The iron arc was

projected above and below the fringe system, as a comparison spectrum, by the mirror Z, and two small mirrors z, one above the other. T_1 is a quartz tube containing sodium and T_2 a similar tube of glass, both closed with glass plates and exhausted. The fringes could be watched through an eye-piece at o focussed on a spectrum of a different order. Photographs were made on a plate at P by rays passing above the prism.

The most extensive and accurate investigation of the dispersion of sodium vapor is that of Filippov and Prokofjew,[1] who employed an interferometer of fluorite and measured, by the "hacken-methode" the dispersion in the vicinity of the first 16 doublets, and observed anomalous dispersion up to the 25th member. Portions of two of their photographs are reproduced in Fig. 298*b*, the lower made by the "hacken-methode," the upper showing the deviations of the interferometer fringes at the lines from the 6th doublet to the end of the series. At the right of the 6th line another absorption line is seen giving reversed deviations. This is the 2536 line of mercury, the reversal of the dispersion being due to the fact that the mercury vapor was in the other path of the interferometer (*i.e.* the vacuum path) having entered from the manometer or pump. They found that the dispersion over the complete spectrum is governed chiefly by the first doublet, the effect of the second member being 70 times smaller. The band of continuous absorption beyond the head of the series was found to be without influence on the refractivity.

Absorption and Dispersion of Electrically Excited Hydrogen. — Hydrogen in the normal or molecular state is one of the most transparent gases known. In the atomic or dissociated state it is capable of absorbing the lines of the Lyman series, while in an electrically excited state the Balmer series lines are also absorbed. Reversal of the hydrogen lines in stellar spectra were observed long ago, and a simple experiment showing the absorption of the light of the red line by the electrically excited vapor was described by the author in 1902. An end-on hydrogen vacuum tube appears pink when viewed laterally and bluish-white when observed end-on, due to the fact that the long column of glowing gas absorbs powerfully the light of the red line and little or no absorption for the other rays. Ladenburg and Loria were the first to study the phenomenon in a precise way.[2] They used an arrangement similar in principle to one previously employed by Pflüger, reversed both the red and green lines, and obtained photographs showing the selective dispersion and magnetic rotation in the vicinity of the red line

[1] *Zeit. für Phys.*, *56*, 458, 1929.
[2] *Ver. d. D. Phys. Ges.*, *10*, 858, 1908.

The apparatus, as arranged for showing the dispersion of the gas, is shown in Fig. 299. Light from a capillary hydrogen tube E, excited by the discharge of a coil with Leyden jars in circuit, is rendered parallel by a lens L, and divided into two beams by the mirror of a Jamin (or other) interferometer P_1. One beam passes down the second (absorption) hydrogen tube A, the other along its side, traversing, however, the projecting parts of the glass plates used for closing the tube. The beams are united by the second

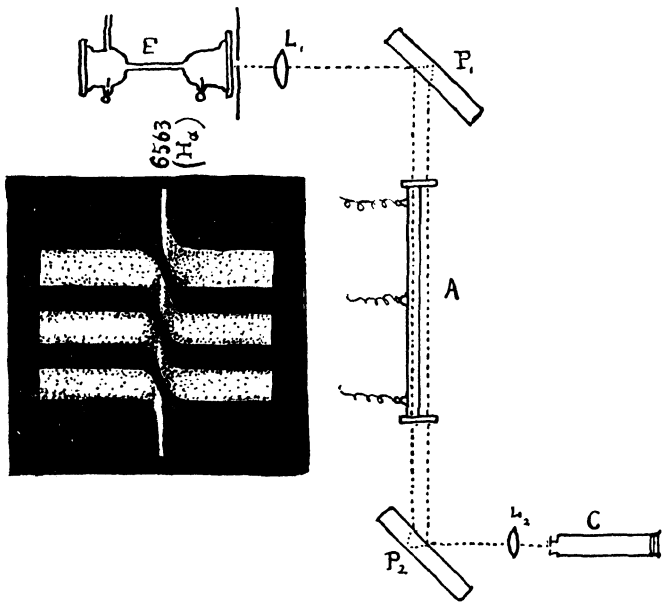


FIG. 299

Jamin plate P_2 and the horizontal interference fringes focussed on the slit of a spectroscope C by the lens L_2. On account of its small bore, the tube E emits, in addition to the hydrogen lines, a good deal of continuous spectrum; consequently we see in the spectroscope a spectrum traversed by horizontal interference bands. If now the absorption tube A, which has a wide bore, and contains hydrogen at 3 or 4 mms. pressure, is excited by being included in the same electrical circuit with E, the interference bands are bent away in opposite directions, to the right and left of the red hydrogen line Ha, as shown in the figure. A bright line runs down the centre, since only one of the interfering beams passes down the absorption tube. This experiment shows us that the passage of an electrical discharge through hydrogen gives rise to the forma-

tion of dispersion electrons, which are not present in the gas normally.

Hydrogen, then, when ionized, or brought into a state of luminescence by the electrical discharge, has quite different optical properties from ordinary hydrogen. Sodium vapor possesses these properties normally, in the absence of any electrical stimulus.

From the magnitude of the effect in hydrogen, Ladenburg and Loria calculated that the number of dispersion electrons per cubic centimetre was roughly $4 \cdot 10^{12}$, while the number of molecules was

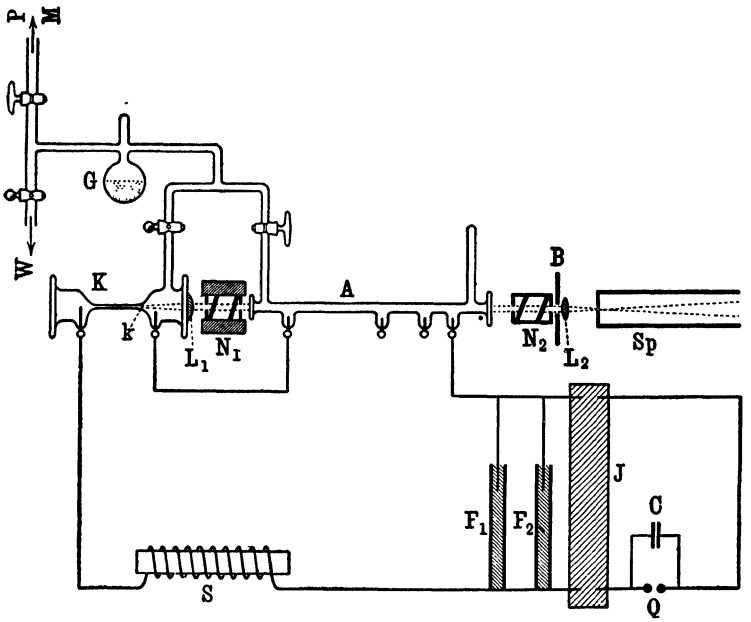


FIG. 300

$2 \cdot 10^{17}$; in other words only one dispersion electron was formed for every 50,000 molecules. In the case of sodium vapor, Hallo and Geiger found a ratio of 200 : 1.

For studying the absorption they employed the apparatus shown in Fig. 300. The intensity of the light from the source K can be regulated at will by turning the Nicol prism N_1. Starting with the two Nicols crossed, by which the light from the source is cut out, one sees in the spectroscope only the bright emission lines from the discharge in the absorption tube A, which contains hydrogen at one mm. pressure. On turning the Nicol, one observed the spectrum of the source K, which is nearly continuous on account

of the high pressure of the gas and the small diameter of the tube. Two dark lines appear at the edges of the bright lines $H\alpha$ and $H\beta$ where they cut across the continuous spectrum, as shown in

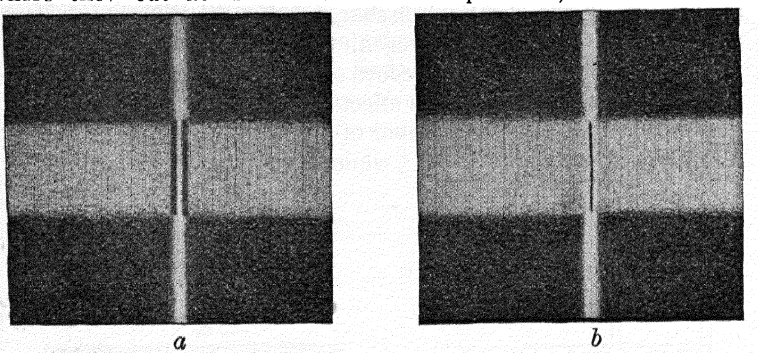


FIG. 301

Fig. 301*a*. By further turning of the Nicol, these lines widen and finally obliterate the bright line lying between them. In the second case, with the absorbing layer at a pressure of 25 mms. and 55 mms. in length, the dark line appeared in the middle of the bright line and gradually widened until the bright line was completely extinguished (Fig. 301*b*).

Dispersion of Nitroso-Dimethyl-Aniline. — The remarkable optical properties of this substance have been investigated by the author.[1] It is of especial interest in that, while fairly transparent for wave-lengths comprised between the red and blue, it has its band of metallic absorption in the violet. This circumstance gives it an enormous dispersive power in the yellow and green, a prism of the substance yielding a spectrum about fifteen times as long as the spectrum given by a glass prism of the same angle. It thus shows visually a spectrum of about the dispersion that would be seen with a prism of carbon bisulphide if our eyes were sensitive to the ultra-violet and could follow the spectrum down to the absorption band that causes the dispersion.

The substance melts at 85° C. and can be formed into prisms between small strips of thin plate glass. The strips should be about two centimetres long, and are best fastened together with one of the small clamps used with rubber tubing. It is best to melt the material on the end of one of the strips, the other being warmed over the same flame, and then clamp the two together with a piece of a match between the ends, to give the required prismatic form. A candle flame viewed through the prism is spread out into a most

[1] *Phil. Mag.*, 1903.

remarkable spectrum. It is instructive to have a prism of the same angle made of Canada balsam or some such substance pressed out between two similar glass strips.

Determinations of the refractive index were made with a spectrometer in the visible spectrum, and in the ultra-violet by photography with a quartz spectrograph by the method of crossed prisms. The fluid being pressed between quartz plates kept warm by an electrically heated wire, and mounted in the prism box of the spectrograph.

The substance was found to become transparent again on the ultra-violet side of the absorption band, and measurements were made in this region by means of photography. A small quartz spectrograph was used, the nitroso prism being mounted with its refracting edge horizontal, immediately behind the quartz prism of the instrument. This device will be at once recognized as the method of crossed prisms.

The undeviated spectrum was photographed by the rays which passed below the edge of the small prism, and by measuring the

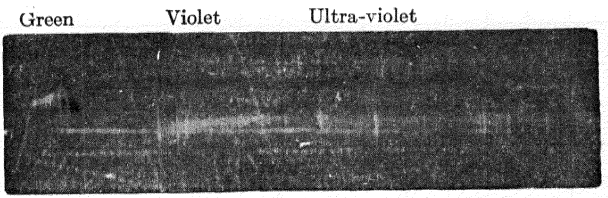


FIG. 302

distances between it and the deviated portion, it was possible to calculate the refractive index. One of these photographs is reproduced in Fig. 302. It will be seen that the deviation is a maximum in the green at the edge of the absorption band, while on the other side of the band the deviation is zero, *i.e.* the refractive index equals unity for this wave-length. The continuity of the dispersion curve can be traced through the absorption band, though the deviated spectrum is so broadened by diffraction that accurate measurements cannot be made in this region. The vertical lines are the bright spectrum lines of the cadmium spark which served as a source of light. The dispersion curve in the visible region is shown graphically in Fig. 303, the position of the absorption band being recorded as well.

Dispersion within the Absorption Band. — In this case the complete formula (13) must be used. Very few experimental determinations of the refractive index within the region of the absorption band have been made. Only very acute prisms can be used

and as the light is transmitted by only the thin edge, the resolving power suffers. Pflüger [1] was the first to undertake a careful series of observations of n and κ in the case of strongly absorbing substances, with a view of testing the dispersion formula near to and within the absorption band. He employed solid bi-prisms of small angle, obtained by evaporating an alcoholic solution of an aniline dye between a glass plate and a curved segment of a glass tube. On removing the segment of the tube, two prisms of small angle remained on the plate, the refracting edges being turned towards the centre. The refractive indices could be obtained by

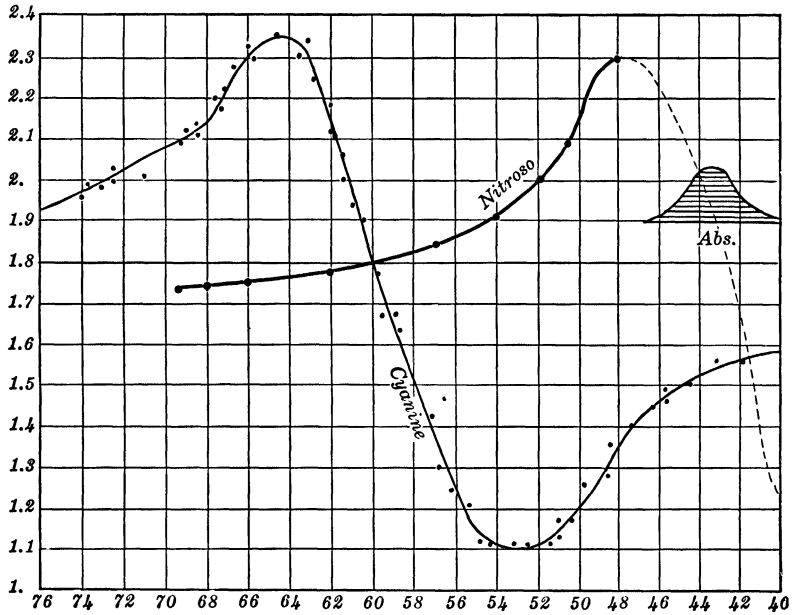


FIG. 303

means of these prisms even at the centre of the absorption band, though in this case the image of the slit of the spectrometer was greatly broadened by diffraction, since only the extreme edges of the prism transmitted the light. The values of the constant κ for the different values of λ were determined by means of the spectro-photometer, thin films of different thicknesses being used.

R. W. Wood and C. E. Magnusson [2] in 1901 made an investigation of the dispersion of cyanine, employing the solid prisms of the fused dye, described in the Chapter on Dispersion, for values

[1] *Ann. der Phys.*, *65*, 113, 1898.
[2] *Phil. Mag.*, *17*. 36

outside of the absorption band, and thin films in conjunction with an interferometer for the region of great opacity. Values were calculated from fringe shifts for ten different values of λ within the absorption band. The dispersion in the ultra-violet was determined by photography. Prisms of as small an angle as 30 seconds were also made by the fusion method, one glass plate being split off after solidification, and with these, values of n were determined through the region of absorption, which agreed closely with the values obtained with the Michelson interferometer. The curve obtained is shown in Fig. 303, the maximum and minimum values of the index n being 2.35 and 1.1. Complete absorption was found for all wave-lengths below 3720 even for a five-hour exposure. Van der Plaats[1] made very accurate measurements of the absorption and dispersion of solutions of certain dyes, and compared observed with calculated values, obtaining very close agreement.

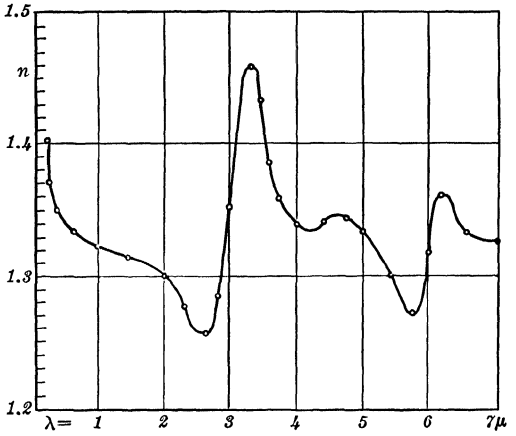


FIG. 304

The dispersion within the absorption band can also be determined by measuring the absorption and reflection coefficient from the formula

$$R=\frac{(n-1)^2+k^2}{(n+1)^2+k^2}.$$

Rubens and Ladenburg investigated the refraction of water from the ultra-violet to 18 μ by this method. The dispersion curve which they obtained is shown in Fig. 304, the ultra-violet values taken from Flatow's measurements. The value at 3.3 μ is higher than that in the ultra-violet, while at 2.5 μ it drops to 1.25. As we shall see presently quartz shows a similar relation.

Selective Reflection by Absorbing Media: Surface Color. — Attention was drawn by the author in articles on "Surface Color" [2] and "Surface Color of Selenium" [3] to the fact that the frequently

[1] *Ann. der Phys.*, *41*, 381, 1915.
[2] *Phys. Rev.*, June, 1902.
[3] *Phil. Mag.*, June, 1902.

occurring statement that the most powerfully reflected radiations were those most strongly absorbed was not only misleading but absolutely incorrect, and curves for the reflection, dispersion and absorption of cyanine calculated from Pflüger's values of n and κ were given. The centre of the absorption band of this substance is in the yellow, not far from the D lines, yet the color of the reflected light is purple, not very different in hue from that of the transmitted light. The spectrum of the reflected light showed a dark band in the green, centred at $\lambda = 5000$, due to the very low value of the refractive index at this point, in other words the color of the reflected light was determined by an abnormally low reflecting power, rather than by an abnormally high one. The reflecting power for the green light was found to be scarcely 2%, or only one-half that of a single glass surface.

Nevertheless for many years it appears to have been assumed that experimental determinations of the regions of powerful selective reflection in the infra-red, located the wave-length range most powerfully absorbed, or in other words determined the "free-period" of the absorbing oscillators. It was pointed out that if the dye was spread on the face of a glass prism of small angle, the reflection from the glass-to-dye surface showed the yellow color characteristic of the region of strongest absorption. In the following year it was shown that nitroso-dimethyl-aniline, with a heavy absorption band in the violet, reflects a very pure violet when in contact with glass, but nearly white light when in contact with air.

The best way of showing the effect is to coat one face of a prism of 5° to 10° with the absorbing medium and view the reflection through the opposite face, the advantage of a prism over a glass plate lying in the absence of reflection from the front surface, this light being thrown to one side. Nitroso-dimethyl-aniline (fused) shows a splendid violet color, in contrast to the complete absence of color when the reflection occurs at an air-nitroso surface. This blue color is not to be confused with the blue or violet color seen reflected at large angles of incidence through a Nicol prism, which is due to the dispersion of the angle of polarization. The white reflection from an air-nitroso surface results from the high value of the refractive index over the whole range of the visible spectrum. A strong solution of uranine (sodium salt of fluorescein) evaporated on the face of the prism gives a yellow-green reflection on the air side and a beautiful sky-blue color on the glass side. This method has been more recently used by others for the study of selective reflection in the ultra-violet. The condition necessary for selective reflection depending on k is a high value of the ab-

sorption coefficient, that is the waves must be stopped within a distance comparable to the wave-length.

Schaefer has drawn attention to the circumstance that the absorption bands in the visible and ultra-violet are produced by electrons instead of ions, and if one assumes the damping factor g for each to be of the same order of magnitude, the actual damping, measured by g/m (m=mass of electron or ion) will be several thousand times as great for an electron as for an ion, and the curve expressing the absorption coefficient k will be broad and flat, the effect of k^2 in the formula for reflection $R=[(n-1)^2+k^2]/[(n+1)^2+k^2]$ will be insignificant and *selective* reflection will not occur.

Selective Reflection in Ultra-Violet and Visible Spectra. — Martens [1] and Flatow [2] examined the reflecting power of various liquids covered by quartz plates and found evidences of selective reflection without, however, considering the factor depending upon the difference of the refractive indices of the media.

Schaefer [3] employing Flatow's method of multiple reflections within a quartz plate, the lower surface in contact with the liquid, and the upper surface silvered save for two narrow strips for the entrance and emergence of the light (Fig. 305), examined the spectra of the light reflected from various absorbing solutions. He took full account of all of the factors involved, and showed that any one of three distinct conditions might prevail, as shown in the figure in which the solid lines represent the dispersion curves of the absorbing media (dyes) and the dotted line the curve for the quartz.

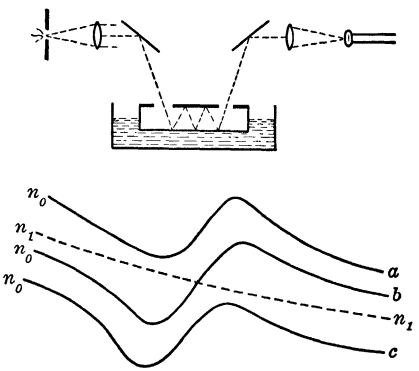


FIG. 305

It is clear from the figure that, in case "a" the position of maximum reflection is determined by the maximum value of n_1 since the difference in refractive index is greatest at this point: for case "c" the point of minimum value of n_1 locates the reflection, while in case "b," *two* maxima of reflection occur corresponding to the maximum and minimum values of n_1.

Schaefer found that an alcoholic solution of fuchsine illustrated

[1] *Ann. der Phys.*, *6*, 603, 1903.
[2] *Ibid.*, *12*, 85, 1903.
[3] *Zeit. für Phys.*, *75*, 687, 1932.

case "*a*" nitroso-dimethyl-aniline case "*c*" and aniline case "*b*," showing two maxima at 2500 and 3000, the absorption band extending from 2700 to 2800. Chinolin, with absorption between 2500 and 2600 gave also two reflection maxima at 2400 and 3100.

Thin films of the aniline dyes show surface color by reflection, as illustrated by the green reflection from letters written with purple ink. This is a phenomenon related to the infra-red selective reflection of crystals, but brought about by a different absorbing mechanism.

The reflection of an absorbing medium is represented by the formula $R=[(n-1)^2+k^2]/[(n+1)^2+k^2]$ if the reflection is from an air surface R being the ratio of the reflected to the incident intensity at normal incidence. The variation of R with the frequency can be calculated if we have given the dispersion and extinction curves. If the reflection occurs within glass of refractive index n' the formula becomes

$$R=\frac{(n-n')^2+k^2}{(n+n')^2+k^2}.$$

It is instructive to calculate the reflection curve for both types of reflection from Pflüger's determination of n and κ for cyanine remembering that $k=n\kappa$.

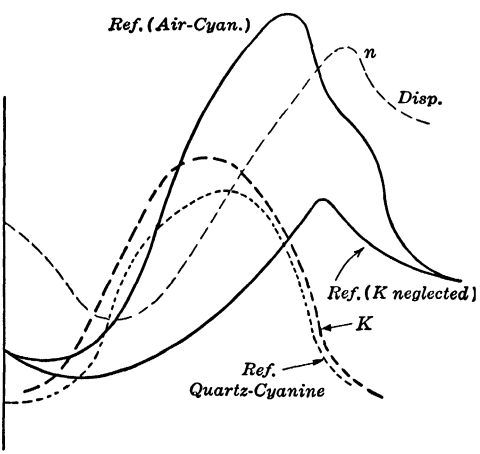


FIG. 306

In Fig. 306 three calculated reflection curves are shown, two real and one imaginary. The imaginary one was calculated from the refractive index alone, using the formula for transparent media. It parallels the dispersion curve very closely of course, and is given to show the effect of neglecting k, and the change which results with its introduction into the reflection formula. The dispersion and absorption curves are given as well, and the reflection curves for air to cyanine, and quartz to cyanine. It should be noted that the air-cyanine and the imaginary curve unite in the regions lying outside of the absorption band.

It was pointed out by Försterling [1] that the point of maximum

[1] *Ann. der Phys.*, *61*, 577, 1920.

reflection fell on the short-wave-length side of the centre of the absorption band in the case of media in which dispersion and absorption result from oscillators of a single frequency, that is a single absorption band.

This can be seen more easily from the following elementary considerations than from the equations given in his paper. If a single frequency only is involved in the dispersion of the substance, the refractive index will be less than unity on the high frequency side of the absorption band, and greater than unity on the other, as in the case of sodium vapor at the *D* lines. The shift of the region of maximum reflection to the short-wave-length side of the absorption band results from the values of the refractive index which are less than unity. If we calculate the reflection for regions not involving the narrow absorption band we can neglect k in the formula, and assuming a medium having indices varying from 0.4 close to the absorption band on the short-wave-length side to 1.6 on the other side, we find the following values:

n	R	R	n
1.6	.05	.2	.4
1.5	.04	.11	.5
1.4	.027	.06	.6
1.3	.017	.031	.7
1.2	.008	.012	.8

Rock salt in the infra-red has a dispersion curve not very dissimilar from that of the assumed medium. For example the values of 1.6 and .4 corresponding to the refractive index for two values of λ symmetrically located to the right and left of the absorption, give reflections of 5% and 20% respectively, which results from the circumstance that $(n-1)=.6$ in the former case and $-.6$ in the second, their squares being .36 in both cases while the denominators $(n+1)^2$ are $(2.6)^2$ and $(1.4)^2$ respectively.

The effect cannot be shown with sodium vapor of course, but recent work on selective reflection in the infra-red region has shown that there are many cases in which the shift to shorter wave-lengths is conspicuous. The shift in the opposite direction, to which attention has already been drawn, occurs, when all values of the refractive index are greater than unity.

An interesting case in which the region of selective reflection is shifted towards the region of shorter wave-lengths was observed by the author in the case of very dense mercury vapor enclosed in a bulb of fused quartz. A description of this experiment will be found on page 537.

Absorption and Selective Reflection by Crystals. — This subject, in which very rapid advances have been made in recent

years, may be said to have been opened up by the discovery (in 1897) of E. F. Nichols [1] working in Rubens's laboratory in Berlin, that quartz had a reflecting power, for certain wave-lengths in the infra-red, very nearly as great as that of the most highly reflecting metals, while in adjacent regions it reflected only 4% or less of the incident energy. This was a phenomenon analogous to the surface color of strongly absorbing media, such as many of the aniline dyes, which had been known for a long time, but the reflecting power was much greater than in these cases. This discovery led immediately to the development, by Rubens and Nichols [2] of the method of residual rays or Rest-strahlen, by which these strongly reflected frequencies were sifted out from the more intense radiations of shorter wave-length by multiple reflections from a number of surfaces of the same substance.

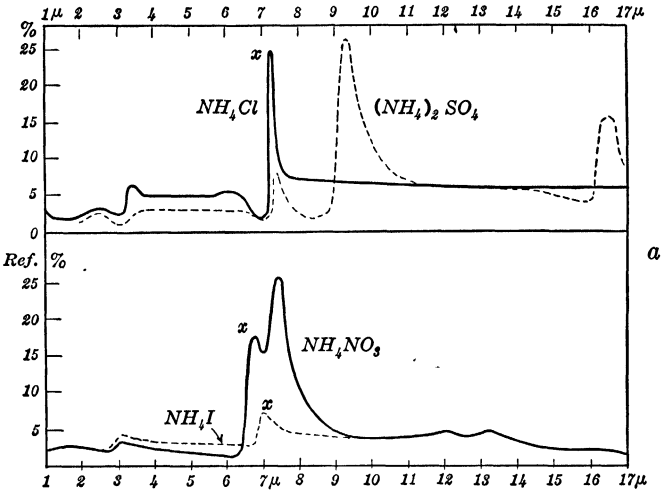
It was then supposed that atomic vibrations were concerned in the production of this powerful reflection, but no very definite theory was proposed to account for the difference in behavior of the infra-red absorption bands and those in the visible spectrum. The conception of two radically different types of absorption (other than electronic vibrations) seems to have been first conceived by Rubens and G. Herz,[3] as a result of their work on the absorption of quartz, fluorite and sylvite at different temperatures.

The absorption coefficient for certain regions of the spectrum decreased with lowering temperature approaching 0 as a limit for the absolute zero of temperature. In other regions the limit approached had a finite and fairly large value. This was the case for absorption in the region $7\ \mu - 15\ \mu$ and this type was assumed to result from vibrations of atoms within the molecule, while the other type, which showed the greater temperature effect, was attributed to vibrations between adjacent molecules. The present theory recognizes these two types of absorption but specifies the nature of the vibration. The low frequency vibration, which is the most sensitive to temperature, is considered as an oscillation of the anion-kation lattice, which gives rise to selective reflection in the region between 20 μ and 160 μ, as illustrated by Iceland spar (Ca—CO_3) in which the calcium lattice vibrates with respect to that of CO_3. Radiations reflected from a vibratory system of this type are studied by the method of residual rays which will be described presently. The other type of free-period, which corresponds to frequencies in the infra-red region between the visible spectrum and, say, 20 μ persists when the crystal is dissolved in water or

[1] *Wied. Ann.*, *60*, 401.
[2] *Wied. Ann.*, *60*, 418.
[3] *Berl. Ber.*, *256*, 1912.

fused and is attributed to the vibration of the atoms which make up the ion.

The resulting absorption bands are characteristic of the ions and any given band is found in the case of all salts in which the



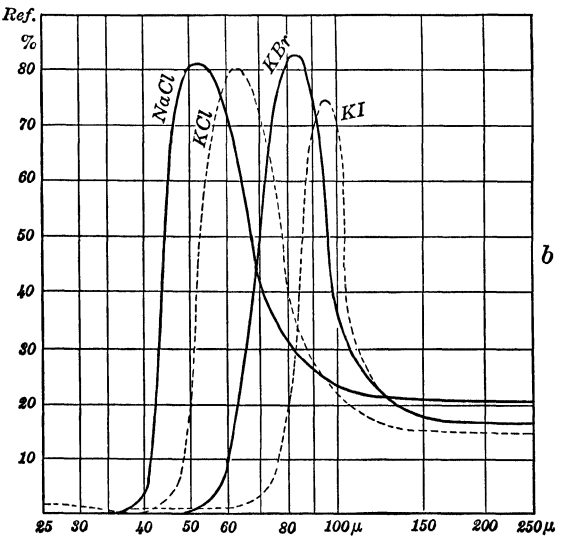


FIG. 307

ion, responsible for it, forms one constituent. The two types of absorption are shown in Fig. 307*a* from a paper by Reinkober.[1]

[1] *Zeit. für Phys.*, *3*, 1, 1920.

The band x due to vibration within the NH_4 radical is common to all. The SO_4 group gives bands at 9 and 16.3 μ, and the NO_3 group at 7.4 μ. In Fig. 307*b*, the selectively reflected radiations due to the vibrations of the crystal lattice are shown for a number of halogen salts of NH_4 plotted on a log scale of wave-lengths which compresses the long-wave-length side of the band making it appear less unsymmetrical.

Residual Rays. — The method of residual rays (Rest-Strahlen) was developed by Rubens and Nichols [1] in 1897 as a result of observations made by the latter of the absorbing and reflecting power of quartz in the infra-red. Rubens recognized that the very high reflecting power in certain regions of the spectrum (80 to 90%) in contrast to the very low reflecting power over the larger part of the spectrum (4%) could be made the basis of a method of isolating radiations of very long wave-length from the continuous spectrum of a white-hot body. Quartz was found to have two bands of metallic reflection, one at 8.5 μ, the other at 20 μ. The intensity at these points of the spectrum is small in comparison to the intensity in the visible and near infra-red, but by reflecting the radiation in succession from a number of flat plates of the substance, the intensity of the short-wave-length radiation was enormously reduced, while the feebler 8.5 μ and 20 μ radiations were reduced scarcely at all.

For example three reflections at 4% reduces the radiation to $(0.4)^2 = .00006$. In brief, the method consisted in reflecting the radiant energy, coming from a zirconia button heated in the oxyhydrogen flame, in succession from several polished surfaces of quartz. The energy after five reflections was examined with a wire diffraction grating, and found to consist principally of waves of length 8.5 μ. In addition to these waves, the grating showed that wave-length 20 μ was also present in the reflected energy, conclusive proof that there was a second absorption band at 20 μ.

The method is of great use in the study of dispersion for two reasons. First it enables the regions of strong absorption or rather selective reflection to be directly determined, and secondly it furnishes a means of obtaining approximately monochromatic radiation of long wave-length for the study of the dispersion and absorption of other substances. The region of strongest reflection does not, however, coincide exactly with the region of strongest absorption (spectral position of the free period) for the refractive index of the medium enters into the expression for the reflecting power, as well as the coefficient of absorption.

[1] *Weid. Ann., 60,* 418.

Interferometer Study of Residual Rays. — Another method for the study of residual rays was developed subsequently by J. Koch, and used by Rubens and Holnagel, a quartz interferometer being employed in place of the wire grating for measuring the wave-length of the selectively reflected radiation.

This investigation resulted in the discovery of heat-waves 96 μ in length, the longest observed at the time. The arrangement of the apparatus was as follows. Light from a Welsbach lamp, without a chimney, rendered parallel by a mirror, was passed through the interferometer. The plates of this instrument were of quartz, cut perpendicular to the axis, .6 mm. in thickness. Thin plates of quartz are fairly transparent for the very long waves, and on account of the high value of the refractive index in this region,

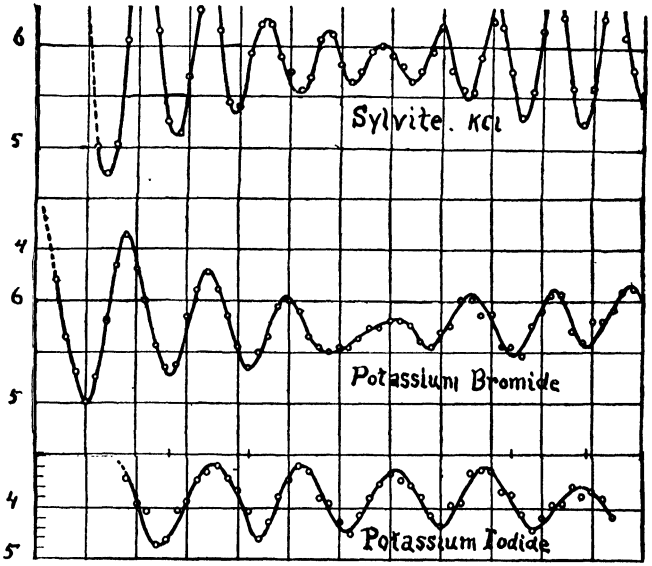


FIG. 308

possess a very high reflecting power. Previous investigations by Rubens and Aschkinass had shown that the refractive index for the waves reflected from sylvite ($\lambda = 56\ \mu$) is as high as 2.18, which gives a reflecting power of nearly 14%. After passage through the interferometer the rays suffer a fourfold reflection from the reflecting surfaces (which are made of crystal plates, or slabs of fused salts), and are then focussed upon the radiomicrometer. The quartz plates were first placed in contact, and the scale deflection observed. They were then separated by constantly increasing

amounts, readings being taken for each position. With properly adjusted plates, an interference maximum or minimum could be made to cover completely the sensitive surface of the recording instrument, and by plotting the scale deflections against the readings of the interferometer wheel an intensity curve was obtained precisely similar in every respect to the visibility curves obtained by Michelson with his instrument. The substances investigated were rock-salt, sylvite and bromide and iodide of potassium. The curves obtained from some of these substances are shown in Fig. 308. From the similarity to the visibility curves obtained with sodium light we recognize the presence of two wave-lengths. It was subsequently found that the duplicity of the bands was due to absorption by water vapor in the air.

From the known separation of the quartz plates for each maximum and minimum, the intensity curve of the radiation can be calculated. The longest waves were obtained from plates of fused iodide of potassium, which gives us Rest-Strahlen 95.6 μ in length.

Later on radiations of even greater wave-length were found. A list of residual rays obtained up to the present time is given in the following table:

Quartz	8.5 – 20 μ	Rubidium chloride	73.8
Calc spar	6.76, 28, 90 μ	Silver chloride	81.5
Lithium fluoride	17. μ	Thallium chloride	91.9
Sodium fluoride	35.8	KBr.	81.5
Thallium fluoride	48.6	Thallium brom.	117.
Fluor spar	2.4–31.6	KI	94
NaCl. Rock salt	52 μ	Silver brom.	112.7
KCl. Sylvite	63	Thallium iodide	151.8

A further improvement in the method was introduced by Czerny [1] for the purpose of avoiding the loss of energy by repeated reflections in cases of substances of only moderate reflecting power. By employing radiation polarized with the electric vector parallel to the plane of incidence and reflecting it from the substance at its polarizing angle for the shorter waves which are to be suppressed, these latter penetrate the surface and are lost by transmission or absorption, while the residual ray groups are reflected. As a polarizer Czerny used a selenium mirror or in some cases a plate of the substance under investigation.

Absorption and Selective Reflection of Quartz. — The optical properties of crystalline quartz, a material very extensively used in the construction of optical apparatus, are now very accurately known over a wide spectral range. Its band of ultra-violet absorption extends from the shortest wave-lengths of light which are known up to about wave-length 2000 for moderate thicknesses,

[1] *Zeit. für Phys., 16,* 321, 1921.

while the 1849 mercury line is passed by a plate a few millimetres thick, and 1500 by very thin plates. Nearly perfect transparency exists from this limit up to 2.5 μ, where an absorption band with a double maximum occurs. This band is equally strong in the case of fused quartz, hence it is not due to a vibration of the SiO_2 crystal lattice. It does not give rise to marked selective reflection. It was found as a reflection *minimum* in the case of a flat layer of powdered quartz by A. Trowbridge, while a maximum was found at 8.5 μ, the first residual-ray band. This is very interesting, and shows us the difference in the behavior of the two types of absorption bands. The one at 2.95 μ is weak, and only manifests itself when a considerable thickness is traversed. The quartz therefore shows "body color" by reflection, *i.e.* a color resulting from absorption, the energy penetrating deeply into the powdered mass, and finally emerging as the result of repeated reflection and refractions among the particles, robbed of energy of wave-length 2.95 μ. The band at 8.5 μ is metallic in character, and the energy is selectively reflected from the upper surfaces of the particles. If the particles were very small and the surface flat, specular or regular reflection would appear. Trowbridge used a rather coarse powder, and studied the diffuse reflection. The distinction between the two types of bands must be taken into account in all investigations made with surfaces which permit of the formation of "body color," otherwise the results will appear to be very anomalous in character.

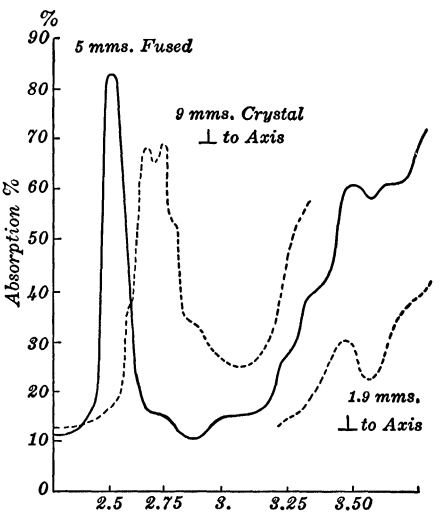


FIG. 309

The character of this band was investigated by Dreisch [1] for crystallized and fused quartz. The curves are shown in Fig. 309. The band at 2.7 μ for fused quartz 5 mms. thick is narrower and more intense than in the case of crystal quartz of nearly twice the thickness. The vibration concerned in the production of this band is obviously not due to the lattice, as it persists in the fused state, and is also found in glass. It may be due to a vibration of the SiO_2 molecule.

[1] *Zeit. für Phys.*, *42*, 426, 1927

Dispersion of Quartz. — The refractive indices of quartz have been very accurately determined from $\lambda = .1980$ in the ultra-violet to 313 μ in the infra-red. It was one of the substances first investigated with a view of verifying the Helmholtz dispersion formula, which as modified by Ketteler is very similar to the formula which has been deduced from electromagnetic considerations.

By measuring the dispersion in the visible spectrum and determining the constants, it is possible to calculate the positions of the absorption bands in the infra-red and ultra-violet, even if we cannot observe them.

In this way bands of absorption have been definitely located in the infra-red region of various media, and subsequently found by experiment.

Much of this early work was done by E. F. Nichols and Rubens. Discrepancies between the observed and calculated curve were found to be due to assuming a single absorption band in the infra-red. The position of the second band was calculated and on adding to the dispersion formula, a term representing the effect of this band perfect agreement between observed and calculated values of n resulted. The new bands of absorption were subsequently located by the method of residual rays. A fuller account of this earlier work will be found in former editions of this book. The refractive indices of quartz are given in the following table:

	λ	n Obs.
	.198	1.65070
	.274	1.5875
Visible Spectrum	.358	1.56400
	.434	1.553869
	.534	1.54663
	.656	1.541807
Sarasin	1.160	1.5329
	1.617	1.5271
	1.969	1.5216
	2.32	1.5156
	2.65	1.5081
	3.094	1.497
Nichols	3.63	1.48
	3.96	1.468
	4.20	1.457
	4.5	1.450
	5.	1.417
	6.45	1.274
	7.	1.167
	7.2	1.080
	7.4	1.00
	7.5	.930
	7.7	.798
	7.9	.611

λ	n
8.	.478
8.05	.366

Absorption Bands to 32 μ
Dispersion Anomalous Here

33 μ	2.13
52	2.03
63	2.00
83	1.99
110	1.96
313	1.94
∞	1.94

(Determined from observations of reflection by Rubens)

Rubens and Aschkinass [1] isolated radiations around 56 μ with a very acute quartz prism and in this way determined n by prismatic deflection, obtaining the value $n=2.18$ in fair agreement with the value found by reflection measurements.

The values given in this table are of interest in connection with the method of isolating very long heat-waves by means of quartz

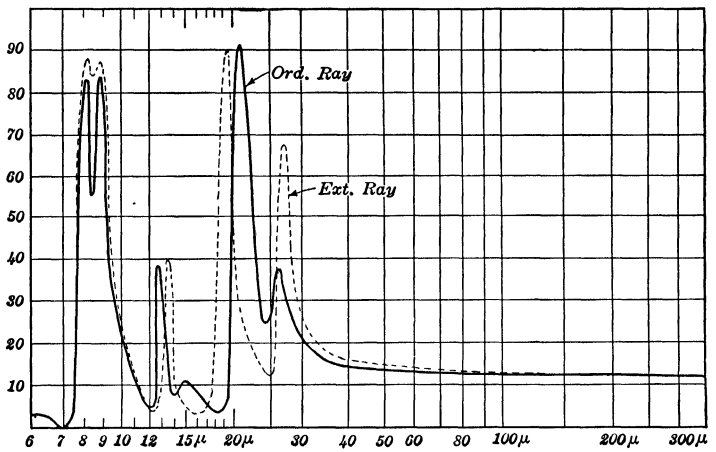


FIG. 310

lenses. The next bands occur at 8.5 μ and 12.3 μ, the former double. A second double band comes in the region between 18 and 25 μ. These are located by the powerful selective reflection which

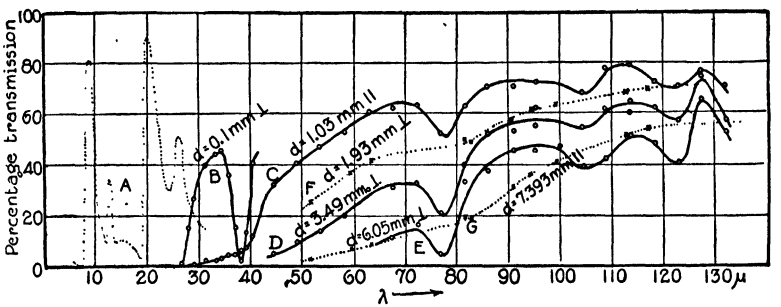


FIG. 311

they give rise to, and their location varies slightly with the direction of the polarization of the radiation, as shown in Fig. 310 from a paper by Rubens. The wave-lengths are plotted on a logarithmic scale.

[1] *Wied. Ann.*, *67*, 459, 1899.

These bands also appear in the reflection of fused quartz, and as there appears to be no doubt but what they are due to a vibration of the crystal lattice, it seems possible that even in the fused material there is a micro-crystalline structure. The transmission of crystal quartz with plates of varying thickness was studied by R. B. Barnes [1] as far as 130 μ. His curves are reproduced in

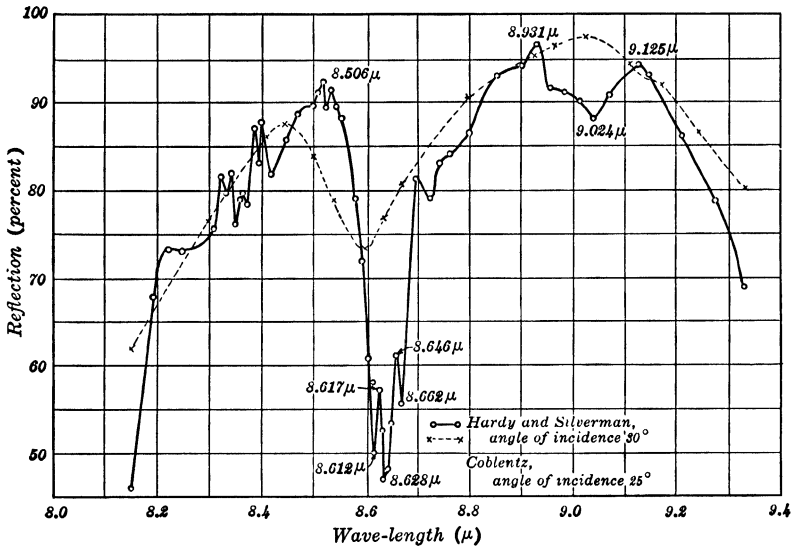


FIG. 312

Fig. 311 in which wave-lengths are plotted on a logarithmic scale, a method first suggested by Lord Rayleigh [2] and one in which every octave occupies the same space.

Fine Structure of Reflection Bands. — Some very remarkable results have been recently reported by Hardy and Silverman.[3] They measured the reflection of a quartz plate cut perpendicular to the axis with an echelette grating 5″ in diameter, made especially for the work by the author. The grating was ruled with 1312 lines to the inch, the sloping sides of the grooves reflecting the energy in the direction of the first order spectrum for 8.5 μ, and the resolving power was found to be about 80% of the full theoretical, or 4000. With this instrument the quartz double band at 8.6 μ was found to have a fine structure as shown in Fig. 312, two of the minima at the centre being separated by a distance of only .016 μ. A similar fine

[1] *Phys. Rev.*, *39*, 562, 1932.
[2] *Nature*, 1883.
[3] *Phys. Rev.*, *37*, 176.

structure was subsequently found by Silverman in the 7 μ band of calcite, reflecting the light from a cleavage plane. The same fine structure was found by transmission through powdered calcite, in this case recording minima instead of maxima. These experiments should be repeated using polarized light.

The Longest Heat-Waves: Focal Isolation. — The longest heat-waves are obtained by a method devised by Rubens and R. W. Wood.[1] The radiations from a Welsbach light *A* (without chimney) pass through a hole 1 cm. in diameter in a screen *B* made of two sheets of tin, and are received by the quartz lens *C*. For the longest heat-waves previously measured, Rubens found that the refractive index of quartz was about 2.2, and the screen *E*, also perforated with a 1-cm. hole, was accordingly mounted at such a

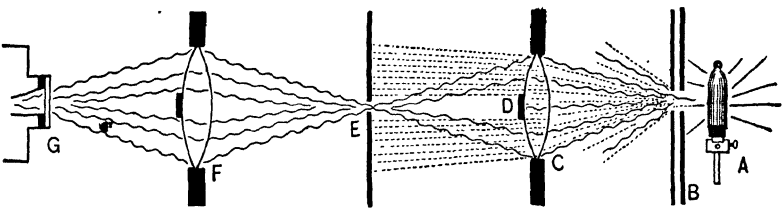


FIG. 313

distance from the lens that the two screens were at conjugate foci, for radiations for which the refractive index of the lens had this value (Fig. 313).

Owing to the much smaller value of the refractive index for the shorter heat- and light-waves, these actually diverge after leaving the quartz lens. The paths of the two types of rays are indicated in the figure, the long heat rays being represented by wavy lines, the short-wave rays by dotted lines. The aperture *E* is screened from the central portion of the short-wave ray-bundle by a small coin *D* fastened to the surface of the lens with wax. A single lens arranged in this way enables us to obtain at once the longest heat-waves of all, but to make matters sure, the second lens *F*, arranged in the same way, was used to focus the radiation upon the thermo-element of the radiomicrometer *G*. This still further purified the radiation, though its use was not imperative. To test whether the radiation is pure, *i.e.* whether we have eliminated completely the short waves, we have only to introduce a plate of rock salt in front of the screen *E*. This material is opaque for these long waves and if the deflection drops back to zero we can be sure that no radiations are present which can be passed by rock salt.

[1] *Berl. Berichte*, 1910. Also *Phil. Mag.*

By introducing the interferometer in the place of the screen E, it was found possible to measure the wave-length of the radiation isolated by the quartz lenses. The maxima and minima obtained as the distance between the quartz plates was increased were very pronounced, but only three could be recorded, owing to the wide spectral range of the transmitted radiation. The curve obtained is shown in Fig. 314, and its analogy to the intensity curve obtained in all interference experiments with white light is at once apparent. The mean wave-length of the radiation can be determined from the distance between the maxima, and was found to be 107 μ or more than one-tenth of a millimetre. There are still longer waves present, as we know from the damping of the intensity curve; in other words 107 μ is merely the centre of a rather broad spectral region isolated by the quartz lenses. We thus have experimental evidence of waves probably as long as 130 μ or possibly 150 μ in the radiation from the Welsbach lamp.

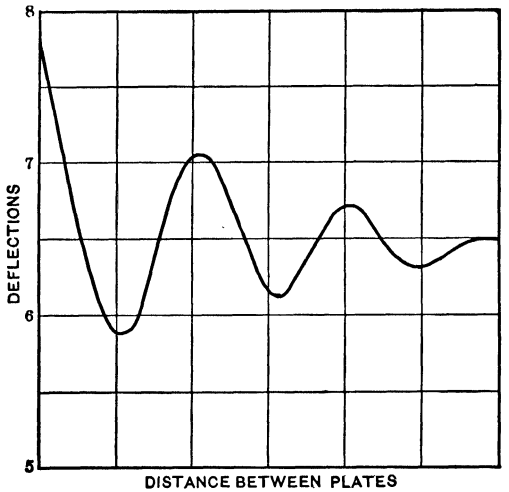


FIG. 314

With the same apparatus Rubens and von Baeyer [1] obtained radiations of still greater wave-length from a high-pressure quartz-mercury arc. There were two groups of radiation of mean wave-length 218 μ and 343 μ. Rubens [2] subsequently investigated this region with a wire grating finding for the longest waves from the Welsbach mantle 145 μ and from the Hg arc 325 μ.

It has recently been found that these waves are radiated by mercury molecules in rotation. Kroebel [3] showed that the radiation is absorbed by mercury vapor distilling from an arc, or excited by 4.9 volt electrons, and is destroyed by the introduction of hydrogen into the mercury arc, both of which experiments confirm the view

[1] *Berl. Ber.*, *339*, 666, 1911.
[2] *Berl. Ber.*, 1921.
[3] *Zeit. für Phys.*, *56*, 114, 1929.

of Franck and Grotrian [1] and Houtermans [2] that the radiation is a vibration-rotation band of a metastable mercury molecule formed by the union of a metastable excited atom and a neutral atom. It is now well known that hydrogen prevents the formation of metastable atoms. These very long heat-waves have some very curious properties. They pass through thin, black paper almost without absorption. It was in fact found possible to isolate them by means of a concave mirror combined with two or three sheets of black paper. The spectrum range obtained in this way was, however, very much wider than the range obtained with quartz lenses. Thick films of smoke absolutely opaque to light were also transparent to them.

The absorbing powers of quartz and water for these radiations are shown in the following table, q being the absorption coefficient in the formula $i/J = e^{-qd}$.

Quartz			Water		
Thickness of Plate	Absorption %	q	Thickness of Film	Absorption %	q
.6 mm.	5.4	.105	.0147 mm.	58	.059
2.00 "	18.7	.103	.0294 "	77.5	.051
3.03 "	27.	.104	.0442 "	88.2	.048
4.03 "	33.6	.102			
7.26 "	51.7	.100			
12. "	67.6	.094	Extinction coef. $\kappa = \frac{q\lambda}{4\pi}$		.420
14.66 "	69.4	.081			

The absorption of water was determined by separating the interferometer plates a known distance, measured by sodium fringes, and then allowing a drop of water to creep in between them, watching the fringes to make sure that no change in the distance resulted from capillary attraction. The extinction coefficient of water for the rays from KBr($\lambda = 87\ \mu$) is $\kappa = .66$, while for the new rays ($\lambda = 107\ \mu$) it is only .42.

Absorption of Fluorite, Rock Salt and Sylvite. — Fluorite is transparent between wave-lengths 9 μ and a point in the ultra-violet as low as 1200 A.U. in the case of the purest crystals. Rock salt is transparent between 1720 A.U. and 20 μ and sylvite (KCl) between 1810 A.U. and 25 μ.

Dispersion of Fluorite, Sylvite and KBr. — Fluorite is transparent from a region in the ultra-violet around $\lambda = 1200$ A.U. to about 10 μ in the infra-red. Its transparency is practically 100% up to 8 μ where it drops to 85% for a plate 1 cm. thick. At 9 μ

[1] *Zeit. für Phys.*, *4*, 89, 1921.
[2] *Zeit. für Phys.*, *41*, 140. 1927.

it transmits 50% and at 10 μ 15%. Its dispersion in the infra-red is shown by the following table:

FLUORITE

λ	n	λ	n
.5890	1.434	5.482	1.378
.8840	1.429	7.072	1.368
1.473 μ	1.4264	8.25	1.344
2.058	1.4236	8.84	1.331
3.241	1.416	9.43	1.316
4.714	1.402		
5.304	1.395		

Sylvite, or natural KCl is transparent from 1810 A.U. to 22 μ. Large crystals of this substance have been prepared by R. Pohl and are now available for spectroscopic work.

The dispersion which was measured by Paschen to 11 μ, and by Rubens with a prism of 12° as far as 22.5 μ, is as follows:

SYLVITE

λ	n	λ	n
.5893	1.490	5.893 μ	1.469
.8840	1.481	8.250	1.463
1.178 μ	1.478	10.018	1.456
1.768	1.476	12.965	1.443
2.357	1.475	14.144	1.437
2.946	1.474	15.912	1.426
2.356	1.473	17.68	1.414
4.714	1.471		
5.304	1.470		

Potassium bromide has also been prepared by Pohl as single crystals of large size suitable for prisms. It can be used to 30 μ. Its dispersion, determined by Gundelach [1] follows. The curve is shown in Fig. 315 which is typical for all three substances.

λ IN μ	n
.769	1.5492
.982	1.5437
1.179	.5414
1.768	.5385
2.357	.5367
3.536	.5350
4.714	.5359
5.893	.5325
8.250	.5290
10.018	.5255
11.768	.5209
12.965	.5177
14.143	.5146
15.910	.5080
18.100	.4983

KBr Dispersion — E. Gundelach, *Zeit. für Phys.*, *66*, 775, 1930.

[1] *Zeit. für Phys.*, *66*, 775, 1930.

Effect of Density and Pressure on Absorption Lines and Bands. — Liveing and Dewar [1] observed that the absorption of pure oxygen at a pressure corresponding to its partial pressure in the atmosphere was less than the absorption of an equivalent amount of the gas mixed with enough nitrogen to raise the total pressure to 760 mms.; they found a similar effect by compressing the oxygen, thus establishing an exception to Beer's law. Ångström [2] in 1901 made a more thorough investigation of the matter with carbon dioxide, which has two strong absorption bands in the infra-red at wave-lengths 2.8 μ and 4.3 μ. The apparatus consisted of a

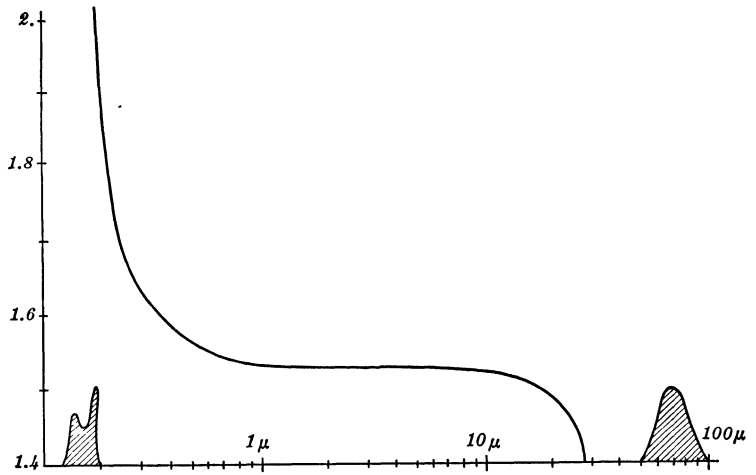


FIG. 315

long glass tube divided by a rock-salt plate into two compartments, 3 and 30 cms. in length respectively. The two compartments could be placed in communication by a glass tube furnished with a stopcock. The investigation was carried on in the following way:

(*a*) The two compartments were exhausted and the spectrum of the transmitted light investigated with a thermo-element.

(*b*) The small compartment was filled with carbonic acid at a pressure p_1 and the absorption a_1 determined, the other compartment remaining vacuous.

(*c*) The stopcock between the two compartments was now opened, and the gas was allowed to fill the entire tube. The pressure is now $p_2 = p_1 \frac{3}{33}$. The absorption a_2 was now determined and found to be less than a_1.

(*d*) A non-absorbing gas was now introduced into the tube until

[1] *Proc. Roy. Soc.*, *46*, 222, 1889.
[2] *Ann. der Phys.*, *6*, 163.

the total pressure of the gas mixture had the same value p_1, as that of the carbonic acid in experiment a. The absorption in this case a_3 was found to be equal to a_1, from which the following law was deduced: As the pressure decreases, the product of pressure and length of column remaining constant, the absorption decreases, returning, however, to its original value when an amount of a non-absorbing gas is added sufficient to give to the mixture a pressure equal to the original pressure. The absorption curves obtained

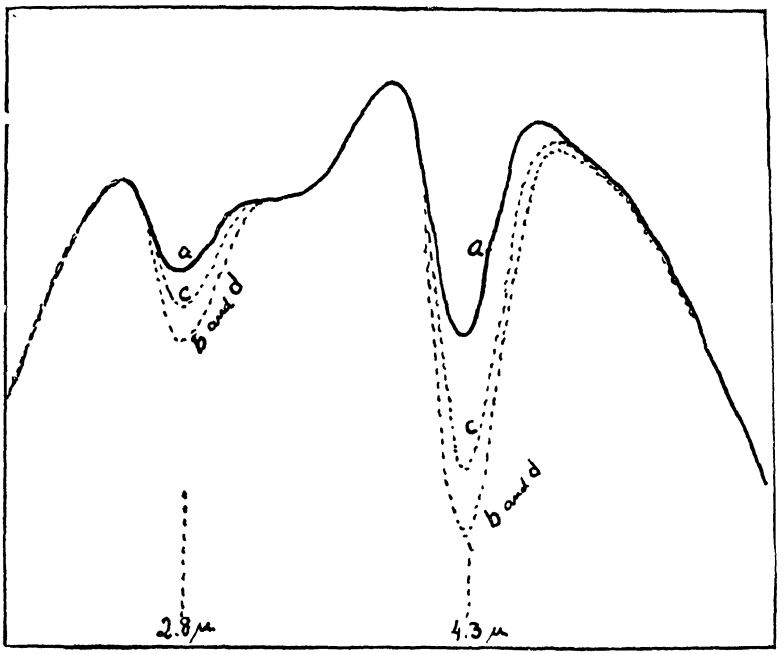


FIG. 316

in this way are shown in Fig. 316, in which the letters a, b, c and d refer to the cases mentioned. Similar results were obtained with carbon monoxide. The presence of the CO_2 bands in the curve a are due to the gas normally present in the air of the room. Rubens and Ladenburg [1] found that the band at 14.7 μ behaved in a quite different manner, Beer's law holding, and the addition of another gas producing no effect.

Experiments by the author commenced in 1900 showed that the D lines of sodium were enormously widened by the introduction of hydrogen, but the experiment was very inconclusive, as the hy-

[1] *Ver. d. D. Phys. Ges.*, *3*, 170, 1909

drogen prevented rapid diffusion of the sodium to the cooler parts of the tube and consequently increased the number of atoms present in the absorbing column. Results less open to criticism were obtained with mercury vapor at room temperature in a steel tube 3 metres long, closed with quartz windows. In vacuo the 2536 absorption line was much narrower and less conspicuous than when air was introduced.

Fuchtbauer [1] made similar experiments with iodine vapor with an apparatus capable of resolving the component lines of the broad absorption band. He found that the individual lines were broadened by the admission of another gas, but that the absorption at the centre of each line was weakened. For a thick layer of vapor this would mean an increased total absorption.

Reflection and Dispersion by Rock Salt. — The most recent and instructive investigation of the optical properties of a medium in relation to the dispersion formula are those carried out by Czerny [2] for rock salt, which is especially adapted to the verification of the theory on account of the simplicity of its chemical and crystalline structure. The reflection curve is shown in Fig. 317

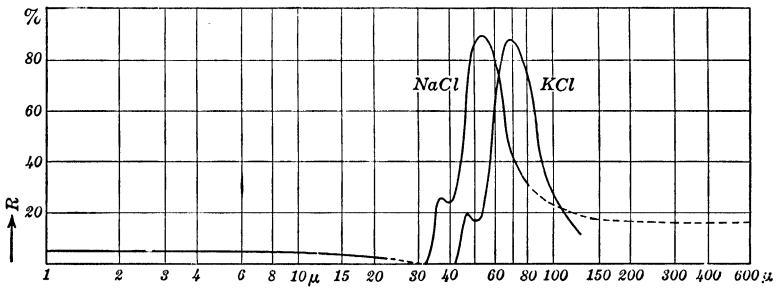


FIG. 317

together with that of sylvite, KCl. Rubens determined its reflection at 23 μ and 33 μ as 2.1% and 1.7%. The portion between 30 and 70 μ is from Czerny's work with homogeneous radiation obtained with a wire diffraction grating, while the region between 70 and 300 μ was investigated by Rubens with residual rays. The region between 23 and 35 μ has not been completely investigated.

The older curves for the residual rays, showing two maxima of nearly equal height are incorrect, the minimum being due to absorption by water vapor in the air. All infra-red work should be done in absolutely dry air if possible.

[1] *Phys. Zeit.*, *12*, 722, 1911.
[2] *Zeit. für Phys.*, *65*, 600, 1930.

The reflection formula $R=[(n-1)^2+k^2]/[(n+1)^2+k^2]$ if solved for n gives

$$n=\frac{1+R}{1-R}\pm\sqrt{\frac{4R}{(1-R)^2}-k}.$$

For cases in which k^2 is small in comparison to $4R/1-R^2$ Czerny develops the square root term in powers of k^2 obtaining

$$n_1=\frac{1+\sqrt{R}}{1-\sqrt{R}}-\frac{k^2}{4}\frac{1-R}{\sqrt{R}}-\frac{k^4}{64}\frac{(1-R)^3}{\sqrt{R}}$$

$$n_2=\frac{1-\sqrt{R}}{1+\sqrt{R}}+\frac{k^2}{4}\frac{1-R}{\sqrt{R}}+\frac{k^2}{64}\frac{(1-R)^3}{\sqrt{R}}$$

n_1 being for the + and n_2 for the − sign of the ± sign of the $\sqrt{\ }$ term, the former giving $n>1$, the latter $n<1$. For vanishingly small values of k the two n values are reciprocal. This simplifies the calculation as one has only to determine $(1+\sqrt{R})/(1-\sqrt{R})$ and $1-R/\sqrt{R}$ to determine both n values.

For the determination of k the transmission, which is diminished by reflection as well as by absorption, is measured. It is not sufficient to measure the intensity of the radiation transmitted through a single thin plate, and allow for the loss due to the reflection at the two surfaces, for interference may be taking place, as in the case of thin films with visible light.

Czerny prepared plates by shaving a rock-salt crystal with a microtome constructed especially for the purpose, arriving at a

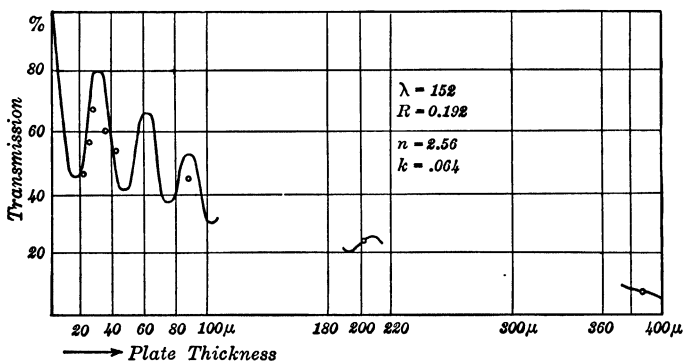


FIG. 318

thickness as small as 8 μ. The transmission curve for residual rays from thallium iodide (152 μ) is shown in Fig. 318. Ordinates represent transmission in %, abscissae thickness of the salt plate. Starting with 100% transmission for zero thickness, since in this

case there is no reflection loss (Newton's black spot on soap-films) he found a minimum at a thickness of 20 μ, a maximum at 30 μ, etc. The refractive index, calculated from the reflection coefficient .192 is $n=2.56$, and the wave-length in the medium is therefore $152/2.56=60$ μ. The first minimum of transmission, corresponding to the first maximum of reflection should be for a thickness equal to one-quarter of 60 μ or 15 μ. Czerny's lowest value was for a plate of 20 μ thickness. The minimum for reflection should be for $60/2$, and maximum transmission was found for 30 μ. And now comes a very important point. The value of n calculated from the reflection may be either greater or less than unity according as we employ the + or − sign in the formula. The curve of Fig. 318 shows that the value of n is much greater than one: if less than one, the wave-length in salt would be greater than in air, and the maxima and minima would be much farther apart.

With increasing plate thickness the maxima and minima become less pronounced, owing to the wide spectral range of the residual rays, and k was determined from the observation with the thicker plates in which case interference played no part.

Determinations were made with the following wave-lengths: 152, 94, 70, 65, 52, 43 and 38 μ. As an interesting example of the way in which the results are to be interpreted we may take the case of the determinations of n and k for $\lambda=38$ μ, obtained with a wire grating applied to residual rays from strontianite. The reflection coefficient was found to be $R=0.27$ from which the refractive index was calculated to be either $n=3.14$ giving $k=0.3$ or $n=0.40$ giving $k=0.5$. Observations were made with plates of four different thicknesses and the results are tabulated as follows:

d	D	k FOR $n=3.14$	k FOR $n=0.4$
24 μ	0.045	.32	.43
19	.06	.38	.49
14	.14	.28	.49
8	.345	.09	.53
8	.38	.04	.50

This shows that the value $n=0.4$ is the correct one, as the calculated values of k must come out the same for plates of all thicknesses.

The following formula, which represents the dispersion from the ultra-violet to 22 μ contains no damping factor.

$$n^2=1+\sum_i \frac{C_i}{\nu_i^2-\nu^2}=1+\sum_i \frac{\frac{1}{c^2}C_1}{\frac{1}{\lambda_i^2}-\frac{1}{\lambda^2}}$$

and was checked by prism measurements by Fuchs and Wolff,[1] ν_i is the frequency of the ith natural vibration, and c_i the corresponding intensity constant. They took three natural frequencies in the ultra-violet and one in the infra-red. Their values of the constants are as follows:

C_1	$3.99 \cdot 10^{30}$	$\lambda_1 = 347$ A.U.
C_2	$7.68 \cdot 10^{30}$	$\lambda_2 = 1085$ A.U.
C_3	$.972 \cdot 10^{30}$	$\lambda_3 = 1584$ A.U.
C_4	$8.37 \cdot 10^{25}$	$\lambda_4 = 61.67\ \mu$

The complete formula with constants is as follows, the last term representing the infra-red band.

$$n^2 = 1 + \frac{43.33}{830.5 - \frac{1}{\lambda^2}} + \frac{85.33}{84.95 - \frac{1}{\lambda^2}} + \frac{10.8}{39.86 - \frac{1}{\lambda^2}} + \frac{9.30 \cdot 10^{-4}}{2.629 \cdot 10^{-4} - \frac{1}{\lambda^2}}.$$

For all values of $\lambda > 14\ \mu$ the sum of the three ultra-violet terms for all values of λ is 1.3276 (for $5\ \mu$ it is only 1.3283) consequently the formula can be reduced to

$$n^2 = 2.3276 + \frac{9.30 \cdot 10^{-4}}{2.629 \cdot 10^{-4} - \frac{1}{\lambda^2}} \quad (\mu \text{ as unit of } \lambda).$$

For $\lambda = \infty$ this gives as the refractive index for infinitely long waves

$$n_\infty^2 = 5.865 = \epsilon \text{ the dielectric constant.}$$

As an illustration of the accuracy with which this formula represents the results of observations we have from Czerny's paper, Fig. 319*a* for the refractive index and *b* for the reflection coefficient R. The curve marked "observed" was calculated from observations of R, the other two curves from the dispersion formula, in one case taking only a single infra-red term, in the other case two terms, corresponding to the double maximum of the reflection curve. The latter shows a very close agreement with the observed curve.

On the long-wave-length side the results are still more interesting:

λ	n (Cal.)	R (Cal.)	R (Obs.)
$300\ \mu$	2.454	17.7%	17.9
117	2.688	20.9	19.9
94	2.922	24.0	24.3
83	3.198	27.4	27.5
70	4.259	38.4	39.0
65	6.149	51.9	56.5

In the above table the values of n were calculated from the above formula and from these values the reflection was calculated

[1] *Zeit. für Phys.*, *46*, 506.

by the formula $R=(n-1)^2/(n+1)^2$. The results are seen to agree very closely with the observed values of R given in the last column. The curves are shown in Fig. 319*b*. The curves marked "with

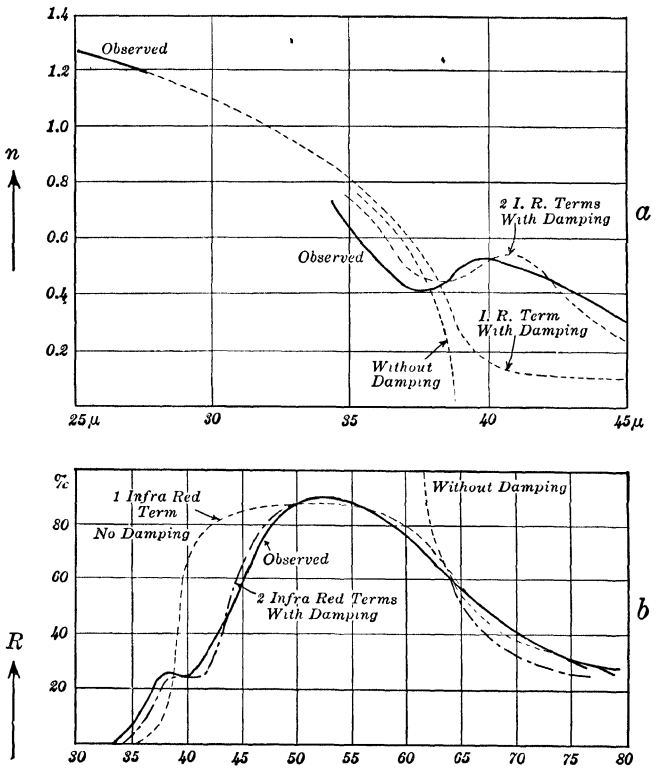


FIG. 319

damping" were calculated by the formula in which the damping factor had been introduced

$$n^2(1-ik)^2=2.3276+\frac{C_4(\nu_4^2-\nu^2)}{(\nu_4^2-\nu^2)^2+b_4^2\nu^2}$$

or separated into the real and imaginary parts

$$n^2(1-k^2)=n^2-k^2=2.3276\,\frac{C_4(\nu_4^2-\nu^2)}{(\nu_4^2-\nu^2)+b_4^2\nu^2}.$$

In a subsequent investigation made by R. B. Barnes and Czerny [1] the transmission of much thinner films of salt, deposited by sub-

[1] *Zeit. für Phys.*, *72*, 447, 1931.

limation on very thin films of celluloid in a high vacuum, were measured. In this way films from .17 to 3.5 μ were examined, and the transmission measured continuously right through the band of absorption. Curves for .17 μ and 3.6 μ are shown in Fig. 320. On the short-wave-length side two secondary minima of transmission were found, the one at 40 μ previously found by Czerny and a new one at 50 μ. The main minimum was accurately located at 61 μ giving the first experimental determination of the exact centre and general form of the absorption band, and the numerical value of the natural frequency of the oscillation of the crystal lattice agreeing closely with the value 61.7 calculated by Fuchs and Wolff. It is interesting to note that, in the case of the .17 μ film, the transmission is 100% at 40 μ in spite of the high value of the reflection coefficient at this point as determined with a thick plate (25%). It is analogous to the black-spot on soap-films of course, resulting from the phase-difference of 180° between the streams reflected from the two surfaces. The thickness of the salt-films were determined from the order of interference color exhibited either by eye observation or in the case of the thicker films by observing the number of dark bands in the spectrum of the reflected light. The celluloid was so thin (0.1 μ) that it gave only a pale gray color with no trace of the straw yellow, the first color to appear as the thickness increases. They found that the thickness in μ was very nearly represented by one-third of the number of dark bands seen in the spectrum, and were thus able to control the thickness by watching the reflected light with a spectroscope during the process of sublimation. It will be remembered that the maximum reflection is at 5.2 μ, there being a shift of 9 μ from the point of the natural frequency at 61 μ, which, as we have seen, is due to the fact that n is very much less than unity in this region.

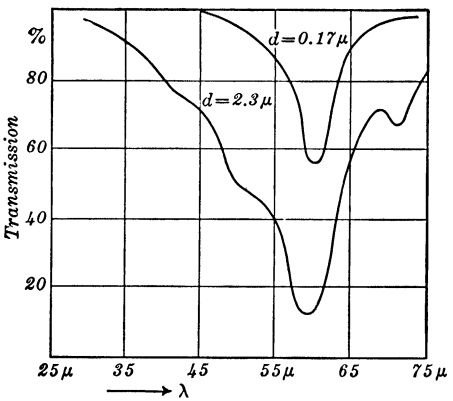


FIG. 320

The final values of the optical constants of rock salt, from the ultra-violet to 1300 μ are given in the table on page 535.

Selective Reflection by an Absorbing Gas. — The first investigation of this subject was made by the author many years before

λ	n	λ	k	n	λ	R	n
2.36 μ	1.526	35 μ	.17	.64	65 μ	17.9	6.15
5.01	1.519	36	.22	.50	70	19.9	4.26
7.66	1.495	37	.34	.42	83	24.3	3.20
10.02	1.472	38	.50	.40	94	27.5	2.92
14.14	1.460	39	.60	.47*	117	39.	2.69
17.93	1.415	40	.68	.52*	300	56.5	2.45
22.30	1.340	41	.70	.48*			
		43	.85	.40			
		46	1.17	.20			

Transmission

λ	D	k	
324 μ	.15	.0079	Nichols and Tear with electric waves, *Astrophysical Journal*, 61, 36.
420	.17	.0094	
1300	.61	.0020	
300		.029	(Rubens).

(Agreement between Rubens and Nichols and Tear poor; should be repeated.)

The values of n indicated by an * indicate a small maximum on the dispersion curve resulting from the small second maximum on the reflection curve.

the advent of the quantum theory, in the belief that a gas such as sodium, which, at low pressure, gives off a diffuse resonance radiation, might at high pressure show selective reflection for the very narrow spectral region covering the absorption line, the lateral diffusion of the light being prevented by interference, owing to the proximity of the resonators, the secondary waves coöperating to form a reflected wave. Nothing could be accomplished with sodium vapor, but the phenomenon was finally observed with mercury vapor in a thick-walled bulb of quartz.

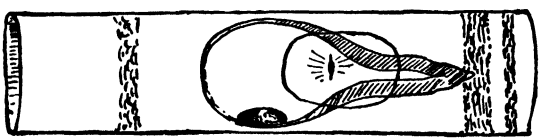


FIG. 321

To separate the images formed by reflection from the inner and outer surfaces of the bulb, the neck was made of very thick-walled tubing, so that the wall of the bulb was prismatic as shown in Fig. 321.

A good-sized globule of mercury was placed in the bulb which was then highly exhausted and sealed. The bulb was mounted in a small tube of thin steel provided with an oval aperture in the side: the ends were closed with disks of asbestos board, one of which supported the neck of the bulb. The steel tube was heated by two Bunsen burners, usually to a full red heat, and the mercury arc placed as close as possible to the aperture and a little to one side, so that its image appeared reflected in the tapering neck of the bulb,

as shown in the figure. The flame of the burner must play over the aperture to prevent condensation of mercury drops at the point where reflection occurs.

By properly choosing the direction in which the bulb was viewed, the reflection from the outer surface disappeared, and the slit of a small quartz spectrograph was directed towards the bright image of the arc reflected from the inner surface of the wall.

A number of photographs of the spectrum of the reflected light were taken, the first with the bulb cold, the succeeding ones at gradually increasing temperatures. It was found that the relative intensity of the 2536.7 line in the spectrum of the reflected light increased rapidly as the temperature of the bulb increased.

This experiment showed that the mercury vapor reflects light of this particular wave-length in much the same way as would a coating of silver on the inside of the bulb.

Experiments were next undertaken to ascertain how nearly the frequency of the light must agree with that of the absorption band in order that metallic reflection should take place.

It was found that the spectrum of the iron arc showed a group of closely packed lines exactly in the region required, and it was accordingly substituted for the mercury arc, and photographs of the reflected image taken with the bulb cold and heated to different temperatures. A very remarkable discovery was at once made, for it turned out that the iron line which was metallically reflected (2535.67) was about one Ångström unit on the short-wave-length side of the absorption line. As the temperature and vapor density increased, a second iron line was strongly reflected, this one coinciding almost exactly with the absorption line. It is in reality a double line, with wave-lengths 2536.90 and 2537.21.

To make absolutely sure that no error had occurred, the spectrum of the iron arc was photographed after passing the light through mercury vapor at different densities.

The iron line which first disappeared was the double one, which was not reflected until the mercury vapor was at its greatest density. The line which was metallically reflected by the vapor at a lesser density was not absorbed by the vapor, even when its density was so great that four or five lines on the long-wave-length side of the line first absorbed were completely blotted out.

The explanation of the more powerful reflection of the iron line of shorter wave-length is as follows:

The 2536.7 line shows powerful anomalous dispersion. The refractive index, in its immediate vicinity, on the short-wave-length side is much below unity, probably as low as .5 or even less close to the line. In the case of light going from a rare to a dense medium,

a high value of the index for the latter is accompanied by strong reflection. When, however, the ray goes from dense to rare (quartz to mercury vapor), as in the present case, it is a low value of the index for the latter which is accompanied by strong reflection. Now the index for mercury vapor is not far from unity for the entire spectrum with the exception of very narrow regions bordering the 2536 line. On the long-wave-length side the index may be about that of quartz, consequently there is no reflection at all here, at least none depending upon the relative values of n. On the other side, where the index is considerably below unity, the reflection will be much more powerful than when the bulb is filled with air. This means that the band of metallic reflection will be shifted slightly towards the region of *shorter* wave-length with respect to the true position of the absorption line, as was found to be the case.

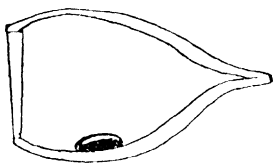


FIG. 322

In a subsequent investigation made in collaboration with M. Kimura [1] an interesting modification of this experiment was tried with a quartz bulb furnished with a slightly prismatic plate of optically worked fused quartz as a reflecting surface, Fig. 322. With this the beams reflected from the inner and outer surface were widely separated. It is obvious from the previous experiment that if we employ the light of two frequencies, one slightly higher and the other slightly lower than the frequency of the 2536 line, the former will be powerfully reflected and the latter not at all. This condition was realized by employing as a source of light a quartz-mercury arc operated at a potential just sufficient to distinctly double the 2536 line by self-reversal, the proper potential being 80 volts, determined by making a number of exposures with the quartz spectrograph, varying the resistance in circuit with the arc. The experiment was made as follows: The light of the lamp running at 80 volts was reflected from the inner surface of the prismatic plate of the quartz bulb into the quartz spectrograph, the slit of which was reduced to a length of 1 mm. by a diaphragm

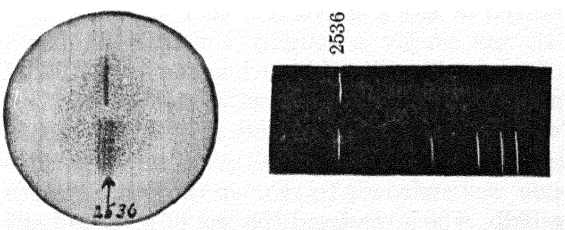


FIG. 323

[1] *Phil. Mag.*, *32*, 329, 1916.

which could be raised by a micrometer-screw. An exposure of one minute was given: the slit diaphragm was then raised 1 mm., the quartz bulb raised to a red heat by a Bunsen burner, and a second exposure of fifteen seconds made. Fig. 323 shows the result of the experiment. The reversed 2536 line appears as a doublet and is indicated by an arrow, the faint companion line, 2534 on the short-wave-length side appearing to its left. This was the exposure made by light reflected from the cold bulb. Above it we have the exposure made with the hot bulb. The light reflected from the hot bulb is seen to consist solely of the short-wave-length component of the doublet (widened and reversed 2536 line), for which the reflecting power of the quartz-mercury vapor surface is high. The long-wave-length component has disappeared entirely, owing to the very low value of the reflecting power for this frequency. The width of the doublet is about 0.8 A.U. It is perhaps worthy of mention that we have here a rather efficient method of isolating from the total radiation of a quartz-mercury arc running at a moderately high temperature, a single line of wave-length about 0.4 A.U. less than that of the 2536 line of a similar lamp running at a low temperature, which might be used in the study of the transition from resonance radiation to Rayleigh scattering. At the right of Fig. 323 we have the spectrum of the reflected radiation above and the total mercury spectrum below. In this same investigation the transition from resonance radiation to selective reflection was more carefully studied.

The diffuse resonance radiation as a function of the density of the mercury vapor in the bulb was first investigated. The light of a water-cooled quartz-mercury arc was passed through a quartz monochromator arranged to give a convergent cone of 2536 monochromatic light. It was simply a roughly constructed quartz spectroscope with a very wide slit and no telescope-tube. The image of the slit formed by the 2536 rays was located in space by means of a strip of uranium glass, and the bulb mounted in such a position that the image fell upon the centre of the prismatic plate The dispersion was sufficient to remove the other images of the slit from the bulb, which obviated the use of a second slit and lens for obtaining the monochromatic illuminating beam. The quartz bulb was mounted over a chimney of thin sheet iron, with a Bunsen burner at its base, and the temperature determined by a nitrogen-filled mercury thermometer, the bulb being in contact with the upper surface of the quartz bulb. A camera of very simple construction, furnished with a quartz lens, was focussed upon the bulb, the process consisting in first focussing it with uranium glass upon the image in space of the 2536 line formed by

the monochromator, and then measuring the distance between the lens and the image in space. This gives us the proper distance of the bulb from the lens to secure a sharp focus.

The arrangement of the apparatus in this experiment is shown in Fig. 324*a*, the rays reflected from the two surfaces of the prismatic plate falling to one side of the lens. In this way we obtain only an image of the scattered resonance radiation from the bulb.

The photographs obtained at different temperatures are reproduced in Fig. 324*b*. The left-hand photograph at room temperature and mercury pressure .0017 shows the bulb filled with resonance

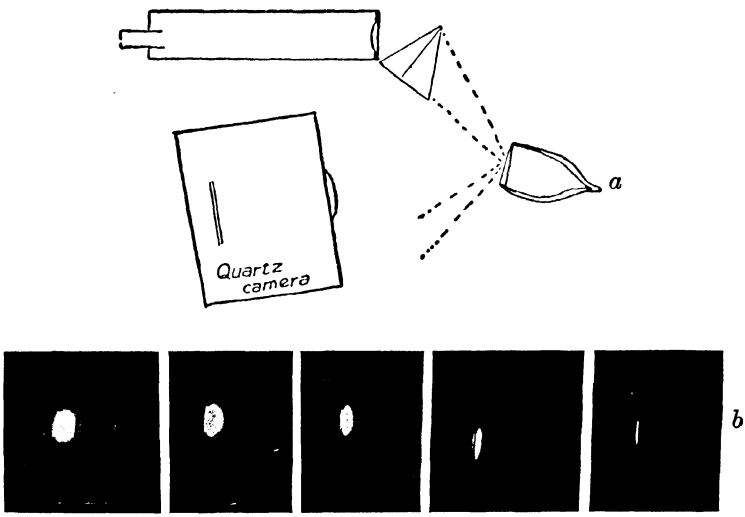


FIG. 324

radiation, the light from the directly excited atoms illuminating and exciting the entire mass of vapor. In the next, a temperature of 40° and .0057 pressure, the radiation is chiefly from the front surface of the bulb, though there is considerable lateral spread of the image of the slit. The following three exposures at temperatures of 76, 100 and 175, pressures .075, .276 and 11.0 mms. show the progressive drawing in of the region of secondary excitation until we have only a narrow line, the image of the slit of the monochromator.

At a temperature of 200°, pressure 18 mms. the intensity of the slit image decreased to about half value, and at 250° and 76 mms. pressure to about one-tenth, while at 270° there was no trace of the image. This is the point at which selective reflection begins as was shown by the following experiments:

In the first experiment the quartz bulb was placed a little inside of the focus of the monochromator, so that the incident radiations came to a focus after reflection from the prismatic plate. A plate of uranium glass was mounted in such a position that the two reflected images were focussed on it. The image formed by reflection from the outer surface was noticeably brighter than the other, owing to absorption by the fused quartz plate, which was twice traversed by the rays reflected from its inner surface. On heating the bulb to a red heat with a Bunsen flame, the latter image brightened up until it appeared to be about three times as bright as the image reflected from the outer surface. In this way it is possible to demonstrate the selective reflection of the vapor to a small audience at close range. A sheet of heavy plate glass must be used as a protection against a possible explosion of the bulb as the pressure may rise to 15 or 20 atmospheres.

By means of photographs made with different exposure times it was found that the reflecting power of the quartz was increased about fivefold by the mercury vapor, *i.e.* raised to a value of about 25%, nearly that of most metals in this region of the spectrum.

The next point to determine was the density at which selective or metallic reflection commenced. Various methods were tried, the following being the one finally adopted.

By employing polarized light and adjusting the angle of incidence on the quartz window, the two reflected beams were reduced nearly to zero, and a small increment in reflecting power due to mercury vapor became more conspicuous. The two reflected beams were received by the lens of the quartz camera and exposures made with the bulb at different temperatures. The image due to the internal reflection was distinctly brighter than the other when the mercury vapor had a pressure of 12 cms. and at 25 cms. was enormously brighter.

In subsequent experiments made by the author the source of light was a mercury resonance lamp, excited by a water-cooled magnetically deflected mercury arc (as described in the Chapter on Resonance Radiation) and rough measurements were made of the reflecting power of the vapor at 1 atmosphere pressure. A value of 0.25 was obtained.

The effects of an admixture of hydrogen with the mercury vapor were studied by O. Schnettler.[1] Hydrogen at 0.2 mm. reduces the intensity of resonance radiation to one-half, while hydrogen at 50 cms. was almost without influence on the intensity of the reflected radiation, the latter being due to forced vibrations which are coherent with the light, while the resonance radiation is in-

[1] *Zeit. für Phys.*, *66*, 55, 1930.

coherent and due to Bohr transitions, the atoms having a definite lifetime in the excited state. He assumes two kinds of radiation, fluorescent (*i.e.* resonance radiation) and scattered, and considers that the probability of the latter is increased by a coupling of the mercury atoms at higher pressures, but that there is no coupling with a foreign gas.

With 50 cms. of hydrogen in his prismatic reflecting bulb he found the following reductions of reflecting power:

Temp.	Ref. Reduced to
371	73%
415	78
481	84

The bulb was in communication with a pump and a barometer column, so that he could exhaust or introduce a gas. The barometer column was frozen with liquid air in lieu of a stopcock, so that high mercury vapor pressure could be developed in the bulb after the admission of the hydrogen.

The subject was further investigated by W. Rump [1] who found evidences of selective reflection at lower pressures than 10 cms., using also a resonance lamp as the source, and receiving the reflected ray with a photo-cell. A filter of mercury vapor was introduced alternately between the lamp and the reflecting bulb, and between the bulb and thermo-cell, larger deflections being obtained in the latter case when only resonance radiation was being emitted. This experiment should be repeated; in the opinion of the author there is no reason for expecting a difference depending on the position of the absorption cell.

[1] *Zeit. für Phys.*, *29*, 196, 1924.

CHAPTER XVI

THE OPTICAL PROPERTIES OF METALS

Of the various optical properties of metals, the most interesting is the relation between the reflecting power and the electrical conductivity. It can be shown theoretically that the proportion of the incident radiation which penetrates the metal surface and is absorbed, divided by the square root of the electrical resistance, is very nearly a constant for all metals. This relation is confirmed by experiment if observations are confined to wave-lengths greater than 20 μ, though it does not hold in the visible and near infra-red, for reasons which we shall see. Rubens and Hagen found platinum films less transparent than gold and silver in the visible, but more transparent in the infra-red region. A bismuth film .09 μ thick transmitted less than 0.01% in the red, while at 4 μ it transmitted 10%. From this they concluded that the influence of molecular free periods (as they termed them) gradually disappeared with increasing λ.

If polarized light is reflected from a metal at other than normal incidence, the components of the incident electric vector are reflected with a phase-difference, giving rise to elliptical polarization, from the study of which the optical constants of the metal n and κ can be determined.

The optical properties of metals are explained by the presence of both free and bound electrons, the latter having natural free periods as in the case of transparent and semi-transparent media. For the low frequencies, corresponding to the infra-red, the free electrons come into play, treating waves of all lengths very much alike, and giving rise to powerful reflection and absorption. Silver for example reflects over 98% at 1.5 μ, 95% in the red, 85% in the extreme violet, 14% at 3260 and only 4% at 3160. That portion of the energy which is not reflected, enters the surface and is completely absorbed after traversing a distance of a few wave-lengths. Metals vary, however, in respect to the ratio of reflecting to absorbing power, and for any given metal the ratio varies with the wave-length. For example for silver at $\lambda = 2880$ the absorption constant $a = 19$, the reflection in per cent $R = 34$, while for $\lambda = 3750$ $a = 19$ while $R = 74$, which accounts for the failure of the silvered plates of the Fabry and Perot interferometer to give a large number of multiple reflections in the ultra-violet.

Apparently the free electrons give high reflecting power with only a moderate absorption coefficient, this relation holding for the longer wave-lengths, while the bound electrons, operating in the short wave-length region, give only a very moderate reflecting power for the same absorption. The narrow transparent region of silver can be regarded as a gap between the regions of the two types of absorption. Homogeneous films of the sodium and potassium metals, deposited on the wall of a quartz bulb at $-183°$ were found by the author [1] to be highly transparent to the ultra-violet between 2300 and 3600 while quite opaque to visible light. Here we probably have a gap between the free and bound electron absorption bands similar to, but much wider than, the one shown by silver.

The subject was taken up again in 1932, all of the alkali metals being studied in relation to their reflecting and transmitting power. The very remarkable fact was discovered that the optical properties of these metals change from those of a metal to those of a dielectric as we decrease the wave-length of the light. In the ultra-violet they give plane-polarization by reflection, and the dispersion can be determined by observation of the angle.

The equations which have been developed in the previous chapter can, with certain modifications, be applied to metals, in which case the optical properties depend, at least for long wave-lengths, on the presence of the free or conducting electrons, and can therefore be expressed in terms of the electrical conductivity of the metal. For visible and ultra-violet light metals in general show more or less selective properties due to the presence of bound electrons having natural free-periods. These give rise to selective absorption, which is very marked in the case of gold and copper. Glass plates can be easily gilded either by cathodic sputtering or more rapidly by evaporation from a gold plated tungsten filament in a high vacuum. The films are very transparent for green light and show strong selective reflection for the red and yellow region. If two plates are opposed and the light of a tungsten lamp multiply reflected between them, the image of the filament appears as red as when viewed through ruby glass. This enhancement of the surface color is analogous to that obtained by the residual ray method of Rubens. The same phenomenon can be seen by looking into the interior of a gilded goblet, the bottom appearing of a deep red color. Silver has the very property of reflecting powerfully the entire visible and near ultra-violet regions, with the exception of a narrow gap at wave-length 3200, where the reflecting power is only 4%, about that of a single glass surface. A silver film of such thickness that the filament of a nitrogen-filled tungsten lamp is barely

[1] *Phil. Mag.*, *38*, 98, 1913.

visible through it, transmits this narrow region of the spectrum with sufficient intensity for photographic purposes. It was with a quartz lens heavily silvered that the first ultra-violet photographs of the moon, terrestrial landscapes, flowers and other objects were made by the author [1] in 1908. As has been said the alkali metals are still more remarkable in this respect.

According to our present views, we regard metals as substances in which free electrons exist, which are capable of continuous movement under the action of a steady electromotive force. Heretofore we have regarded the electrons as bound to positions of equilibrium by forces of restitution, experiencing only a slight change of position under the action of a steady electric force. Upon the removal of the force, the electron returns to its original position.

We will now investigate the behavior of free electrons under the action of the rapidly alternating electrical forces of light-waves.

Electromagnetic Theory of Metallic Absorption. — We have already discussed the propagation of waves in a medium which is a perfect insulator, in which the current is proportional to $\partial X/\partial t$. In such a medium the current may consist of two parts, a displacement current in the ether represented by $1/4\pi(\partial X/\partial t)$, and a convection current due to the motions of the electrons inside the atoms. It is clear that the displacement current will cease as soon as the electric force ceases to vary. We will now investigate the propagation of waves in a medium which is not a perfect insulator. In such a medium a current will be set up under the influence of a steady electric force, which will be proportional to the force X instead of to the rate of change of X. We may think of this current as due to the motion of free electrons, which will drift along under the influence of the force, giving rise to a conduction current represented by σX, in which σ represents the absolute conductivity measured electrostatically.

If we have a periodic electric force, as in light-waves, both currents will be present, and we may have absorption or a transformation of energy from two distinct causes. If the vibration of the electrons which are not free is accompanied with something akin to friction, there will be a heating similar to the heating of the dielectric of a condenser when it is rapidly charged and discharged. This type of absorption has been discussed in the Chapter on Dispersion. There may in addition be an ohmic heating, similar to the heating of wires by steady currents. This we may think of as due perhaps to the impacts of the free electrons with atoms.

[1] *Roy. Inst. Proc.*, 1911.

If we limit ourselves to plane-polarized plane-waves, we may write for the total current parallel to the x axis

$$j_x=\frac{\epsilon}{4\pi}\frac{\partial X}{\partial t}+\sigma X. \tag{1}$$

No further modifications than the one introduced into equation (1) are needed, and the fundamental Maxwell equations

$$\frac{4\pi j_x}{c}=\frac{\partial \gamma}{\partial z}-\frac{\partial \beta}{\partial y}, \text{ etc., and } \frac{1}{c}\frac{\partial \alpha}{\partial t}=\frac{\partial Y}{\partial z}-\frac{\partial Z}{\partial y}, \text{ etc.,}$$

still hold, if we write the permeability $\mu=1$.

This we are justified in doing, even in the case of the strongly magnetic metals such as iron, nickel and cobalt, for experiments indicate that the ferromagnetic effect is unable to follow the very rapid changes involved in the case of light-waves.

The boundary conditions may be written as before, $X_1=X_2$, $Y_1=Y_2$, $\alpha_1=\alpha_2$, $\beta_1=\beta_2$.

For the present we shall concern ourselves only with the absorption due to ohmic heating, *i.e.* resulting from the term σX, the conduction current, for the case of a plane-wave polarized with the electric vector parallel to x and propagated in the z direction.

Substituting from (7) to (5) in previous chapter

$$\frac{\partial \beta}{\partial t}=\frac{\partial Z}{\partial x}-\frac{\partial x}{\partial z}=-\frac{\partial x}{\partial z} \qquad \text{since } Z=0.$$

Differentiating (4) $\quad \frac{4\pi}{c}\frac{\partial j_x}{\partial t}=\frac{\partial}{\partial y}\frac{\partial \gamma}{\partial t}-\frac{\partial}{\partial z}\frac{\partial \beta}{\partial t}=\frac{-\partial}{\partial z}\frac{\partial \beta}{\partial t}$

since $\gamma=0$ $\quad \therefore \frac{4\pi}{c}\frac{\partial j_x}{\partial t}=-\frac{\partial}{\partial z}\left(-c\frac{\partial X}{\partial z}\right)=c\frac{\partial^2 X}{\partial z^2}$

$$\frac{\partial j_x}{\partial t}=\frac{c^2}{4\pi}\frac{\partial^2 X}{\partial z^2}$$

if $\quad j_x=\frac{\epsilon}{4\pi}\frac{\partial X}{\partial t}$, we have $\frac{\partial^2 X}{\partial t^2}=\frac{c^2}{\epsilon}\frac{\partial^2 X}{\partial z^2}$. (eq. 12)

If $\quad j_x=\sigma X \quad \frac{\partial j_x}{\partial t}=\frac{\sigma \partial X}{\partial t}$ and $\frac{\sigma \partial X}{\partial t}=\frac{c^2}{4\pi}\frac{\partial^2 X}{\partial t^2}$.

If finally $\quad j_x=\sigma X+\frac{\epsilon}{4\pi}\frac{\partial X}{\partial t}$ we have

$$\frac{\epsilon}{c^2}\frac{\partial^2 X}{\partial t^2}+\frac{4\pi\sigma}{c^2}\frac{\partial X}{\partial t}=\frac{\partial^2 X}{\partial z^2}. \tag{2}$$

When applied to harmonic motion, this equation has for its solution

$$X = Ae^{i\frac{2\pi}{T}(t-mz)}, \text{ in which } m \text{ is complex.} \tag{3}$$

Differentiating (3), and substituting in (2), we get

$$\frac{\epsilon}{c^2}\left(-\frac{4\pi^2}{T^2}\right)+\frac{4\pi\sigma}{c^2}\frac{2\pi i}{T} = -\frac{4\pi^2}{T^2}m^2$$

$$\text{or,} \qquad \frac{\epsilon'}{c^2} = m^2 \tag{4}$$

in which the complex dielectric constant $\epsilon' = \epsilon - 2i\sigma T$ and since ϵ is complex, m is also complex.

Now m has the dimension of a reciprocal velocity, and we may write $m = (1-i\kappa)/V$, in which V is the velocity of propagation of the wave in the absorbing medium.

Substituting this value in (2),

$$X = Ae^{\frac{2\pi i}{T}\left(t-\frac{(1-i\kappa)}{V}z\right)} = Ae^{\frac{2\pi i}{T}\left(t-\frac{z-i\kappa z}{V}\right)} = Ae^{\frac{2\pi it}{T}-\frac{2\pi iz}{TV}-\frac{2\pi\kappa z}{TV}},$$

in which $TV = \lambda$.

$$X = Ae^{-2\pi\kappa\frac{z}{\lambda}}e^{2\pi i\left(\frac{t}{T}-\frac{z}{\lambda}\right)}.$$

In this expression $Ae^{-2\pi\kappa\frac{z}{\lambda}}$ represents the amplitude, which clearly decreases as z increases.

After traversing a thickness equal to the wave-length λ, the amplitude has decreased by the amount $e^{-2\pi\kappa}$. The constant κ is the measure of the absorption, and is called the absorption index.

If we call $c/V = n$ the refractive index of the medium, we have from equation (4)

$$\frac{\epsilon'}{c^2} = \left(\frac{1-i\kappa}{V}\right)^2,$$

$$\epsilon' = \frac{c^2}{V^2}(1-\kappa^2-2i\kappa);$$

$$\therefore \epsilon' = n^2(1-\kappa^2-2i\kappa); \text{ or } \sqrt{\epsilon'} = n(1-i\kappa) \tag{5}$$

and since $\epsilon' = \epsilon - i2\sigma T = n^2 - n^2\kappa^2 - 2n^2 i\kappa,$

we get by equating the real and imaginary parts,

$$n^2(1-\kappa^2) = \epsilon, \; n^2\kappa = \sigma T. \tag{6}$$

This last relation is in agreement with facts only for long waves, as we shall see presently. The reason of this is that, in the present treatment, we have not taken into account the influence which the bound electrons have upon the propagation of the disturbance. We have shown that, in the case of imperfect insulators, we have a complex dielectric constant, due to the conduction term σX. In the treatment of dispersion we have seen that a complex dielectric constant results from the presence of vibrating dipoles, without the presence of the term σX, and absorption will occur if the vibration of the dipole is damped by radiation or otherwise. There are thus two distinct types of absorption, one caused by conducting, the other by bound electrons.

Metallic Reflection. — We will now deduce the formula which expresses the reflecting power of a metal for normal incidence, in terms of its electrostatically measured conductivity σ and the periodic time T of the vibration, following Drude and Planck.

For transparent media we found for the ratio of the reflected to the incident amplitude $(n-1)/(n+1)$. In the present case we write for n the square root of the complex dielectric constant $\tilde{\epsilon}$, given by (5)

$$\frac{\tilde{R}_p}{E_p}=\frac{R_p e^{id_p}}{E_p}=\frac{n(1-i\kappa)-1}{n(1-i\kappa)+1}.$$

Multiplying this by its complex conjugate $(n+ni\kappa-1)/(n+ni\kappa+1)$ gives for the reflecting power R measured in per cent of incident *intensity*, since multiplying by the complex conjugate is equivalent to the squaring necessary to change amplitude to intensity.

$$R=\frac{R_p^2}{E_p^2}=\frac{n^2(1+\kappa^2)+1-2n}{n^2(1+\kappa^2)+1+2n}=\frac{(n-1)^2+n^2\kappa^2}{(n+1)^2+n^2\kappa^2} \tag{7}$$

which expression differs from that for transparent media only by the presence of $n^2\kappa^2$ in the numerator and denominator. From the first equation of (6) it follows that the maximum value of κ is 1 otherwise we obtain a negative dielectric constant which has no meaning.

If we put $\kappa=1$ it follows from (6) that

$$n=\sqrt{\sigma T}. \tag{8}$$

As we shall see presently the measured values of κ for visible light for metals are invariably greater than unity. This may be due to the presence of bound electrons, which we are neglecting in the present treatment.

Changing the sign of $2n$ in the numerator of (7), which requires the subtraction of $4n$ and dividing, gives

$$R=1-\frac{4n}{n^2+2n+1+n^2\kappa^2}=1-\frac{4n}{(n+1)^2+n^2\kappa^2}.$$

If $\kappa=1$, $n>1$

$$(n+1)^2+n^2\kappa^2=2n^2.$$

$$\therefore \quad R=1-\frac{4n}{2n^2}=1-\frac{2}{\sqrt{\sigma T}} \qquad \text{from (8).}$$

Introducing, instead of σ and T, the specific resistance r, in chms, of a metre wire of 1 mm.^{2} cross section, and λ measured in μ, we have, since $\sigma T=30\lambda/r$, an approximation formula for the amount of radiation which penetrates the surface of the metal

$$1-R=0.365\sqrt{\frac{r}{\lambda}},$$

a formula arrived at by Rubens and Hagen from measurements of the reflecting power. The formula shows that the reflecting power increases with the conductivity. As written above, the development based on the assumption that $\kappa=1$ holds for wave-lengths greater than 12 μ. Between 5 μ and 12 μ a formula of two members is required

$$1-R=.365\sqrt{\frac{r}{\lambda}}-.0667\frac{r}{\lambda}. \qquad (8a)$$

Below 5 μ the reflection can no longer be calculated from the conductivity. The single term formula can be written $(1-R)/\sqrt{r}=.365/\sqrt{\lambda}$ which shows that for a given wave-length, the amount of radiation penetrating the metal divided by the square root of the resistance is a constant.

The following table by Rubens and Hagen shows the values determined by experiment.[1]

$\frac{100-R}{\sqrt{r}}$ for	4 μ	8 μ	12 μ	25 μ	
Silver	14.9	9.8	9	7.3	
Copper	20.6	10.6	12.1		
Gold	21.9	17.4	13.8	7.1	
Nickel	23.9	13.6	12	7.3	
Steel	27.3	15.7	11	6.9	
Bismuth	22.7	16.9	16.3	8.35	(molten)
Calculated from (8a)	18.25	12.9	10.5	7.29	

[1] *Ann. der Phys.*, *8*, 1; *11*, 874, 1903.

In the older tables the abnormally high value for bismuth was due to imperfect polish and oxidation. A consistent result was obtained with a clean molten surface.

The reflecting powers were measured by means of the apparatus shown in Fig. 325. A Nernst lamp B was fixed on a turn-table in such a position that its image A, formed by the concave mirror D, was symmetrically located with respect to the centre of the turn-table. By turning the table the filament of the lamp could be brought into the position previously occupied by its image. The surface of the mirror D was composed of the metal under investigation. The rays from the filament or its image were focussed upon the slit of a reflecting spectrometer, furnished with a fluorite prism, and the spectrum thrown upon the thermopile at T. It is clear that by this arrangement we have a means of comparing the incident with the reflected energy, since the loss by reflections from the surfaces of the optical parts of the instrument is the same in each case. They investigated a large number of metals and alloys of known conductivity, obtaining curves of the type shown in Fig. 326.

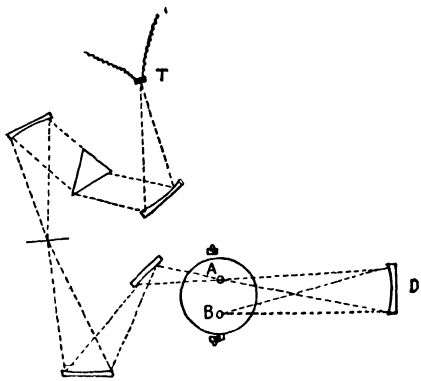


FIG. 325

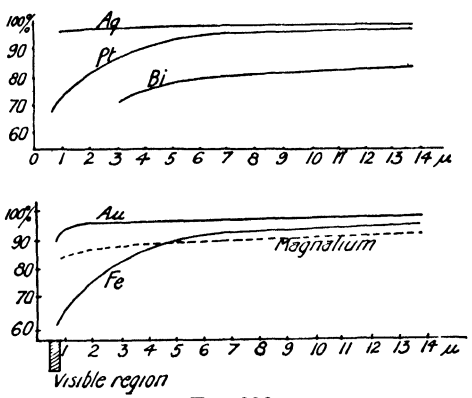


FIG. 326

It is worthy of note that magnalium, the most brilliantly reflecting alloy which we have, is surpassed by iron, in reflecting power, for wave-lengths greater than 4 μ.

The reflecting powers of various metals are given in the table on page 550 by Rubens and Hagen.

It is apparent from the table that the reflecting power of silver in contact with glass is somewhat less than that of silver in contact with air.

The same is true for mercury.

An easy way of exhibiting the loss of light by reflection from a metal is to half fill a test-tube with mercury and plunge it in a jar of clean water; the light reflected from the metal will appear quite dull in comparison with the light totally reflected at the glass-air surface.

For $\lambda =$	251	288	305	316	326	338	357	385	420	450	500	700	1000
Silver	34.1	21.2	9.1	4.2	14.6	55.5	74.5	81.4	86.6	90.5	91.3	94.6	96.6
Platinum	33.8	38.8	39.8	—	41.4	—	43.4	45.4	51.8	54.7	58.4	69	75.5
Nickel	38	43	44	—	45	46	49	49	56	59	61	69	73
Steel	33	35	37	—	40	—	45	48	52	54	55	58	62
Gold	39	34	32	—	28	—	28	27	29	33	47	92	97
Copper	26	24	25	—	25	—	27	28	33	37	44	83	93
Brashear alloy — 68 Cu, 32 Sn	30	38	42	—	—	—	51	53	56	60	63	67	74
Schröder alloy — 66 Cu, 22 Sn, 12 Zn	40	48	50	—	54	—	57	60	62	63	63	67	75
Magnalium — 69 Al, 31 Mg	67	70	72	—	75	—	81	84	83	83	83	83	84
Glass backed by silver										79 to 85	81 to 86	84 to 89	
Mercury			70	(O'Brien)			78		77		75		
Glass backed by mercury										73	71	73	

Determination of R from Emission. — On account of the relation between emission and absorption, expressed by Kirchhoff's law, we can determine $(100-R)$ by comparing the emission from the substance of waves of given length, with the corresponding emission of a perfect black body. Twenty different metals were investigated by this method by Rubens and Hagen. The radiations from the black body (see Chapter on Laws of Radiation) and from the metal, both heated to the same temperature, were allowed to fall alternately upon a thermopile, after reflection from four fluorite surfaces, which eliminated all wave-lengths except 25.5 μ. The temperature employed was 170°, maintained by an electrically heated bath of aniline, and the conductivity of the metal used in the calculations was of course the conductivity for this temperature. The results appear in the table, on page 551 in which x is the conductivity.

It is at once clear that a further increase in the length of the wave has improved matters, as is shown by the more nearly constant value of the products in the last column. The mean value is 7.34, while that calculated is $c_{25}\mu = 36.5/\sqrt{25.5} = 7.23$.

These results show that it is possible to determine the specific electrical conductivity of the metals by purely optical means, if we limit our observations to waves of sufficient length.

TEMPERATURE 170°

	$\sqrt{x}$	EMISSION $J=100-R$	$(100-R)\sqrt{x}$
Silver	6.26	1.13	7.07
Copper	5.70	1.17	6.67
Gold	5.21	1.56	8.10
Aluminum	4.52	1.97	8.91
Zinc	3.19	2.27	7.24
Cadmium	2.86	2.55	7.29
Platinum	2.44	2.82	6.88
Nickel	2.29	3.20	7.33
Tin	2.24	3.27	7.32
Palladium	2.11	3.58	7.53
Steel	1.81	3.66	6.62
Mercury	.957	7.66	7.33
Bismuth	.716	25.6	18.3

Determination of n and κ. — It has been shown that it is possible to develop equations expressing n and κ in terms of the conductivity of the metal.

We will now consider the method by which they can be determined optically, and a comparison of the values determined by the two methods will furnish a test of the theory. To determine the refractive index n and the absorption index κ we have only to measure two especial angles. If plane-polarized light in azimuth 45° is incident on a metallic mirror at other than normal incidence, the components of the electric vector parallel and perpendicular to the plane of incidence are reflected with a phase-difference, and the reflected beam is elliptically polarized. The ellipticity will be greatest for the angle of incidence for which the phase-difference $\Delta=\pi/2$. This angle is called the ANGLE OF PRINCIPAL INCIDENCE, and circularly polarized light, reflected at this angle, will be converted in plane-polarized light. This angle is easily determined by reflecting circular vibrations obtained with a quarter-wave plate or Fresnel rhomb from the metallic surface and finding the angle at which the reflected beam can be quenched by a Nicol.

The PRINCIPAL AZIMUTH is the angle which the direction of the resulting plane vibration (short diagonal of the Nicol) makes with the plane of incidence. Equations can be developed in which n and κ are expressed in terms of these two angles.

We have for the refractive index

$$n=\frac{\sin\Phi}{\sin\chi} \text{ or } \sin\chi=\frac{\sin\Phi}{\sqrt{\epsilon}}, \text{ since } n=\sqrt{\epsilon}\ (\epsilon=\text{dielectric constant}).$$

As we have seen the dielectric constant is complex in the case of metals, we therefore write for the present case

$$\sin\chi=\frac{\sin\Phi}{\sqrt{\epsilon'}} \tag{9}$$

in which $\sin\chi$ is complex.

Let light polarized at an angle of 45° with the incidence plane be reflected from a metal.

In the previous chapter we found for the ratio of the reflected components

$$\frac{R_p}{R_s}=-\frac{E_p\cos(\Phi+\chi)}{E_s\cos(\Phi-\chi)}=-\frac{\cos(\Phi+\chi)}{\cos(\Phi-\chi)}$$

of $E_p=E_s$ as is the case for azimuth 45°. The amplitudes of the reflected components are complex in the present case, and we therefore have

$$\frac{R_p'}{R_s'}=\rho e^{i\Delta}=-\frac{\cos(\Phi+\chi)}{\cos(\Phi-\chi)} \tag{10}$$

in which ρ=the ratio of the real amplitudes of the p and s components of the reflected light and Δ their *relative* phase-difference, as one sees if one puts

$$R_p'=R_pe^{i\delta_p};\ R_s'=R_se^{i\delta_s} \text{ in which } R_p,\ R_s,\ \delta_p \text{ and } \delta_s \text{ are real.}$$

$$\rho=\frac{R_p}{R_s} \text{ and } \Delta=\delta_p-\delta_s.$$

Since the right-hand member of the above equation is complex, Δ must differ from zero, and there is a phase-difference between the two components of the reflected light, which produces elliptical polarization.

We will now determine how this phase-difference and the accompanying elliptical polarization vary with the angle of incidence. We must first get rid of χ in equation (10) since we have no means of measuring the angle of refraction in the case of a metal. We then develop an equation expressing n and κ in terms of the angle of principal incidence.

Writing (10) in the form

$$\rho e^{i\Delta}=-\frac{\cos\Phi\cos\chi-\sin\Phi\sin\chi}{\cos\Phi\cos\chi+\sin\Phi\sin\chi};$$

and multiplying this equation by the denominator and transposing the terms gives us

$$\frac{1+\rho e^{i\Delta}}{1-\rho e^{i\Delta}}=\frac{\sin\Phi\sin\chi}{\cos\Phi\cos\chi}=\tan\Phi\frac{\sin\chi}{\cos\chi}=\frac{\tan\Phi\sin\Phi}{\sqrt{\epsilon'\left(1-\frac{\sin^2\Phi}{\epsilon'}\right)}};$$ by (9)

$$\therefore\frac{1+\rho e^{i\Delta}}{1-\rho e^{i\Delta}}=\frac{\sin\Phi\tan\Phi}{\sqrt{\epsilon'-\sin^2\Phi}}. \tag{10}$$

At normal incidence $\Phi=0$, $\rho e^{i\Delta}=-1$, *i.e.* $\Delta=0$ and $\rho=-1$, or the wave is reflected with a change of sign, but with no phase-difference between the components; the light therefore remains plane-polarized. The reflected waves form, by interference with the incident waves, a system of stationary waves, and since the reflection is accompanied by a change of sign, we shall have a node at the reflecting surface.

At grazing incidence $\Phi=90$, $\rho e^{i\Delta}=1$, *i.e.* $\Delta=0$ and $\rho=1$, or reflection occurs without change of sign and without elliptical polarization.

The ellipticity will be greatest for the angle of incidence for which $\Delta=\pi/2$. At this angle we have $e^{i\Delta}=i$, since $e^{i\Delta}=\cos\Delta+i\sin\Delta=0+i$.

This angle is termed the angle of principal incidence, and we will designate it by $\overline{\Phi}$.

If in (10) $\sin^2\Phi$ is disregarded in comparison to ϵ' we have from (5)

$$\frac{1+\rho' i}{1-\rho' i}=\frac{\sin\overline{\Phi}\tan\overline{\Phi}}{n(1-i\kappa)}. \tag{11}$$

Multiplying by the complex conjugate gives

$$1=\frac{(\sin\overline{\Phi}\tan\overline{\Phi})^2}{n^2(1+\kappa^2)}$$

$$\text{or } \sin\overline{\Phi}\tan\overline{\Phi}=n\sqrt{1+\kappa^2}. \tag{12}$$

This same equation can be obtained without disregarding $\sin^2\Phi$ in (10) by a somewhat longer process (see Drude's *Optics*).

If by means of a Babinet compensator, which annuls the phase-difference introduced by the metallic reflection at any angle of incidence, we convert the elliptical vibration into a plane-polarized one, the plane of polarization will make an angle Ψ with the plane of incidence, and we have the relation $\rho=\tan\Psi$ since ρ is the ratic of the p and s components of the reflected light.

If we can establish equations connecting n and κ with Φ, Δ and

Ψ, it is clear that we can determine the refractive index n and the absorption index κ of a metal by observations made on the state of polarization of the reflected light. Such methods are called katoptric methods in contrast to the dioptric methods employed in the case of transparent substances. They have been largely used in the determination of the optical properties of intensely opaque matter. They are obviously not as reliable as the dioptric methods, for, as we have seen, the ellipticity of the reflected light is profoundly affected by the presence of surface films, and we can seldom be certain that such films are not present in the case of metallic surfaces, since any polishing process may produce them. A freshly cleaved surface of a single metallic crystal would, however, be free from them.

We will now establish relations between n and κ, and Φ, Δ and Ψ.

We require first an expression for $1-pe^{i\Delta}/1+pe^{i\Delta}$, in terms of Ψ and Δ, which expression we shall substitute in equation (10).

Since $\rho=\tan\Psi=\sin\Psi/\cos\Psi$, we may write the above expression,

$$\frac{1-\tan\Psi(\cos\Delta+\mathrm{i}\sin\Delta)}{1+\tan\Psi(\cos\Delta+\mathrm{i}\sin\Delta)}=\frac{\cos\Psi-\sin\Psi(\cos\Delta,\text{ etc.})}{\cos\Psi+\sin\Psi(\cos\Delta,\text{ etc.})}.$$

A short transformation leads to

$$\frac{1-\rho e^{i\Delta}}{1+\rho e^{i\Delta}}=\frac{\cos 2\Psi-\mathrm{i}\sin 2\Psi\sin\Delta}{1+\sin 2\Psi\cos\Delta}=\frac{n(1-i\kappa)}{\sin\Phi\tan\Phi},$$ [as in (11) neglecting $\sin^2\Phi$]. $\therefore$ equating the real and imaginary parts,

$$\kappa=\sin\Delta\tan 2\Psi,\ n=\sin\Phi\tan\Phi\frac{\cos 2\Psi}{1+\cos\Delta\sin 2\Psi},\qquad(13)$$

$$n^2(1+\kappa^2)=\sin^2\Phi\tan^2\Phi\frac{1-\cos\Delta\sin 2\Psi}{1+\cos\Delta\sin 2\Psi}.\qquad(14)$$

If the light is incident at the angle of principal incidence $\overline{\Phi}$, the corresponding angle $\overline{\Psi}$ is called the principal azimuth. In this case

[since $\sin\Delta=1$ in (13)] $\kappa=\tan 2\overline{\Psi}$ (15)

and by (12) $n\sqrt{1+\kappa^2}=\sin\overline{\Phi}\tan\overline{\Phi}$ [since $e=\cos\Delta=0$ in (14)].

$$\therefore n=\sin\overline{\Phi}\tan\overline{\Phi}\cos 2\overline{\Psi}.\qquad(16)$$

Inhomogeneous Waves and Rigorous Formulae. — The formulae just given for n and κ were derived by neglecting $\sin^2\Phi$. For $\Phi=68°$ an error of 7.5% is introduced but for the larger angles usually found (71 to 80) it is less. The neglect of the $\sin^2\Phi$ has

left out of account a curious phenomenon first treated by Ketteler in his *Theoretische Optik* (1850) namely, that n and κ *depend upon the angle of incidence.* In strongly absorbing media, such as metals, the planes of equal phase are not coincident with the planes of equal amplitude, the angle between them increasing with the angle of incidence. If we draw an oblique refracted wave-front (which is a plane of equal phase) in an absorbing medium, it is clear that one end of it has penetrated to a greater depth than the other and consequently has experienced a greater reduction of amplitude. The planes of equal amplitude are parallel to the boundary. Ketteler showed that a wave of this type, in which the amplitude decreased rapidly along the wave-front was propagated in a manner quite different from that of a wave of equal amplitude. It would even be possible to have a wave of this type in free space, as by the passage of light through a metal prism of small angle, and the equations show that it moves in the normal direction but with a velocity less than that of an ordinary light-wave. In a metal it travels much as does the extraordinary wave in doubly refracting media, that is it slides off to one side as it advances, the ray-velocity differing from the wave-velocity.

The lower velocity of a non-homogeneous wave can be seen without difficulty in the case of total reflection. As we have seen in the Chapter on Reflection Theory, there is a disturbance in the rarer medium which travels along the boundary, the amplitude falling off very rapidly with increasing distance from the boundary. These waves are also non-homogeneous. If we draw two incident wave-fronts, separated by distance λ (in, say, glass) and the two corresponding refracted fronts separated by the greater distance λ' in air, for angles of incidence, less than, equal to and greater than the critical angle, we shall find that the distance between the wave-fronts in air remain unchanged until the critical angle of total reflection is reached, and the refracted wave-fronts are perpendicular to the surface and separated by a distance equal to the distance between the points at which the incident fronts cut the surface. If the angle of incidence is further increased the distance between these points becomes less, and the distance between the waves flowing along the boundary in air, which now replace the refracted wave becomes less in proportion, which means a lowered velocity. At grazing incidence the wave-length of the boundary waves becomes equal to the wave-length in glass. This is of course a purely geometrical way of regarding the lowered velocity.

The mathematical treatment of waves of this type is given in Schuster's *Optics* and in König's article on metal reflection in

Vol. 20 of the *Handbuch der Physik* (1928), and in two papers by Fry.[1]

The relation between the electric and magnetic vectors is very peculiar, for with a plane-polarized electric vector we have an elliptically polarized magnetic vector, one component of the ellipse being in, and the other normal to, the wave-front, which means that the waves are not strictly transverse.

König introduces his treatment of inhomogeneous waves in the new *Handbuch der Physik* (page 194) as follows:

If a plane-polarized plane-wave has entered an absorbing medium, and is advancing in the z direction with its electric vector parallel to x we have

$$E_x = Ae^{-q''z}e^{i(pt-q'z)}. \quad E_y = 0 \quad E_z = 0$$

in which $q' = 2\pi/\lambda$ and $1/q''$ is the distance travelled by the wave for a reduction of amplitude to $1/e$.

If the medium absorbs heavily the amplitude A will depend upon the coördinates x and y, and it follows from Maxwell's equations that, if H represents the magnetic vector

$$H_x = \frac{\delta H_y}{\delta x} = \frac{\delta H_z}{\delta x} = 0.$$

In this case A is a function of y alone. For the magnetic vector we have

$$H_z = 0 \quad H_y = \frac{c}{\mu p}(q' - iq'')Ae^{-q''z} \cdot e^{i(pt-q'z)}$$

$$H_z = -i\frac{c}{\mu p}A'e^{-q''z} \cdot e^{i(pt-q'z)}\left(\text{in which } A' = \frac{dA}{dy}\right).$$

These equations are identical with those for homogeneous waves in an absorbing medium with the exception that H_z is not $=0$ but has a value depending upon the variation of the amplitude in the plane of equal phase.

The complex character of the amplitudes H_y and H_z shows that both have a phase displacement with respect to E_x, that of H_z being retarded by $\pi/2$, and that of H_y by δ, δ being defined by $\tan \delta = q''/q' = \kappa$ in which $q'' = \kappa 2\pi/\lambda$.

It is to be noticed that the magnetic component H_z is in the direction of propagation of the wave.

He derives as the final equations (page 245) for n and κ in terms

[1] *Jour. Am. Opt. Soc.*, *15*, 137, 1927.

of the angles Φ and Ψ, which are determined experimentally

$$n^2=\tfrac{1}{2}\tan^2\Phi\left[1-2\sin^2\Phi\sin^2 2\Psi+\sqrt{1-\sin^2 2\Phi\sin^2 2\Psi}\right]$$

$$\kappa^2=\frac{\sqrt{1-\sin^2 2\Phi\sin^2 2\Psi}-(1-2\sin^2\Phi\sin^2 2\Psi)}{\sqrt{1-\sin^2 2\Phi\sin^2 2\Psi}+1-2\sin^2\Phi\sin^2 2\Psi}$$

in which Φ and Ψ are the angles of principal incidence and azimuth.

From these two equations an expression for $n^2\kappa$ can be determined, namely,

$$n^2\kappa=\tan^2\Phi\sin^2\Phi\sin 2\Psi\cos 2\Psi,$$

which can also be derived from (16) by squaring both sides and multiplying by (15)

$$\kappa=\tan 2\Psi.$$

This shows that $n^2\kappa$ is the same, whether calculated by the original method (neglecting $\sin^2\Phi$) or by the rigorous one.

Determination of Principal Incidence and Azimuth. — Since reflection at the angle of principal incidence converts circularly polarized into plane-polarized light, we can easily determine the angle by reflecting a circular vibration, obtained by means of a quarter-wave plate or Fresnel rhomb, from the metallic surface, and determining the angle of incidence at which the reflected light can be quenched by a Nicol prism. The principal azimuth is the angle which the direction of the plane vibration (short diagonal of the Nicol) makes with the plane of incidence. Or we may start with plane-polarized light and employ a Babinet compensator (with its wedges set so as to displace the central fringe through a distance corresponding to a phase-difference of a quarter of a period) to analyze the reflected light, observing the angle of incidence at which the central fringe returns to the central position. The principal azimuth is determined by observing the angle through which the second Nicol has to be rotated to make the central fringe black. Colored light, obtained by passing sunlight through colored glasses or solutions, should be employed, as the optical constants are a function of the wave-length.

Determinations of the optical constants in the ultra-violet proved very laborious, involving the taking of many photographs for each determination until Voigt originated a very beautiful method of determining the optical constants for each value of λ with a single photograph. The method, in brief, consisted in crossing a double wedge of R and L quartz cut perpendicular to the axis, with a Babinet compensator. By this combination one obtains a double chess-board pattern of dark spots, instead of straight fringes. The shift of one system with respect to the other is determined by the

state of polarization of the light. A full account of the method will be found in a paper by Minor [1] who employed it in his determinations of optical constants of metals over a wide range of wavelengths. Some of his results will be discussed presently.

The values of these angles and the corresponding values of n, κ and R, as determined by Drude [2] for a number of metals are given in the following table:

Metal	$\overline{\Phi}$	$\overline{\Psi}$	n	κ	R
Bismuth	77° 3′	31° 57′	1.90	1.93	65
Lead	76.42	30.45	2.01	1.73	62
Mercury	79.34	35.43	1.73	2.87	78
Platinum	78.30	32.35	2.06	2.06	70
Gold	72.18	41.39	0.36	7.70	85
Antimony	80.26	29.35	3.04	1.63	70
Silver	75.42	43.35	0.18	20.2	95
Copper	71.35	38.57	0.64	4.08	73
Magnesium	77.57	42.42	0.37	11.8	93
Sodium	71	35	0.05	56	99.7

O'Brien [3] has made determinations of the optical constants of a very carefully cleaned mercury surface obtaining the following values, those for 5892 from a paper by Wheeler.

Optical Constants of Mercury

λ	Φ	Δ	Ψ	n	κ	R
5892	60	150	40.7	1.57	4.1	79.6
4358	74.8	92	38	.88	3.94	77
4047	74.8	90.7	38.9	.79	4.30	78.5
3650	74.8	83.3	39.5	.64	4.64	77.7
3130	74.8	74.7	40.8	.44	5.75	78.9
3022	74.8	69.4	39.4	.55	4.09	69.7

These results are interesting as showing the beginning of a minimum of reflecting power at 3022, analogous to the minimum of silver, and as a typical example of the dispersion of a metal. The angle of principal incidence was held constant and a small variation occurred in the azimuth Ψ while the changes of Δ were very large.

Optical Properties of the Alkali Metals. — The metals which exhibit the most remarkable optical properties are those of the alkali group. As was shown by the author [4] many years ago, thin films of sodium and potassium, deposited at liquid air temperature on the inner wall of an exhausted quartz bulb, were highly

[1] *Ann. der Phys.*, *10*, 581, 1903.
[2] *Wied. Ann.*, *36*, 885; *39*, 481, 1890.
[3] *Phys. Rev.*, *27*, 93, 1926.
[4] *Phil. Mag.*, *38*, 98, 1918.

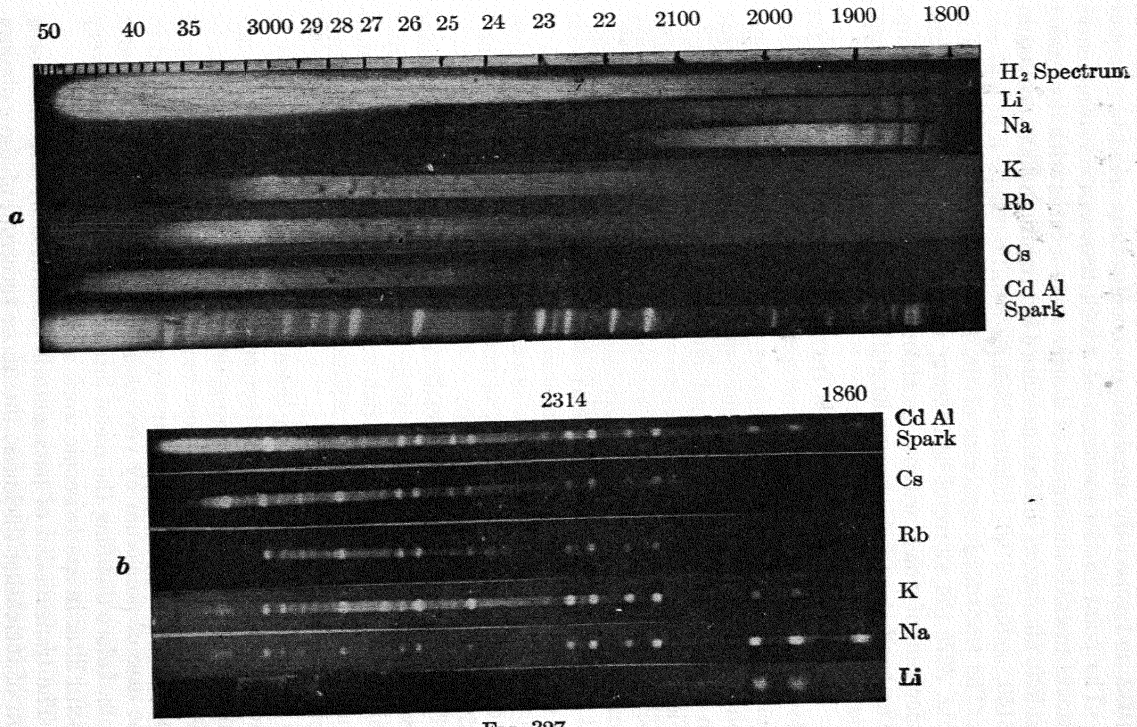

50
40
35
3000
29
28
27
26
25
24
23
22
2100
2000
1900
1800
a
H2 Spectrum
Li
Na
K
Rb
Cs
Cd Al
Spark
2314
1860
b
Cd Al
Spark
Cs
Rb
K
Na
Li

FIG. 327

transparent to the ultra-violet region of the spectrum, while quite opaque to visible light. This investigation was taken up again in 1932, and some very surprising phenomena discovered.[1]

The point in the spectrum at which transparency commences moves towards the region of shorter wave-lengths with decreasing atomic number, as follows: caesium 4400, rubidium 3600, potassium 3150, sodium 2100, lithium 2050.

The transmission spectra of the five metals are reproduced in Fig. 327*a* the source of light being an end-on hydrogen tube, the spectrum of which, without an absorbing film is shown at the top. The bands in the short wave-length region are the oxygen absorption bands of the air.

The transparency continues from the points indicated as far down the spectrum as the investigations have gone at the present time, *i.e.* to about 1860. This is not well brought out in the above photographs on account of under-exposure. It is more apparent in the spectra made with a cadmium aluminium spark shown in Fig. 327*b*. It is clear that we have in this case a very wide gap between the absorptions due to the free and bound electrons. A somewhat similar, but very narrow gap occurs in the case of silver at $\lambda = 3160$, but the transparency in the case of the alkali metals is of a different order of magnitude from that of silver. A film of potassium through which the sun's disk is barely visible (which means a reduction of intensity of perhaps half a million) transmits 25% of everything below 3000.

The case of caesium is especially interesting as its region of high transparency begins in the visible violet and films of the proper thickness transmit light of a rich violet color as deep and pure as that transmitted by a strong solution of cuprammonium or dense cobalt glass.

In the earlier investigation these films were permanent only at low temperatures. At room temperature they broke up into minute globules, the bulb becoming almost completely transparent to visible light. During the more recent investigation it developed that permanency (at least for several days) could be obtained in the case of potassium by very thoroughly out-gassing the quartz bulb, before distilling the metal into it. Such a film makes a very useful filter in spectroscopic work as will appear presently.

The transmission of potassium is illustrated by Fig. 328 for a film of increasing thickness, the upper spectrum of the cadmium spark with 2 seconds exposure made with no film, the following spectra with 8 seconds exposure, a candle flame easily visible through the film in the first case, and the sun's disk absolutely

[1] *Phys. Rev.* 44, 353, 1933.

invisible in the last. It is to be noticed that only a very slight increase in the ultra-violet absorption accompanies this enormous increment in the absorption of visible light. It seemed probable that the high transparency would be accompanied by a very low reflecting power, and this was found to be the case, the reflection of potassium, for example, at wave-length 2147 being only .003 or $\frac{1}{10}$ that of a single surface of fused quartz.

The reflecting powers for different wave-lengths for these metals and the dispersion of K are shown by the curves in Fig. 329.

They were determined by forming a fairly thick film on the inner surface of the quartz bulb, cleaning off the front surface with

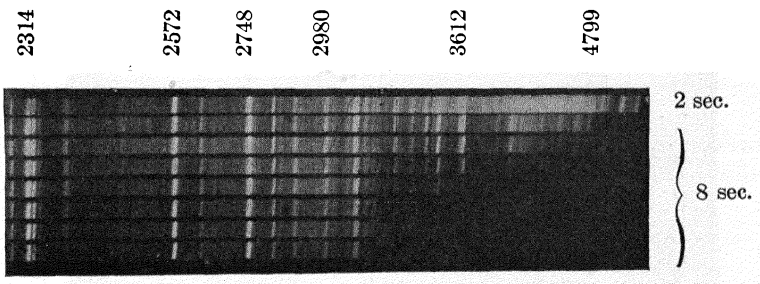


FIG. 328

a small flame, and pointing a quartz spectrograph at the image of a Cd-Al spark reflected in the concave metal mirror. A number of exposures of different times were made, after which the film was removed by the flame and the outer surface of the quartz painted with black Duco. A second set of exposures was now made of the reflection from the single surface of quartz. By comparison of these plates the reflecting power of the metal, in comparison to that of quartz was determined, and the reflecting power of quartz for the different spectral lines having been computed from the table of its refractive indices, the reflecting power of the metal could be at once determined.

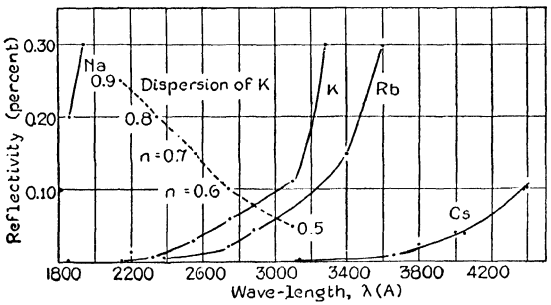


FIG. 329

From the reflecting powers thus determined, the refractive

index of the metal can be determined, as in the case of transparent substances, from the formula $(n-1)^2/(n+1)^2=R$ in which R is expressed in per cent.

More accurate values would result if the complete formula, involving κ were used, but κ has not yet been determined.

Certain anomalies observed at first in the measurements of the reflecting power for different wave-lengths, were suspected to be

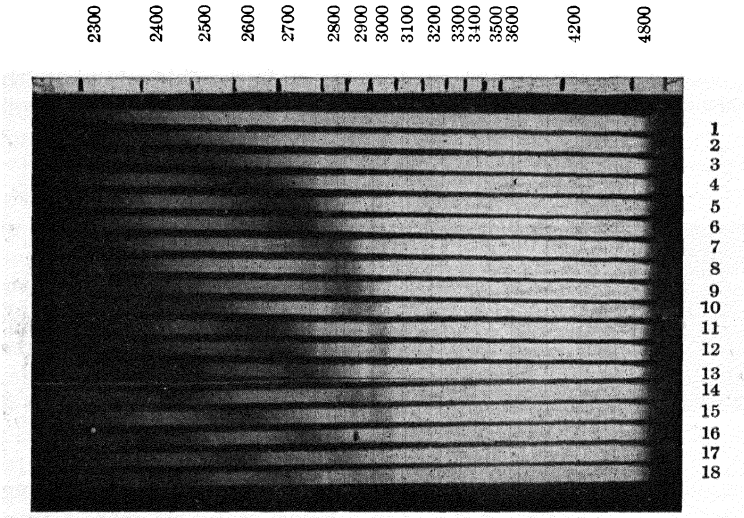


FIG. 330

due to interference between rays reflected from the front and back surface of the metal film, which would be almost certain to occur in the case of a film of such high transparency.

It should then be possible to photograph maxima and minima in the spectrum of the light reflected from the metal film from a continuous source, as in the case of thin films of glass. Substitution of an end-on hydrogen tube, the inner wall silvered to suppress the line spectrum by catalyzing the atomic hydrogen, proved the presence of very distinct maxima and minima, and a series of spectra were then taken in coincidence, the thickness of the film being slightly increased between each exposure. The quartz bulb was backed by a cylindrical cell (a short piece of fibre tube) filled with infusorial earth, which was kept saturated with liquid air.

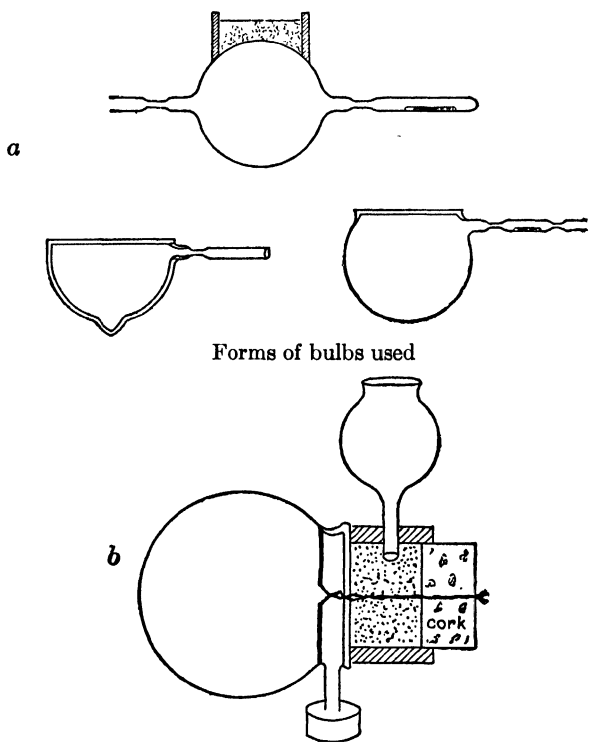
This photograph is reproduced in Fig. 330 which shows in a very graphic manner how the interference minima enter the spectrum at the short-wave-length end and move gradually along the spectrum as the film thickness increases, as described in the Chapter on Interference.

The prediction was now made that these metals would show a Brewsterian angle in the ultra-violet, giving plane polarization by reflection of ordinary light. This was found to be the case, the angle varying with the wave-length, so that it became at once possible to calculate the dispersion from the formula $n = \tan i$, in which i is the polarizing angle.

The values of n for different wave-lengths in the case of potassium are given in the following table, calculated by the formulae $(n-1)^2/(n+1)^2 = R$ and $n = \tan i$.

λ	n FROM R	n FROM TAN i
2147	.89	—
2200	.88	.86
2314	.83	.80
2572	.70	.69
2748	.60	.58
3100	.50	—

For the experiments on polarization a quartz bulb of the form shown in Fig. 331*b* was employed, made by fusing an optically



Forms of bulbs used

FIG. 331

worked circular disk on the end of a tube, inserting a small side tube for exhaustion and introduction of the metal, and then blowing out a thin-walled bulb as shown. (Fig. 331*b*.) This bulb was mounted on a divided circle with the reflecting inner surface

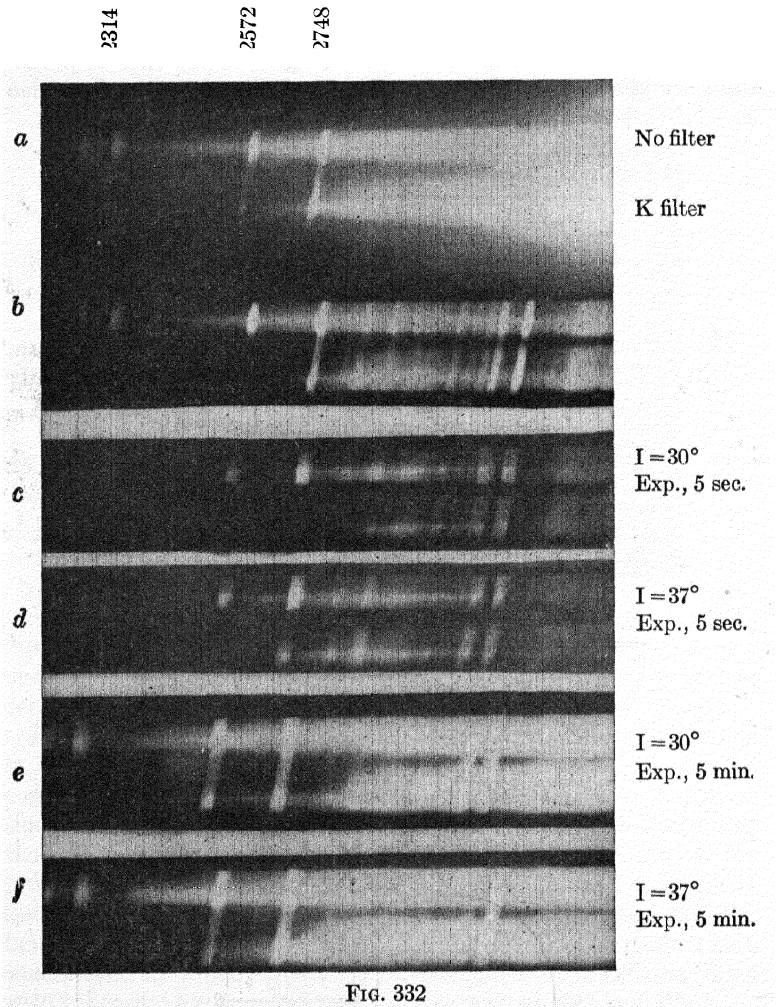


FIG. 332

parallel to and over the axis of rotation. It was backed with a capsule of infusorial earth for liquid air as shown in Fig. 331*b*. The rays from the image of the spark in the metal mirror were passed through a double-image Wollaston prism and focussed on the spectrograph slit by a short-focus fluorite lens. From the rela-

tive intensities of the cadmium lines in the two images the completeness of the polarization could be determined, the ratio being in some cases as 1 : 400.

The Brewsterian angle, in the case of potassium, was found to vary with the wave-length, making it possible to determine the dispersion of the metal in the ultra-violet. Considerable trouble was experienced at first, due to the fog on the plate caused by the powerful reflection of the films in the violet and upper ultra-violet. This was overcome by inserting a second quartz bulb with a fairly thick potassium film in front of the spark as an ultra-violet filter. The advantage gained by the use of this filter is illustrated in Fig. 332, the upper pair of polarized spectra having been made without the filter, the lower pair with the filter inserted. The two polarized spectra have equal intensity above 2748 while the lower spectrum is practically absent below this point. The dispersion is illustrated by figures *c* and *e*, which show that the polarizing angle for 2748 is 30°, the intensity of this line in the lower spectrum

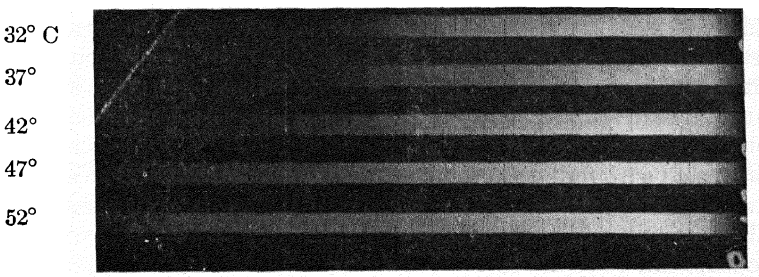


FIG. 333

being about $\frac{1}{40}$ of that in the upper. An exposure of five seconds was given for *c* and five minutes for *e* to bring out the 2314 line, for which the intensity ratio is $\frac{1}{10}$. Figures *d* and *f* show a polarization angle of 37° for 2314, the measured intensity ratio in this case being 1 : 240.

The intensity ratios were measured by giving, say a five-minute exposure to the two spectra, and then making a series of exposures of the brighter spectrum (covering the image of the fainter on the slit) of 1, 2, 4, 8, ··· etc., seconds, picking out the one which matched the intensity of the fainter spectrum made with the five-minute exposure. The results obtained in this way are, of course, not very accurate, but the dispersion calculated from the values of the angle of maximum polarization by the formula $n = \tan I$ agreed surprisingly well with the values obtained from the reflecting power.

Since the refractive index is less than unity for the ultra-violet

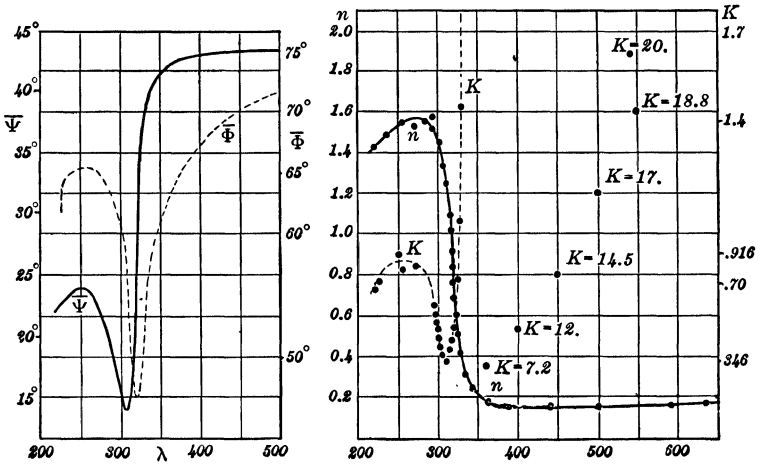
region the metal can be considered as the rarer medium with respect to vacuum, and an angle of total reflection is to be expected, which can be computed from the index. This was found to be the case, the reflecting power, which is of the order of 1% at normal incidence, increasing rapidly with increasing angles of incidence above 40°. The increase is gradual, however, there being no sharply defined angle, as in the case of completely transparent substances. With increasing incidence angle the augmentation moves towards the region of shorter wave-lengths as shown in Fig. 333.

It is necessary to deposit the films at liquid air temperature, for at higher temperatures a colloidal deposit is obtained, the optical properties of which are quite different from those of homogeneous films. The exhaustion is best done with a mercury vapor pump with liquid air trap, the bulb being heated up to the constricted portions over a Bunsen flame or in an electric furnace. After a thorough exhaustion the bulb is allowed to cool, the metal distilled into it and the side tube sealed off. The metal should now be driven back and forth from one side of the bulb to the other with a small flame to remove occluded hydrogen. After sealing off the bulb from the pump a spot is cleaned by a small pointed gas flame and this spot covered by a short piece of fibre tube half filled with infusorial earth (upper part of Fig. 331*a*), or other similar incombustible powder, covered with a disk of filter-paper and then thoroughly saturated with liquid air. The metal is then deposited by local heating of other parts of the bulb with the flame. The thickness of the film is regulated by the area cleaned off. During the operation the capsule can be removed for a few seconds and the opacity observed by viewing a small electric lamp mounted below the bulb.

Optical Constants of Silver and Gold. — The optical constants of silver are of especial interest in connection with its remarkable region of transmission and low reflectivity at wave-length 3200. The constants were determined over a very wide wave-length range by Minor [1] who employed Voigt's method, which will be described presently. The values of $\bar{\Phi}$ and $\bar{\Psi}$, and n and κ are shown by curves in Fig. 334. The κ values for $\lambda > 3500$ are too large for the scale of ordinates and are written in figures at the appropriate λ values. The values of n on the high-frequency side of the band of low reflecting power at 3200 are greater than unity while those on the low-frequency side are less, dropping to $n = 0.18$ in the yellow. It may be of significance that the ordinate $n = 1$ cuts the refraction curve almost exactly at the wave-length value of the minimum reflecting power of 4%. The drop in the value of n from this point on may perhaps be analogous to that discussed

[1] *Ann. der Phys.*, *10*, 614, 1903.

in the case of anomalous dispersion, for we are approaching the region of heaviest absorption from the high-frequency side. In the other direction we are entering the absorption band due



Ψ and Φ for Silver — n and K for Silver

FIG. 334

to bound electrons from the side of lower frequencies and the refractive index increases.

Försterling and Fréedericsz [1] give the following values for infra-red:

λ	n	κ	λ	n	κ	λ	n	κ
65	.19	18.1	1.15	.23	31.2	2.89	1.39	13.7
75	.16	28.6	1.47	.36	23.9	3.38	2.06	11.1
94	.15	36.2	2.10	1.00	14.3	4.04	2.98	9.5
						4.37	4.34	7.4

In the case of gold the ordinate $n=1$ cuts the refraction curve at wave-length 500, while the lowest value of the reflecting power (27%) occurs at 385. In this case we may perhaps consider the two absorption regions as overlapping instead of separated as in the case of silver.

The values of $n\kappa$ can be determined also by measuring the absorption of thin metal films, eliminating the loss by reflection by measuring the intensity of the light transmitted by two films of unequal thickness. If x is the difference of thickness and J_e

[1] *Ann. der Phys.*, *40*, 222. 1913.

the intensity transmitted by the thinner film, we have for J the intensity transmitted by the thicker film

$$J = J_0 e^{\frac{-4\pi n\kappa x}{\lambda}}$$

in which λ is the wave-length in vacuum expressed in the same units as x. The thin film must, however, be of sufficient thickness to give the full reflection of the solid metal, since a metal does not develop its full reflection until a certain thickness is reached, usually of the order of 0.1 μ.

Values of $n\kappa$ obtained by the two methods show good agreement, but the values given in the table cannot be applied to the formulae deduced from the electrical conductivity, as the bound electrons determine in part the values of the optical constants, as measured in the visible region.

Effect of Thickness of Film on Reflection and Transmission. — The most recent work on this subject is by Schultz,[1] who gives both theoretical and experimental curves for metal films not backed by solid matter, and by Murmann[2] who investigated the transmission and reflection of silver films from 5 $m\mu$ to 70 $m\mu$ ($m\mu = 10$ ÅU.) in thickness deposited on glass. His results are shown by Fig. 335 and by the following table, to which has been added the absorption percentages computed by the author from his figures [100 − (Trans. + Ref.) = Abs.]. This column shows that the thinner films absorb more energy than the thick.

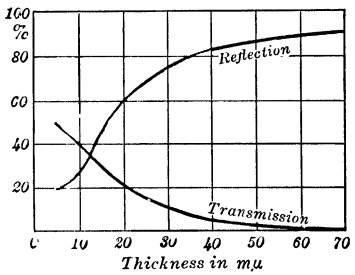


FIG. 335

THICKNESS IN $m\mu$	REF.	TRANS.	ABS.
$m\mu$	%	%	%
5	20	50	30
10	28	40	32
15	44	30	26
20	60	21	19
30	75	10	15
40	83	5	12
60	87	1	12
70	90	0	10

Schultz did not however find this to be the case with his unsupported film, the matter requiring further investigation. Pfund[3] subsequently made measurements which showed 60% absorption for a

[1] *Phys. Zeit.*, *34*, 24, 1933.
[2] *Zeit für Phys.*, *80*, 161.
[3] *J. O. S. A.*, *23*, 272, 1933.

film reflecting 28% in fair agreement with 52% in the computed table. When we come to the subject of thermal emission we shall see that powerful absorption is accompanied with powerful emission and Pfund found that a very thin silver film deposited on quartz radiated more powerfully than a thick one when heated to 150°.

In the case of a thin film there is not sufficient depth for the establishment of the currents in the metal necessary to give the full reflecting power of a massive plate: a greater amount of energy enters the metal, and consequently the absorbing mechanism operates on a larger amount of radiant energy. The reduced reflecting power of a glass surface caused by an exceedingly thin film of silver probably results from a 180° phase-difference between the disturbances reflected from the two surfaces.

Relation between the Emission Coefficients of Metals and the Temperature. — Inasmuch as the optical properties of metals, at least for long waves, depend upon their electrical conductivity, we should expect their emission and reflection coefficients to vary with the temperature, except perhaps in the case of the alloy "constantan," the electrical resistance of which is nearly independent of temperature. Hagen and Rubens [1] made a very exhaustive investigation of the subject and found a most perfect agreement with theory. Up to 5 μ there was very little change in the optical properties of metals, resulting from an increase of temperature. Beyond this point, however, *i.e.* with longer waves, the emission coefficient was found to increase with the temperature, the effect being most pronounced in the case of metals having a large temperature coefficient of electrical conductivity. The metals were heated to known temperatures, and their emission compared with that of a black body at the same temperature. Constantan showed very little change with the temperature, as was to be expected.

In studying the corresponding changes in the reflecting power, two metal plates were mounted side by side in an electrically heated oven, and the light reflected to-and-fro from one to the other. In this way the very slight change in the reflecting power with the temperature was magnified.

The band of low reflectivity of silver at 3200 has been observed by de Sélincourt [2] to shift towards shorter wave-length at low temperature. Between −183 and 150° C. the minimum shifted from 3217 to 3247, the width doubling. The theory of metallic reflection is not complete enough at the present time to explain this phenomenon.

[1] *Sitz Ber. der Preus. Akad.*, XVI, 1909; XXIII, 1910.
[2] *Proc. Roy. Soc.*, *107*, 247, 1925.

CHAPTER XVII

ROTARY POLARIZATION

We have seen that, in general, when a ray of plane-polarized light is passed through a crystal in the direction of its optic axis, there is no double refraction, and the light emerges with its plane of polarization unchanged. The discovery was made by Arago in 1811 that a rotation of the plane occurs when light is transmitted through quartz in a direction parallel to the optic axis. If two Nicol prisms are placed in front of a sodium flame and so oriented as to completely extinguish the light, the introduction between the prisms of a quartz plate, cut perpendicular to the axis, causes the field to become bright again. On turning the analyzing Nicol through a certain angle the light can be completely extinguished, showing that it is still plane-polarized, but that the plane of polarization has been rotated through an angle, which is measured by the angle through which the Nicol has been turned. If white light is used the field appears colored, the colors changing as the analyzing Nicol is rotated, the light never disappearing entirely, as was the case with the monochromatic sodium flame. This is due to the fact that the different colors are rotated by different amounts, the phenomenon being termed rotatory dispersion.

The amount of rotation was found to be proportional to the thickness of the crystal section, which shows that the action occurs within the medium and not at the surfaces. Moreover, some crystals were found to rotate the plane to the right, while others turned it to the left. The former are termed right-handed or dextrogyrate, the latter left-handed or laevogyrate. Simple inspection of the crystals is sufficient to determine their character. One can usually find small hemihedral planes which cut off the obtuse angles at the base of the hexagonal pyramid. The character of the crystal is determined by the relative positions of the surfaces s and x. If s and x lie to the right of R, the surface of the hexagonal pyramid which both touch, the crystal is dextrogyrate, if to the left, laevogyrate, as in the upper figure. The surface x is often absent, in which case we can determine the nature of the crystal by the striae on the surface of s, which always run towards the position which x would occupy if present, as shown in Fig. 336.

The direction of rotation of the plane is not affected by turning the plate around, consequently if the ray is reflected back through the crystal, the plane of polarization is turned back into its original position. This point is very important in connection with the magnetic rotation, which we shall discuss in a subsequent article, for in the case of substances which acquire rotatory power by being placed in a strong magnetic field, the direction of rotation depends on the direction of the magnetic field, and reflection back through the plate doubles the rotation.

The rotatory power of quartz makes it very easy to determine whether a lens is of glass or quartz, a frequent question in the laboratory. We have only to reflect light from a glass plate at about the polarizing angle, view it through a Nicol set for minimum transmission, and interpose the lens. If it is of quartz, colored rings are seen, as these lenses are usually cut perpendicular to the axis.

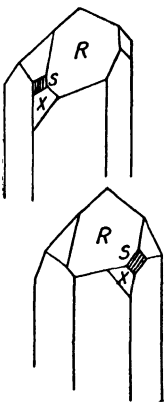


FIG. 336

Rotatory Dispersion. — The phenomenon of rotatory dispersion was investigated by Biot, who found that the rotation was nearly, though not exactly, proportional to the inverse square of the wave-length. Still more accurate measurements were made by Brock, who obtained the following values for a plate 1 mm. thick, for wave-lengths corresponding to those of the principal Fraunhofer lines:

	B	*C*	*D*	*E*	*F*	*G*
ρ,	17° 30′	17° 24′	21° 67′	27° 46′	32° 50′	42° 20′
$\rho\lambda^2$,	7238	7249	7511	7596	7622	7841

The values of $\rho\lambda^2$ are seen to increase with decreasing wave-length, which shows that Biot's law is only approximately followed. Brock's experiment, which is well worth repeating, consisted in passing sunlight through two Nicols, between which a quartz plate was mounted, and then analyzing the light with a spectroscope. On turning the analyzing Nicol, a dark band entered the spectrum from the red end, and passed slowly down towards the violet as the Nicol was rotated. By setting the centre of the band on a Fraunhofer line and reading the position of the Nicol, the rotation for the wave-length of the line was determined.

Stefan calculated from Brock's measurements the following empirical formula $\phi = -1.581 + (8.0403/\lambda^2 10)$, of which the first member represents the departure from Biot's law. If the law held

rigorously it would be equal to zero. Cauchy's formula for ordinary dispersion is

$$n = A + \frac{B}{\lambda^2},$$

in which, however, A is always positive.

The rotatory power of quartz must be kept in mind in all experiments with polarized ultra-violet light. If quartz plates or lenses are used, the planes of polarization of the different lines in the spectrum will be in every conceivable direction owing to the enormous rotatory dispersion. The difficulty can be overcome by employing equal optical paths of right- and left-handed quartz. Lenses can be made which do not rotate the plane, by combining two plano-convex lenses, one of right, the other of left-handed quartz. Optical apparatus is frequently designed without paying due reference to this phenomenon. The following table will be found extremely useful in all polarization work involving the use of quartz plates or lenses:

ROTATION OF POLARIZATION PLANE BY QUARTZ PLATE 1 MM. THICK

ELEMENT	λ	ROT. IN DEGREES	ELEMENT	λ	ROT. IN DEGREES
	2.140	1.60	Cd	3612	63.63
	1.770	2.28		3440	70.59
	1.450	3.43	Cd	3252	80.46
	1.080	6.18	Cd	2746	121.06
	.7600	12.67	Cd	2572	143.26
	.6708	16.54	Cd	2314	190.42
	.6563	17.32	Cd	2266	201.82
Na	.5890	21.68	Cd	2195	220.73
Na	.5896	21.73	Cd	2147	235.97
Hg	.5461	25.70			
	.5270	27.54			
Cd	.5086	29.72			
	.4861	32.77			
	.4340	42.60			
Hg	.4047	48.93			
	.3934	52.15			

If we employ a plate a centimetre or more in thickness the spectrum will be found to be crossed with a number of dark bands which are very nearly equidistant, and which move down the spectrum as the Nicol is turned. The angular rotation for wavelengths corresponding to two adjoining bands differs of course by 180°.

Fresnel's Explanation of the Rotation. — A theory was formulated by Fresnel to account for rotatory polarization, which was

based upon the fact that a ray of circularly polarized light is propagated without change through a quartz crystal in a direction parallel to the optic axis. A linear vibration can be regarded as the resultant of two oppositely polarized circular vibrations, and Fresnel made the assumption that the plane-polarized light, upon entering the crystal, was decomposed into two oppositely polarized circular vibrations, which were propagated with unequal velocities. This inequality in the velocity of propagation will produce a rotation of the resultant, which will amount to 90° after a thickness has been traversed such that one circular disturbance is half a wave-length ahead of the other.

Two points rotating about the same circle in opposite directions with equal velocities represent the two circular vibrations, and their resultant will be harmonic motion along the diameter joining the points at which they pass each other. The amplitude of the resultant plane vibration will be twice the radius of the circle as shown in Fig. 337.

If one point rotates faster than the other, the resultant diameter will rotate in the direction of the faster component. The same will be true if the periods are the same, and one circular component travels faster than the other in the medium. The amplitude relations can be shown as follows.

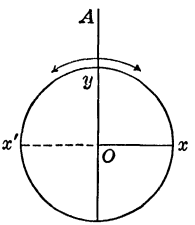


FIG. 337

The components of the two circular vibrations have amplitudes ox and oy (giving counter-clockwise rotation if oy is a quarter period behind ox) and ox' and oy giving clockwise rotation. The resultant is a linear vibration of amplitude OA since the x components are oppositely directed and contribute nothing, while the y components are added.

The rotation is thus seen to be the result of a double refraction for circularly polarized light, the turning of the plane resulting from a progressive displacement of phase of the oppositely rotating circular vibrations into which the plane vibration is resolved.

The existence of these two circular components was shown experimentally by Fresnel, who reasoned that if there existed in reality two circularly polarized disturbances which travelled with different velocities, they should be refracted by different amounts on emerging into the air through an oblique surface; in other words, quartz ought to show feeble double refraction in a direction parallel to the optic axis, and the two images produced thereby should be circularly polarized. Fresnel first tried a single surface, but the effect was too slight to be noticed. By the ingenious device of

building up a compound prism composed of alternate prisms of right- and left-handed quartz the effect was found (Fig. 338). This method permitted the use of very oblique refracting surfaces, since only the small differences in the velocities came into play, while in the case of a single prism we have the refractive index with respect to air to consider, and if the prism angle be too large, light cannot be made to traverse it. By Fresnel's device the separation is increased at each surface, as will be seen from the following consideration. The difference between right- and left-handed quartz lies in the fact that the right-handed circular component is the faster in the former, the slower in the latter, that is, the *R* prisms in Fig. 338 act as the rarer, the *L* prisms as the denser media with respect to this component. The reverse holds true for the left-handed component, the *R* prisms being the denser in this case. The former component is therefore bent down and the latter up, the angular separation increasing at each surface. If the two images of a plane-polarized source seen through the compound prism are examined with a Nicol, they are seen to remain unaltered when the Nicol is rotated. This means that the compound crystal prism has either depolarized the light completely or transformed it into circularly polarized light.

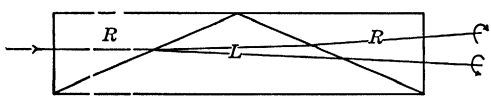


FIG. 338

The introduction of a quarter-wave plate of mica causes the images to disappear in succession as the Nicol is rotated, which proves the two images to be circularly polarized, and in opposite directions, the mica plate transforming them into plane-polarized images, the planes of polarization being at right angles.

Cornu has recently shown that with a single 60° prism of quartz the separation of the rays can be shown, the angle amounting to 27″ for sodium light. It is very much greater for the ultra-violet and the two images can be seen on the uranium glass focussing screen of a small quartz spectrograph if one-half of the Cornu prism is used, the last lines of the Cd spark appearing double.



FIG. 339

By placing the two halves of a Cornu together as in the Fresnel prism, the circular polarization can be observed in regions of longer wave-length. This means that even a quartz prism so cut that the rays travel along the optic axis yields double images. To remedy this defect Cornu devised a prism consisting of two right-angled prisms, the one of right-, the other of left-handed quartz (Fig. 339). It is apparent that the optical path parallel to the base of the prism is the same

for the two circular components; consequently the double refraction, which would otherwise cause a doubling of the spectrum lines, is eliminated.

Prisms of this type are used in practically all of the quartz spectrographs constructed at the present time.

Unequal Absorption of the Circular Components. — The existence of the two circular components was shown in another way by Dove,[1] who found that in colored crystals of quartz (amethyst) the two disturbances were absorbed in different amounts. A similar phenomenon has been more recently observed by Cotton, who found that the strongly colored solutions of certain tartarates absorb right- and left-handed circular vibrations in unequal amounts. These cases will be referred to again when we consider the theory of rotatory polarization.

Other Rotatory Crystals. — Descloizeaux found that crystals of cinnabar have a rotatory power similar to quartz, but fifteen times stronger, some crystals being right-handed, others left. Sulphate of strychnia and sulphate of aethylendiamin have the same property.

The curious discovery was made by Marbach that crystals of sodium chlorate show the phenomenon of rotatory polarization in all directions, sections 1 mm. thick turning the plane (yellow light) 3.7 degrees, no matter how they are cut from the crystal.

Sulphate of strychnia has the power of rotating both in the crystalline state and in solution. As we shall see presently, many solutions exhibit the phenomenon of rotatory polarization, but in general the substances crystallize with two optic axes, the double refraction which occurs in all directions masking any rotatory power which may be present. If, however, the substance can be obtained as an amorphous solid the rotatory power is preserved, if due to the molecular structure. Quartz in an amorphous condition (fused, for example) does not have the rotatory power, neither have solutions of quartz in potash. This makes it quite certain that in the case of quartz and other active crystals, the optical activity depends on the arrangement of the molecules in the crystal, while in the case of substances which are active in the dissolved state the property depends on the arrangement of the atoms making up the molecule.

Rotatory Power of Biaxial Crystals. — The rotation of the plane of polarization by biaxial crystals is more difficult to show as it is masked by the double refraction. Longchambon,[2] however, by using very narrow beams of light measured the rotation of a num-

[1] *Pogg. Ann.*, cx.

[2] *Comptes Rendus*, *172*, 1187, 1921 and *175*, 174, 1922.

ber of biaxial crystals of metallic salts and sugar, the latter showing the remarkably high rotation of 54° along the long axis and −16° along the short, while magnesium sulphate showed a rotation of 20° for both axes. The rotation of sugar is thus seen to be nearly three times as great as that of quartz, and since it also has rotatory power when in solution we can ascribe the high rotation of the crystals as in part due to the molecules and in part to their arrangement.

Rotatory Polarization of Liquids. — The rotation of the plane of polarization by liquids was accidentally discovered by Biot in 1815, while experimenting upon the effect of the surrounding medium upon the colors of thin crystalline plates. The changes which he observed in the colors of plates immersed in oil of turpentine led him to try the effect of the oil alone on polarized light. He found that the turpentine behaved in a manner similar to a quartz plate cut perpendicular to the axis, turning the polarization plane to the left by an amount proportional to the thickness of the fluid traversed, and approximately inversely proportional to the square of the wave-length.

On examining other substances for this property he found that a large number of fluid organic compounds behaved in a similar way, some turning the plane to the left, others to the right. The rotatory power was, however, very much less than that of quartz, a column of turpentine 100 mm. long turning the plane of sodium light 37°. An equal length of a solution of santonid or parasantonid in a mixture of chloroform and bisulphide of carbon, which is the most active liquid substance known, gives a rotation of 446°, while a quartz plate of equal thickness would turn the plane 2167°, or six complete revolutions.

Biot found further that a mixture of an active substance with an inactive one, chemically inert towards it, had a rotatory power proportional to the amount of active substance present; in other words, the rotation was simply proportional to the number of active molecules in the path of the light, and wholly independent of the proximity of similar or different molecules.

The neutral effect of dilution with an inactive substance made the investigation of solid substances possible by bringing them into solution in some optically inactive liquid.

In this way the number of active organic substances was enormously increased, many sugars, gums, albumens, alkaloids, fruit acids, etc., being added to the list. We shall see presently, however, that in some cases the solvent, even if optically inactive, is not without influence on the rotatory power.

Rotatory Power and Change of State. — Biot found that substances which were active in solution preserved their property in

the solid state, provided they were prevented from crystallizing. Sugar and tartaric acid can be obtained as amorphous solids, and are optically active in this state. If the substance crystallizes, the double refraction completely masks the rotatory polarization.

Bringing the substance into the vaporous state does not affect its activity, as Biot ascertained in 1818 by means of a tube 15 metres long filled with the vapor of turpentine. His apparatus took fire before the completion of the experiment and was destroyed, and it remained for Gerney to make careful measurements of the rotatory powers of vapors. He found that the specific activity was the same in the vaporous state as in the liquid, *i.e.* a long column of vapor has the same rotatory power as the short column of liquid into which it condenses (the cross sections being, of course, equal).

Rotatory Dispersion of Liquids. — The rotatory dispersion of liquids was found by Biot to be similar to that of quartz. That the increase in rotation is not strictly proportional to the inverse square of the wave-length he showed by filling a tube with a mixture of dextro-rotatory and laevo-rotatory liquids in such proportion that perfect compensation was secured for a single color. If the law of the inverse squares of λ was strictly followed, the compensation would be perfect for all wave-lengths, which was found not to be the case, the light appearing colored through the analyzer. This is analogous to the Christiansen colors exhibited by a mixture of a powder with a liquid of the same mean refractive index, but higher dispersion, described in the Chapter on Dispersion. The departure from the law is different for different liquids, being very small for the oils and very large in the case of a solution of camphor in alcohol.

Molecular Rotation and the Influence of the Solvent on the Rotatory Power. — The fact that the rotation produced by a given substance appeared to be proportional to the number of molecules in the path of the light, gave rise to the opinion that the rotatory power was inherent in the molecule, and led to the term molecular rotatory power. This we may define as the amount of rotation produced by a column of the solution 1 dm. in length containing 1 gram of the substance per cubic centimetre. If we dissolve p grams of the substance in q grams of the solvent, the density being δ, then $p/(p+q)\,\delta$ is the amount of the substance contained in unit volume of the solution, and if we fill with this solution a tube of length l and observe a rotation for some particular wave-length, we have

$$\rho = [\rho]\,\frac{p}{p+q}\,l\delta,$$

in which $[\rho]$ is a constant for the substance, and is defined as the molecular rotatory power. This constant may also be defined as the rotation produced by a thickness of 1 dm. of the pure substance divided by the density of the substance. Biot found that the rotation was not strictly proportional to the amount of dissolved substance, and that it varied, moreover, with the nature of the solvent. If the change in the rotatory power with changing concentration is continuous, a formula may be deduced by which we may determine the molecular rotatory power of the pure substance from observations of solutions.

This matter was very carefully investigated by Landolt, who worked with fluid substances, so that the activity of the pure substance could be directly determined, and then compared these values with the values calculated from observations made with the substance dissolved in various inactive solvents.

He found that the molecular rotatory power could be expressed as a function of the quantity of active substance contained in solution. If q represents the weight of the solvent in 100 parts by weight of the solution, an equation of the following form could be applied:

$$(\rho) = A + Bq + cq^2.$$

The constants could be determined by making observations with solutions of various concentration, the constant A being the molecular rotation of the pure substance ($q=0$).

In this way both the effects of the nature of the solvent and the degree of concentration are eliminated.

Anomalous Rotatory Dispersion. — In the case of solutions of tartaric acid in water, the dispersion at first increases with decreasing wave-length, reaches a maximum, and then decreases. Measurements made by Arndtsen [1] for 50 parts of crystallized acid in 50 parts of water gave the following values:

C	*D*	*E*	*b*	*F*	*e*
11°.9	13°	14°	13°.7	13°.3	10°.3

The formula for this substance is $[\rho] = A + Be$, in which e is the percentage of water, the value for A and B for the various colors being

	C	*D*	*E*	*b*	*F*	*e*
$A =$	+2°.748	1°.95	0°.153	−°.832	−3°.598	−9°.657
$B =$	9°.446	13°.03	17°.514	19°.147	23°.977	31°.437

From this it is apparent that the pure substance in an amorphous state is dextro-rotatory for all wave-lengths on the red side of a

[1] *Pogg. Ann.*, cv.

point in the spectrum a little below the *E* line, and laevo-rotatory for wave-lengths below this point, a circumstance which had already been noticed by Biot in the case of amorphous plates of the acid, made by fusing the crystals and pouring the liquid on a glass plate.

In the case of active substances which show strong selective absorption, we may have true anomalous rotatory dispersion when we cross the absorption band. Such cases are not to be confused with the one just mentioned, which in all probability owes its peculiarity, as we shall see when we take up the theory of the rotation, to the presence of both a dextro- and laevo-rotatory system within the molecule. We have a somewhat remote analogy in the achromatic prism. Considered as a single dispersing system it may be said to show anomalous dispersion, the outstanding colors due to imperfect compensation being arranged in anomalous order, as we have seen.

Double Refraction of Active Liquids. — The division of the plane-polarized ray into two oppositely polarized circular disturbances was shown experimentally by F. v. Fleischl [1] by a method identical with the one employed by Fresnel in the case of quartz. A long narrow trough was divided into 22 prismatic compartments by means of oblique partitions of plane-parallel glass. These compartments were filled alternately with dextro- and laevo-rotatory liquids of the same refractive index, the best results having been obtained with oil of orange and a mixture of the oils of turpentine and ricinus. A small source of polarized-sodium light appeared doubled when viewed through the compound-fluid prism, and examination with a quarter-wave plate and Nicol prism showed the two images to be circularly polarized in opposite directions.

Rotatory Polarization in Convergent Light. — The behavior of quartz plates cut perpendicular to the optic axis was investigated both theoretically and experimentally by Airy in 1831.[2] Inasmuch as only circularly polarized rays are propagated without change parallel to the axis, and plane-polarized rays perpendicular to the axis, Airy made the hypothesis that in any other direction the only form of vibration capable of being propagated without change was an elliptical one, assuming that a plane-polarized ray incident in a direction inclined to the axis was decomposed into two oppositely polarized elliptical vibrations which travelled through the crystal with different velocities. Airy further assumed that the ellipses were similar, and that the path-difference between the two elliptically polarized rays was the same as in ordinary non-rotatory uniaxial crystals, *increased* by an amount which was independent

[1] *Ber. Wien. Akad.*, 1884.
[2] *Camb. Trans.*, v, Part 1, pp. 79–198.

of direction, and inversely proportional to the square of the wave-length. This means that in active crystals the two wave-surfaces are completely separated, the very small distance between them at the points where they intersect the optic axis being the increment referred to above. The wave-surfaces are shown in Fig. 340.

The sphere is slightly pulled out, and the spheroid slightly retracted as shown in the figure which represents a section of the wave-surface. The elliptical vibrations pass over into plane vibrations at a very small angle with the optic axis, but they have been shown experimentally with an arrangement of prisms similar to the one for showing circular polarization.

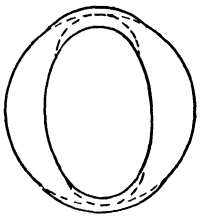


FIG. 340

Physical Explanation of the Rotation. — In the case of rotation by crystals we can refer the phenomenon to the crystalline structure. Ewell [1] has shown that twisted gelatine cylinders show rotatory polarization. If we imagine a bar of elliptical cross section which has been twisted torsionally, we have a rough analogy, which may help us to understand how a spiral arrangement of the axes of maximum and minimum elasticity may account for the rotation of a plane-polarized vibration.

In the case of liquids and solutions, however, we must necessarily refer the rotatory power to the structure of the molecule. The majority of rotatory liquids contain carbon, and their power has been ascribed by Le Bel and Van't Hoff to the quadrivalence of this element.

If four different atoms or radicals, in combination with the carbon atom, form the corners of a regular tetrahedron, we can arrange them in two different ways, such that one is the looking-glass image of the other, and yet no amount of turning enables them to be brought into coincidence. One is dextro-, the other laevo-rotatory. Certain nitrogen and phosphorus compounds have recently been found showing rotatory power, aminoxyde and phosphinoxyde for example.

The earliest attempts to explain rotatory polarization were based upon an experiment made by Reusch, who found that, if thin mica plates were superposed, each plate having its principal section turned through a definite angle either to the right or left, with reference to the principal section of the plate below, the combination imitated the behavior of a quartz plate cut perpendicular to the axis, rotating plane-polarized light to the right or left according as the pile of plates were built up clockwise or counter-clockwise. The thinner the plates and the greater their number, the more nearly the pile imitated a rotatory crystal. It was quite

[1] *American Journal of Science*, 8, 89, 1899.

natural, in view of this very suggestive experiment, to ascribe a somewhat similar structure to quartz.

Theory of Rotatory Polarization. — The modern theory of the rotatory polarization by molecules is due chiefly to Born [1] and Oseen.[2] The theory differs from the dispersion theory in that it considers the molecular diameter as not negligibly small in comparison to the wave-length, and that the oscillators are bound together by a coupling force, and act on each other reciprocally. This coupling force has the directional properties of a screw, *i.e.* a displacement of one of the particles parallel to the x axis causes a displacement of the other along y. The cause of this will appear presently. A very complete mathematical treatment of the whole matter has been given by W. Kuhn.[3] The action of the resonators can, however, be followed without the use of equations. As the simplest model of a molecule having rotatory power consider one having two oscillators (ions) of different natural frequencies, one of which can vibrate only along the x axis, the other only along the y. They are coupled, so that a vibration of one of them starts a small forced oscillation of the other. The equilibrium position of the one capable of vibration along x is at the origin of coördinates, that of the other, on the z axis, above the origin as shown in Fig. 341. These oscillators we will designate as X and Y, and their natural frequencies ν_1 and ν_2. The separation of the two oscillators by a distance comparable to the wave-length is essential. For very high and very low frequencies the rotation is zero. It will now be shown that this system responds in a different degree to right- and left-handed circular vibrations, and the scattered radiation which it emits, and which alters the phase and hence the velocity of the primary beam (as in dispersion theory) is different in the two cases.

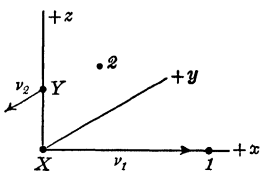


FIG. 341

Rotation of the plane of polarization occurs only as a result of absorption bands caused by coupled oscillators of this type, which explains why certain absorption bands are found to be without influence in governing the unequal refractive indices for the circular vibrations, while others are directly contributory.

The force which operates on the first resonator is $Kx_1 = -k_1x_1 - k_{12}y_2$ in which x_1 and y_2 are the displacements of the two resonators, k_1 the force along x for unit displacement, and k_{12} the coupling force. The minus sign is taken since the direction of the

[1] *Phys. Zeit.*, *16*, 251, 1915.
[2] *Ann. der Phys.*, *48*, 1, 1915; *55*, 177, 1918.
[3] Freudenberg's *Stereochemie*, 3 Lieferung 348.

force is opposite to the direction of displacement. For the second resonator we have $Ky_2 = -k_2y_2 - k_{12}x_1$. In both cases the coupling force acts in the same direction as the restoring force if all displacements are positive. The force acting on each resonator thus depends upon the position of the other. If k_{12} and y_2 are positive, an additional force acts on particle X in the negative x direction, and when x_1 is positive the additional force acts on particle Y in the negative y direction. We will designate the compound vibration, in which particle X of frequency ν_1 executes the large amplitude, giving rise to a small forced oscillation of Y, as *vibration X*, while vibration Y designates the compound vibration of particle Y with frequency ν_2. We shall consider $\nu_2 > \nu_1$.

As a cause of the circumstance that the two ions can vibrate only along the x and y axes, and that a motion of X along x in the positive direction is accompanied by a small motion of Y along y, in the negative direction, it is necessary to suppose two other ions 1 and 2 present in the positions shown in Fig. 341. This is the simplest model that will function as required, and experiments show that nothing simpler than a molecule made up of four or five atoms (for example, a carbon atom at the centre of a tetrahedron is capable of causing rotation of the plane of polarization).

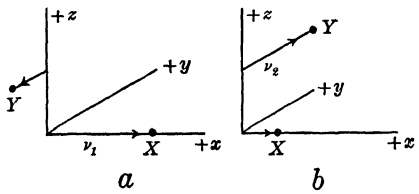


FIG. 342

In the case of a resonator excited by light we have seen, in the Chapter on Dispersion, that the resonator executes forced vibrations in phase with those of the light, if the excitation is by a lower frequency than the natural frequency of the resonator, but with a phase-difference of 180° if the excitation is by a higher frequency. In the present case the relation is reversed, since a + displacement of particle X causes a − displacement of particle Y, whereas in the case of displacements due to light-waves a resonator is displaced in the direction of the electric force. If $\nu_2 > \nu_1$ and we consider particle X in vibration, the forced vibration of particle Y, being excited by a frequency less than its natural frequency, will vibrate with a phase-difference of 180°, *i.e.* as particle X moves in the $+x$ direction, particle Y will move along $-y$ (Fig. 342*a*). If however we consider particle Y as exerting the driving force, we are now exciting by a higher frequency than the natural frequency of particle X and as Y moves in the $+y$ direction, X moves in phase with it, *i.e.* along $+x$ (Fig. 342*b*).

As a result of this coupling the two compound vibrations X and Y respond differently according as they are excited by right- or left-handed circularly polarized light. As it is difficult to represent the model by a two-dimensional diagram, the stereoscopic photograph reproduced in Fig. 343 was prepared in which the spiral wire represents the passage of right- and left-handed circular light across the coupled resonators.[1]

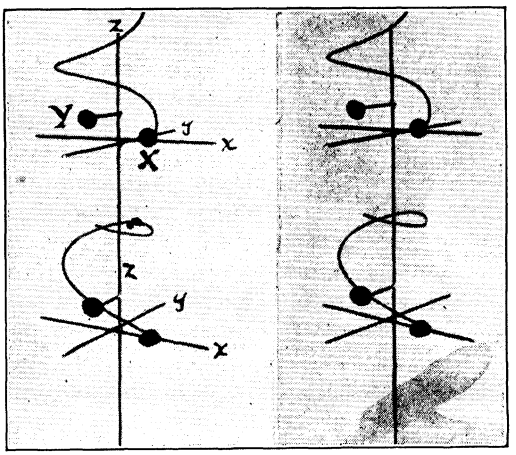


FIG. 343

It will be remembered that by right-handed rotation we mean clockwise rotation as viewed from the side towards which the light travels. In the stereoscopic figure the light is descending, *i.e.* travelling along the z axis towards the origin. Our viewpoint is therefore from below. The direction of the electric field in any plane parallel to the xy plane is given by a line drawn in the plane from the z axis to the spiral. As the upper spiral is pushed down (without rotation) the line from z to the spiral represents an electric vector of constant value turning to the right, *i.e.* right-handed circularly polarized light.

In the figure the positions of the two particles are represented for the case in which particle Y is being driven by particle X (phase-difference of 180°). In the upper figure particle X has been displaced along the x axis in the + direction by right-handed circular light, represented by the upper spiral, while particle Y has moved in the $-y$ direction as a result of the coupling force. The direction of the electric force of the light, in any plane parallel to the xy plane is given by a line drawn from the z axis to the spiral, and we thus see that the electric force of the light is operating in a direction opposite to the direction of displacement of Y, at the moment at which it is directed along x, as in the upper diagram, while in the lower diagram, which represents excitation by left-handed light, the force is acting in the same direction as the displacement of Y. The response will therefore be more vigorous in the latter than in the former case, and the secondary wave given

[1] See footnote, page 407, for directions for viewing.

out, due to the compound vibration X, will in consequence have a greater intensity. This will result in a higher refractive index on the long-wave-length side of the absorption band ν_1 for left-handed than for right-handed circular light, which is all that is required to account for the rotation of the plane of polarization.

If the coupling force is negative, we have all of the above conditions reversed, and the rotation of the plane is in the opposite direction.

If we consider the response of compound vibration Y, however, the model shows that the condition is exactly the opposite, for in this case the upper resonator would be in the $+y$ direction instead of the $-y$ (as shown in the figure) and the stronger response would be for the right-handed polarized light. The result of this is that, if absorption bands ν_1 and ν_2 are both in the ultra-violet, the rotation in the visible spectrum will be the difference between the positive rotation due to one and the negative rotation due to the other.

We also see from the diagram that if both particles are in the xy plane the electric force of the light operating on particle y is along x, *i.e.* in a direction in which it is incapable of moving.

The Secondary Wave from the Compound Resonator. — In the Chapter on Dispersion Theory we have seen that refraction is accounted for by compounding the secondary waves scattered by the molecules of a medium with the primary waves. The difference in velocity of right- and left-handed circular light in a medium containing coupled resonators can be shown in the same way.

If the X vibration of the coupled resonators is excited, it is clear from the stereoscopic figure that, owing to the accompanying small vibration of Y, the coupled resonators will emit a wave which is the resultant of two linear vibrations at right angles to each other, one large and one small, with a path-difference $\lambda/4$, the distance of Y above the origin. This resultant will be a left-handed elliptical vibration when viewed from the direction towards which the exciting light is travelling. Excitation of the Y vibration gives right-handed elliptical light. In both cases the direction of revolution of the exciting light plays no part. The elliptically polarized light from a single molecule or a collection of similarly oriented molecules cannot be compounded with the circular exciting wave to account for the dependence of velocity on the direction of rotation of the electric vector in the light. All that we should have in this case would be double refraction light, *i.e.* compounded of a large component along x and a small one along $-y$ retarded by $\lambda/4$, if this is the distance between the particles when both are on the z axis. A left-handed elliptical vibration cannot compound

to a circular resultant with right-handed circular light, but if there are a number of molecules oriented at random at the point $z = 0$, a resultant right-handed circularly polarized wave will be emitted, which by compounding with the primary wave causes the phase retardation that constitutes refraction.

This can be most easily seen from the following considerations. We will suppose two molecules situated at the origin similarly oriented except that one is rotated around the z axis 90° with respect to the other. We will at first neglect the vibration of resonator Y which produces the left-handed ellipticity, and consider only the action of the light on resonator X. The two resonators, started one after the other as the circularly polarized wave passes them, will execute linear vibrations at right angles to each other, with a phase-difference of a quarter of a period, owing to the restrictions placed on their direction of motion, and the resultant wave due to the two molecules will be right-handed circular light if the exciting light is right-handed. The same will hold for two elliptical vibrations of large eccentricity (such as we are dealing with), regardless of their direction of rotation, in other words the circumstance that they compound into a right-handed circular vibration does not result from the fact that the ellipses are left-handed.

If, however, we have two molecules at the origin similarly oriented, except that one is rotated about the z axis through an angle of 90° with respect to the other, and excite them with left-handed circular light, the secondary wave given off by the pair will be circularly polarized (left-handed) and by combining with the exciting wave give rise to refraction.

Consider first the vibration of the two X vibrators, one vibrating along x the other along y (owing to the 90° rotation of the molecule). Referring to the lower stereoscopic model of Fig. 343 we see that the X vibration of the rotated molecule will be a quarter period behind that of the other molecule, since the descending spiral cuts the $-y$ axis after an advance of $\lambda/4$ from the position shown. The two X vibrations thus compound into a left-handed circular wave. This, by combining with the exciting wave, would cause ordinary refraction. These two vibrations, however, set the two Y vibrators in motion, owing to the coupling force, and these will give rise to a small circular vibration, also left-handed, but with a phase advance of a quarter of a period with respect to the vibration given out by the X vibrators, as shown in Fig. 344*a*, in which the dotted lines represent the vibrations of the molecule which has been rotated 90°. In this figure the vibration of the Y vibrators is to be regarded as above that of the X, *i.e.* above the plane of the paper, and the viewpoint is below the paper in de-

termining the direction of rotation. Since the disturbance from the Y vibrators (in the $-z$ direction) is $\lambda/4$, behind the disturbance from the X vibrators the phase advance will be compensated by the path-difference and the amplitude of the wave from the X vibrators will be increased by the small circular vibration.

If, however, we excite by right-handed circular light, the upper stereoscopic diagram shows that the circularly polarized wave given out by the two Y vibrators is $\frac{1}{4}$ of a period behind the circular wave from the X vibrators (Fig. 344*b*) and the addition of another quarter period due to the path-difference brings it into opposite phase with respect to this wave and causes diminution of amplitude, in other words the secondary wave from the pair of molecules has a smaller intensity when excited by right-handed circularly polarized light. There is in addition the less vigorous response of each molecule in this case, due to the circumstance that the electric force of the light is in a direction opposite to the displacement of the Y vibrator which is due to the coupling force, which causes a further reduction of intensity.

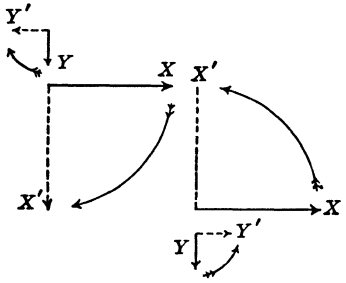


FIG. 344

The above conditions hold as well for a large collection of molecules oriented fortuitously. The circumstance that the scattered light is less intense when excited by right-handed light, means a lower refractive index than when the excitation is by left-handed light. This accounts for the rotation of the plane of polarization.

CHAPTER XVIII

RESONANCE RADIATION AND FLUORESCENCE OF ATOMS

In this chapter we shall consider the transformation which radiation undergoes when traversing a gas or vapor capable of responding to one or more of the frequencies present in the beam. If the light which has passed through the vapor is examined with a spectroscope, dark lines, usually termed absorption lines, appear in the spectrum, corresponding to the frequencies abstracted from the light by the gas. As we shall see presently, however, there is frequently no true absorption in these cases, the frequencies, absent in the transmitted light being scattered laterally, sometimes with and sometimes without a change of wave-length. True absorption, or the change of the radiant energy into heat, results only as a result of collisions between the atoms or molecules, consequently if we wish to observe the fundamental processes we must operate with the gas at such low pressure that the atoms are, for the greater part of the time, beyond the influence of the electrical fields of their neighbors. This means that the pressure must not be greater than a very small fraction of a millimetre of mercury.

The absorption of light by a gas was explained on the classical theory as a resonance effect similar to the excitation of a tuning-fork by sound-waves of the same frequency as its own, and one would naturally expect the excited atoms to emit light of the same frequency as that abstracted from the incident beam in a manner analogous to the case of the acoustically excited fork.

This later reëmission of light by a resonating gas escaped detection for a long time owing to the fact that the necessary condition of very low pressure was not realized.

The phenomenon was first detected in 1904 [1] by the author and named resonance radiation, during an experiment on the fluorescence of sodium vapor in a highly exhausted glass bulb, illuminated by a monochromator giving light of wave-length equal to that of the *D* lines. It was observed to still better advantage soon after in a test-tube containing a fragment of sodium and closed at the top with a flat plate of glass. The tube was exhausted, heated at the bottom with a small flame, and the light of an oxyhydrogen

[1] *Proc. Am. Acad.*, xl, 396.

flame rich in sodium focussed on the axis of the tube above the metal. On looking down into the tube the track of the illuminating beam was seen as a cone of yellow light, which presently drew back until there remained only a thin skin of yellow glow lining the inner wall at the point where the exciting beam entered.

This was due to the circumstances that, as the vapor density increased, the penetration of the light became less, until finally it was stopped by a thin layer of atoms adjacent to the glass wall of the tube. The same phenomenon was discovered by photography in the case of mercury vapor for the ultra-violet line of wave-length 2537, in which case the vapor was contained in a highly exhausted bulb of fused quartz containing a drop of mercury, which required no heating, as the resonance radiation appeared at its maximum intensity at room temperature, that is at a pressure of about 0.001 mm.[1]

We will now take up the study of different types of resonance radiation beginning with that of mercury in which the phenomenon is exhibited by the vapor in its normal or unexcited state. Other gases such as helium or neon which, in the normal state, show no absorption, at least in the visible region can be made to exhibit visible resonance radiation by bringing the atoms into an excited condition by means of feeble electrical discharges. This distinction will be more fully explained presently.

Resonance Radiation of Mercury Vapor. — A tube of fused quartz, closed by flat end-plates, and containing a small drop of

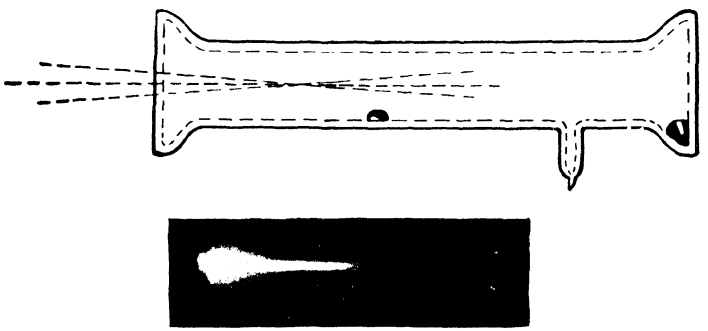


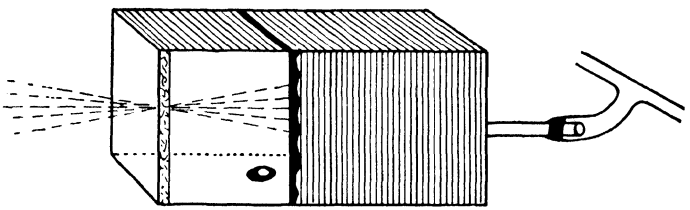
FIG. 345

mercury was exhausted and sealed. The light of a quartz-mercury arc was now focussed along its axis and the tube photographed against a curtain of black velvet from the side with a camera furnished with a lens of quartz. The photograph showed an image of the cone of rays traversing the high vacuum precisely as if the tube

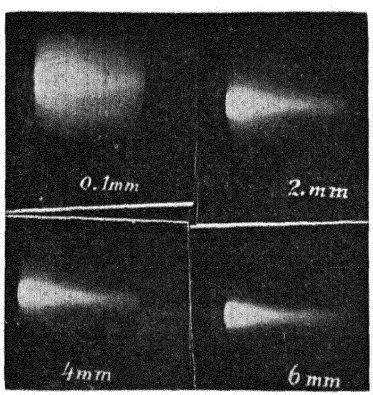
[1] R. W. Wood, *Phil. Mag.*, May, 1912.

were filled with smoke (Fig. 345). The image of the cone of rays on the plate was absolutely black while the region outside of the illuminating cone also showed evidences of powerful radiation. We may term this surrounding luminosity secondary resonance radiation since it is excited by the primary resonance radiation coming from the illuminated track.

A larger chamber for the mercury vapor was constructed of quartz plates and brass in the form of a box, which was put in communication with the pump and a manometer (Fig. 346*a*).



a



b

FIG. 346

It was found that unless the pressure of the air was less than 3 or 4 millimetres no trace of any secondary radiation was present. On lowering the pressure of the residual air it developed rapidly, however, and after the pump had been in operation for several minutes the luminous glow filled the entire interior of the box, the luminous cone being nearly lost on the strong luminous background of secondary radiation. Photographs taken at four different pressures are reproduced in Fig. 346*b*. For quantitative study of these phenomena it is necessary to have a source of light capable of exciting powerful resonance radiation which will operate continuously. After considerable experimenting the following method has been found to give the best results. The vertical type of Cooper-Hewitt quartz-mercury arc, furnished with an anode of tungsten, is immersed in running water to a depth of about one centimetre above the cathode bulb. One pole of an electromagnet, formed by a straight bundle of soft iron wires surrounded by a coil of wire which is included in circuit with the arc,

is now brought up to the side of the lamp in such a position as to cause a deflection of the discharge against the front wall of the tube. This technique was first employed by Kirschbaum. The water reservoir prevents the mercury vapor from acquiring any considerable density, and the magnet, by squeezing the discharge against the wall, still further prevents the reversal of 2537 line. The lamp should be operated with considerable self-inductance in the circuit and the maximum resistance consistent with continuous operation, the current being from one and a half to two amperes. If the best results are to be secured it is better to subject the light from the lamp to a preliminary analysis with a quartz monochromator, illuminating the resonance tube with a monochromatic beam of wave-length 2537. Such an instrument can be easily improvised with a pair of quartz lenses and a quartz prism, a screen being placed in front of the lamp perforated with an aperture a millimetre or two in width and ten millimetres high. The spectrum which is formed in space can be easily located and a black screen mounted just inside the focus in such a position as to intercept all radiations except those which form the 2537 image on the uranium glass screen.

If now we allow the monochromatic cone of rays to pass axially down the quartz resonance tube previously described, we find, on holding the uranium glass behind the tube, that a considerable portion of the monochromatic radiation passes through it without absorption. Photographs show that resonance radiation is emitted only by the first few centimetres of vapor traversed by the beam. This means that the radiation which is capable of passing on through the tube is incapable of exciting resonance radiation. Though we have spoken of the light as monochromatic, we must bear in mind that all lines have a finite width, and that it is only the centre or *core* of the line which is stopped by the mercury vapor, the edges or wings passing on without sensible absorption. In order to get some idea of the stopping power of the mercury vapor for light more nearly homogeneous than that emitted by the arc, a mercury resonance bulb was used as a source of light. This consisted of a small spherical bulb of quartz containing a drop of mercury and highly exhausted, which was illuminated by the converging beam from the monochromator. The radiations from the bulb, which we may term a resonance lamp, fell upon a screen of barium-platino-cyanide, exciting it to luminosity. On placing the mouth of a small bottle containing a little mercury between the resonance lamp and the screen and blowing into the bottle through a glass tube the mercury vapor rising from the neck appeared as a cloud of black smoke on the barium-platino-cyanide

screen, in other words, the mercury vapor at room temperature cast a distinct shadow in the monochromatic light.

Measurements of the intensity of the radiation along the track of the primary beam showed that the core of the line or the radiation capable of exciting resonance was reduced in intensity by 50% after traversing a path of only 5 mms. in mercury vapor at room temperature.

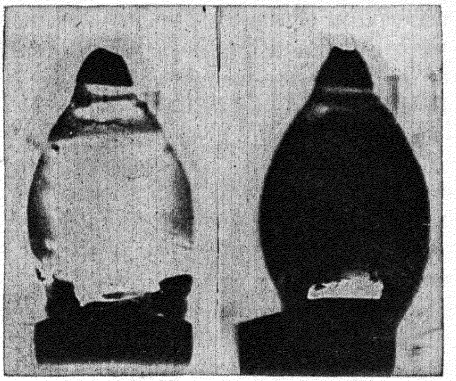


FIG. 347

Two shadow photographs of a quartz bulb, in one case empty and in the other containing a small drop of mercury, cast by the light of the resonance lamp on a photographic plate are reproduced in Fig. 347. The intense opacity of the mercury vapor at room temperature is evident.

Absorption by a Resonating Gas. — If the pressure of the air is gradually increased from zero, the secondary resonance radiation becomes increasingly weak in comparison to the primary, and at pressures above a few centimetres the primary radiation vanishes also. Energy is still being abstracted from the exciting beam, as the absorption spectrum of the transmitted light indicates, and it is clear that the effect of introducing the air is to lessen resonance radiation and increase true absorption. That this is the case is very clearly indicated by the behavior of the secondary resonance radiation. Suppose that, by the introduction of air at the pressure of 3 millimetres, the intensity of the light emitted by the directly excited atoms is reduced to one-third of its original value, the rest of the abstracted energy being absorbed, *i.e.* spent in increasing the velocity of the atoms. By means of a threefold intensity increase of the exciting light we can raise this intensity to its original value, so that the same amount of light is available for the excitation of the secondary radiation as before. The intensity of the secondary radiation excited under these circumstances will, however, be only one-third of its former value, since two-thirds of the energy received from the directly excited atoms is transformed into heat by the true absorption which has been introduced by the presence of the air. The intensity of the secondary radiation in comparison with that of the primary will consequently be much

less (one-third) than when the mercury vapor is in a high vacuum.

Photometric measurements were made by photography of the intensities of the primary resonance radiation with air at different pressures in the bulb. From these values the percentages lost by absorption of the total energy abstracted from the primary beam were calculated on the assumption that at zero air pressure all of the abstracted energy was reëmitted. The values found are given in the following table:

Air Pressure in mms.	Intensity of Resonance Radiation	Percentage of True Absorption
0	100	0
0.45	77	23
1.10	66	33
2.20	56	43
6.2	33	66
9.5	23	77
14.2	17	83
18.	13	87
32.	4	96

Polarization of the Resonance Radiation. — If the temperature of the resonance tube is gradually raised, the cone of resonance radiation draws back toward the window through which the exciting light enters until finally all that remains of the resonance radiation is the thin layer of glowing vapor which lines the inner surface of the front plate. This can be photographed through the end window of the tube through which the light enters. On testing this luminosity for traces of polarization with a Savart plate, the quartz camera being equipped with a polarizing prism, no traces of polarization were found even with complete polarization of the incident light. This was very surprising, inasmuch as strong polarization had been previously found in the case of the fluorescent light emitted by iodine and sodium vapor.

This failure to detect polarization was due to the circumstance that the vapor was employed at too high a density, for the phenomenon was subsequently discovered by Lord Rayleigh who examined the light scattered laterally from the path of a beam traversing mercury vapor at room temperature. He found polarization considerably stronger at some distance from the entrance window than in the portions of the vapor first traversed by the light which would appear to indicate the presence of a stray magnetic field, since recent experiments by Wood and Ellett [1] have shown that the polarization of the resonance radiation is practically destroyed by a magnetic field of the order of magnitude of the

[1] *Proc. R. S., 103*, 396, 1923

earth's field. This remarkable phenomenon will be more fully discussed in the Chapter on Magneto-Optics. With the earth's magnetic field compensated, and the tube cooled in ice to still further lower the vapor pressure, and the illuminating beam plane-polarized to start with, 90% of polarization may be observed in the resonance radiation. If the temperature is gradually raised above room temperature or if small traces of inert gases are allowed to enter the bulb, the resonance radiation shows a decreasing degree of polarization which vanishes at moderate temperatures or pressures.

For controlling the pressure of the mercury vapor in these and similar experiments the resonance tube or bulb should be provided with a small lateral branch which may be dipped in a cooling mixture, the remainder of the tube being at room temperature. Resonance radiation has been observed in mercury vapor with the lateral branch cooled to $-50°$, which shows that there is an appreciable amount of vapor given off by the solidified metal 10° below its freezing point!

Resonance Radiation of Sodium Vapor. — The resonance radiation of sodium vapor described in the earlier part of the chapter was studied more in detail and with improved apparatus by L. Dunoyer,[1] who employed glass bulbs containing a small amount of metallic sodium, highly exhausted and heated to a uniform temperature in the hot air rising from a tall chimney surrounding a large Bunsen burner. All of the phenomena found by photography in the case of mercury can be rendered visible by means of sodium since the vapor resonates to the light of the two yellow sodium lines.

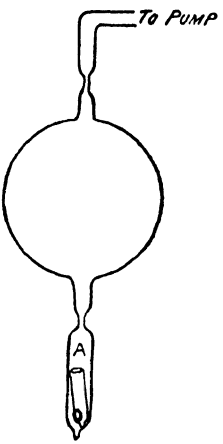


FIG. 348

A thin-walled bulb of pyrex glass of the form shown in Fig. 348 with a small piece of sodium in a minute tube in the lower branch is highly exhausted with a mercury pump, preferably protected by a liquid air trap. Molten sodium can be sucked up into a long thin-walled capillary glass tube of about 1 mm. bore (using a long rubber tube between the mouth and the glass tube). Pieces of the sodium tube can be cut off with scissors as required, protecting the open end of the long tube with soft wax. The bulb should be strongly heated by a Bunsen flame during the exhaustion, care being taken, however, not to heat the sodium. After the air and other adsorbed gases have been eliminated from the walls of

[1] *LeRadium*, *9*, 177, 209, 1912; *10*, 400; *Jour. de Phys.*, *4*, 17, 1914.

the bulb the flame is cautiously applied to the bottom of the tube and the sodium evaporated into the bulb. The silvery deposit should now be driven about from side to side by the Bunsen flame, the pump operating all the time. It is well to continue this process for five or ten minutes as the sodium in condensing on the wall carries down much hydrogen with it. The constricted portion at the bottom is now sealed off and the bulb sealed from the pump. It is held in a light pair of tongs made of iron wire and supported over a stove-pipe heated at the bottom by three or four Bunsen burners.

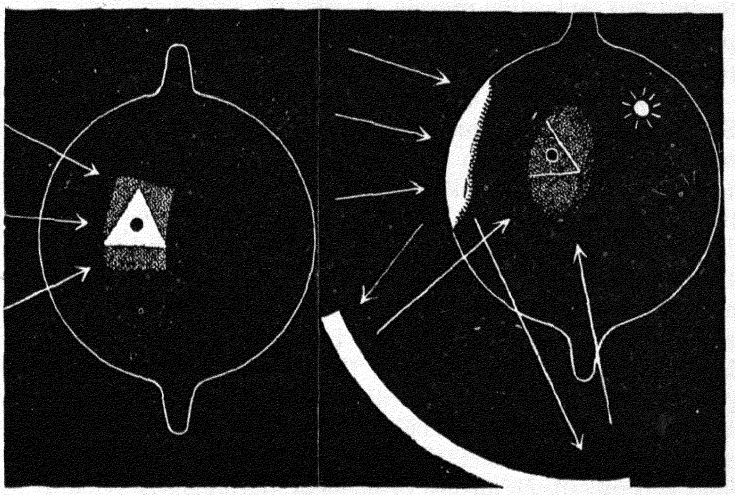
The best source of light for exciting the resonance is the recently developed hot-cathode sodium lamp (Pirani) though fairly satisfactory results can be obtained with a Meker burner charged with sodium by bringing a piece of asbestos paper which has been dipped in brine against the side of the flame which is away from the resonance bulb: this abolishes the self-reversal of the *D* lines. A screen, perforated by an aperture of about 1.5 square cms. is placed close to the brightest part of the flame, and an image of it projected on the bulb by a large lens of the shortest possible focus.

As the temperature of the bulb rises we presently see the path of the illuminating beam glowing with a feeble yellow light which extends clear across the bulb. As the temperature goes up and the density of the vapor increases, this cone of light retreats towards the point at which the light enters the bulb and finally there remains only a sharp image of the circular aperture on the inner surface of the bulb. Only a small part of the sodium light coming from the flame is arrested by the bulb as can be seen by viewing the aperture through the bulb, for just as in the case of mercury vapor it is only the centre of the line (or core) which is capable of exciting resonance. This means that the patch of light seen on the surface of the bulb when the vapor is dense is very much less brilliant than an image of the aperture formed on a piece of white paper which reflects all of the light of the flame. If strictly homogeneous light were employed for the excitation of the vapor the difference of intensity between the resonance radiation and the light diffused by white paper would be much less, and might even disappear entirely if there were no loss by absorption. With the bulb at a lower temperature, such that the luminous track extends clear across the bulb a feeble glow can be seen coming from the vapor which lies outside of the track of the illuminating beam, as shown in Fig. 4 of the colored frontispiece. This is the secondary resonance radiation which has been described already in the case of mercury vapor. It is extremely brilliant with the Pirani lamp.

By studying the diffusing power of dense sodium vapor for the light of the flame in comparison with the diffusing power of mag-

nesium oxide we have a means of determining to what extent true absorption occurs if it exists. This matter was investigated by Wood and Dunoyer.[1]

The relative reflecting powers of magnesia and dense sodium vapor were first determined photometrically, and it was found that the ratio depended upon the amount of sodium vapor in the flame. With the lowest intensity with which it was possible to work, the magnesia reflected four times as powerfully as the vapor, while with an intense flame, giving broad lines, the value rose to 19. To secure a more highly monochromatic source the principle of the resonance lamp (previously described in the case of mercury resonance) was employed. The experiment was made by utilizing the patch of superficial resonance on the sodium bulb as a source of light for exciting the vapor at a different point on the surface of the bulb. The arrangement of the apparatus is shown in Fig. 349.



a *b*

FIG. 349

A small triangular spot of magnesia was formed on the surface of the sodium bulb with a black dot of lampblack at its centre to indicate its position. The image of the sodium flame was thrown upon this spot and the magnesia triangle showed up brilliantly upon the less intense background of the resonance radiation (Fig. 349*a*). The bulb was now rotated until the triangle of magnesia was in darkness, and an image of the spot of resonance radiation thrown back upon it by means of a large concave mirror formed by silvering one surface of a double convex lens (Fig. 349*b*).

[1] *Phil. Mag.*, *27*, 1018, 1914.

Under these conditions the magnesia triangle and the resonance radiation which surrounded it had practically the same intensity.

Bohr's Theory of Resonance Radiation. — While the classical theory of radiation gave a satisfactory explanation of resonance radiation it failed wholly to account for the presence in the emitted light of radiation of longer wave-lengths than those of the exciting light, such as occur in fluorescence and in the case of sodium vapor which emits yellow light when illuminated by the ultra-violet zinc line as we shall see later on. Bohr advanced the theory that the Hg atom in its normal state is able to absorb a quantum of radiant energy corresponding to wave-length 2536, by which the electron is carried from the inner, or 1 $^1S_0$ orbit, to the 2^3P_1 orbit. It is the return of the electron from this orbit to the inner orbit that results in the emission of a quantum of resonance radiation of the same wave-length as that of the light absorbed. The newer theory thus makes the process of emission quite distinct from the process of absorption, whereas on the classical theory the two processes were supposed to occur simultaneously.

Separate Excitation of Centres of Emission of the *D* Lines of Sodium. — The question as to whether both of the *D* lines of sodium would appear in the spectrum of the resonance radiation when the vapor was excited by one *D* line only was investigated by Wood and Dunoyer. The device employed for the removal of D_1 or D_2 from the exciting beam has been described in detail in the Chapter on Polarization (page 349) and was designed by the author for an earlier investigation of the anomalous dispersion of sodium. As used in the present investigation it consisted of a large block of quartz 32 mms. thick, cut parallel to the axis, placed between the two halves of the condenser which focussed the flame on the resonance bulb. Two double-image prisms served as the polarizing system, yielding two images containing only D_2 light superposed, D_1 images of one-half the intensity lying to the right and left. By this expedient the D_2 image had the full intensity of the source except for the loss by reflections from the six surfaces of the prisms and quartz block.

The experiment indicated clearly that vapor emitted only the radiation present in the exciting light, no trace of D_1 being observed when the excitation was by the light of D_2. The experiments were repeated some years later by Wood and Mohler,[1] the result confirming the earlier investigation, but furnishing the additional information that, if hydrogen at a pressure of .025 mm. was present in the bulb, excitation by *one* of the sodium lines gave a resonance radiation which showed both sodium lines. The same

[1] *Phys. Rev.*, *11*, 70, 1918.

effect was found if the bulb was heated to a temperature of 300° C. In other words a transference of energy from one centre of emission to the other could be produced either by the presence of a foreign gas or by working with sodium vapor at considerable density.

No explanation of the effect was possible at the time, but we now know that collisions shift electrons from the level to which they have been raised by the exciting radiation to adjacent levels.

The effects of various gases on the ratio of intensity of D_1 and D_2 when excited by the light of either radiation alone was investigated by Lochte-Holtgraven.[1] The apparatus was similar to that just described except that a very powerful sodium vacuum tube was used in place of a flame. This lamp, designed in collaboration with Cario [2] avoids self-reversal of the D lines by the use of a "gas-window." The form of the tube is shown by Fig. 350. Argon, at a pressure of a few millimetres, is circulated continuously by a mercury vapor pump through the anode bulb, and the sodium vapor is unable to diffuse against this gas current, and consequently is confined to the cathode chamber.

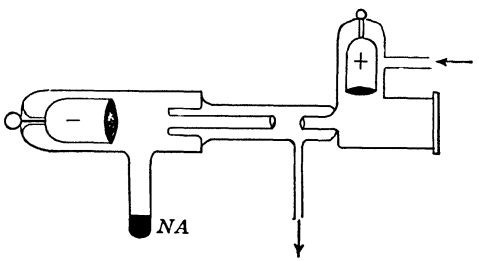


FIG. 350

Exciting the sodium vapor, in argon at 20 mms., by the light of D_1 they observed the appearance of D_2 with an intensity of 1.3 times that of D_1. That the normal ratio of 2 : 1 was not observed was due to the stronger absorption of D_2 during the passage of the resonance light through non-luminous vapor. With nitrogen at 8 mms. and hydrogen at 1 mm. a value of only 0.8 was obtained, these gases quenching the radiation by collisions which return the electrons to the ground level without radiation, the energy being expended in increasing the velocity of rebound.

By increasing the density of the sodium vapor to .002 mm. (by increase of temperature) or by adding potassium vapor at .003 mm. or argon at 0.6 mm. an intensity ratio of 2 : 0.36 was obtained by employing D_2 excitation. The vapors of the alkali metals thus appear to be about 200 times as efficient as argon in transferring the electrons from $2p_1$ to $2p_2$.

The Imprisonment of Resonance Radiation. — It was shown by the author that resonance radiation spreads out into the non-

[1] *Zeit. für Phys.*, *47*, 362.
[2] *Zeit. für Phys.*, *42*, 22. 1927.

illuminated gas mass by repeated atomic absorptions and re-emissions, giving rise to what has been called secondary resonance. The idea was extended by K. T. Compton who suggested that a quantum of radiation was absorbed and reëmitted many times by the mercury vapor before it finally escaped from the boundary of the enclosing vessel. If now the absorption and reëmission occupies a finite time, the illumination of the vapor should persist after the cutting off of the exciting light for a duration of time equal to the sum of the successive absorptions and emissions multiplied by the time taken for each. This should amount to phosphorescence of the vapor, though the mechanism involved is undoubtedly different from that which obtains in the case of most phosphorescent substances. The subject was investigated by Miss Lucy Hayner at the suggestion of Prof. H. W. Webb, who had made some calculations on the rate at which radiation should arrive at the boundary. Miss Hayner employed a hot-cathode mercury arc in a quartz vessel arranged in such a way that it could be struck and cut off very rapidly by a commutator. The light from this arc passed through a quartz cell containing a drop of mercury and highly exhausted, behind which was a rotating disk provided with a number of circular apertures so arranged that the light from the resonance radiation in the quartz cell passed through them into the slit of the spectrograph at the moment when the arc was shut off. The slit of the spectrograph was wide open and arranged in such a way that each hole passed vertically down it forming a trace on the photographic plate the length of which was proportional to the duration of the illumination of the mercury vapor in the quartz cell. Observations were made first on the light of the arc, the quartz cell being absent, and it was found that the illumination here fell off very rapidly after the breaking of the current. With the quartz cell containing mercury vapor interposed, the duration of the illumination was very much longer, and with a cell of double thickness the time was still further increased. The results apparently showed that the resonance radiation fell off exponentially, with an exponential constant that was independent of the vapor density and inversely proportional to the thickness of the layer traversed by the light.

A mathematical investigation of what we may term the diffusion of imprisoned resonance radiation was made by Milne.[1] He considered the vapor enclosed between two infinite planes, one of which was illuminated by monochromatic light, and obtained an expression for the radiation arriving at the other face after the exciting light had been cut off. His results showed that

[1] *Jour. London Math. Soc.*, No. 1, 1926.

this radiation should fall off exponentially with an exponential constant inversely proportional to the square of the number of absorbing atoms per c.c. and the square of the thickness of the slab of vapor. As this was not in agreement with the experimental results found by Miss Hayner the subject was investigated anew with improved apparatus and methods in the same laboratory by M. W. Zemansky.[1] Quartz cells of two different thicknesses were employed and they were used at temperatures varying between 60 and 130°. A somewhat surprising result was obtained, that as the number of absorbing atoms increased the resonance radiation remained imprisoned for a longer and longer time, reaching a maximum value after which any further increase in the number of atoms caused a diminution in the time of imprisonment of the radiation in the vapor, which was attributed to the effect of atomic collisions of the second kind which transform the energy of the excited atoms into heat without the emission of radiation. For low pressures the results were in good agreement with the calculations of Milne.

Step-Up Excitation of Spectra. — As we have seen, mercury vapor in the normal state is able to absorb only wave-lengths corresponding to that of lines originated by electron falls terminating on the innermost or 1^1S_0 orbit. Of all the lines in the mercury spectrum but two originate in this way, namely, the 1849 and 2536 lines, the former produced by a fall from the 2^1P_1 orbit, the latter by a fall from the 2^3P_1, the resonance radiation which we have just studied. The light corresponding to all of the other lines in the mercury spectrum passes through the vapor as readily as through air. This, however, is no longer the case if the mercury vapor is in the process of absorbing the 2536 radiation, for by this process the electrons are raised to the 2^3P_1 orbit, and are thus in a condition to absorb the light of any of the other lines produced by electron falls terminating on this orbit. (See Fig. 352.) By this second absorption process the electrons are carried to higher levels, from which they may fall back in various ways, giving rise to a large number of lines in the emission spectrum of the resonating gas. An excitation of this type, which we may call "step-up," since it results from the successive absorption of light of a number of different frequencies was first observed by Fuchtbauer [2] who, by the illumination of mercury vapor at room temperature in an exhausted quartz tube, surrounded by a mercury arc, showed that the cold vapor radiated the complete arc spectrum. On substituting a glass tube for the one of quartz no radiation appeared, since

[1] *Phys. Rev.*, *29*, 513.
[2] *Phys. Zeit.*, *21*, 635, 1920.

the glass absorbed the 2536 radiation necessary for the initial transition. The experiment did not actually show that the emission resulted from the simultaneous illumination by a number of different frequencies, though no other interpretation was consistent with the Bohr theory. It appeared, at all events, worthwhile to develop the method in such a manner as to render possible the control of the orbital transfers or radiating states of the electrons, and a long and somewhat elaborate series of experiments was carried out by the author [1] in the hope of gathering evidence on the transition probabilities between different levels, the manner in which the transitions were affected by small traces of other gases, and whether or not the transitions actually follow the course predicted for them by the theory of Bohr.

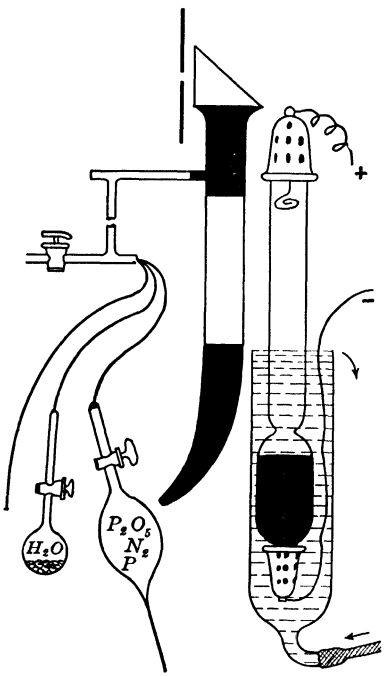


FIG. 351

The apparatus was designed in such a way that the vapor could be illuminated by the 2536 light of the arc in combination with other frequencies selected by suitable filters. It is shown in Fig. 351. The vapor is contained in a vertical tube of fused quartz, drawn off obliquely at the bottom and painted black to secure a dark background, and furnished at the top with a right-angled prism of crystal quartz cemented with piccin. This tube contains a drop of mercury and is in communication with a mercury vapor pump, which operates continually during the experiment. Alongside of, and close to, this tube the quartz-mercury arc is mounted with its negative electrode bulb immersed in running water, to keep the vapor at low pressure and prevent reversal of the 2536 line. A magnetic field is employed as in the experiment with resonance radiation. A small quartz spectrograph is now mounted with its slit (reduced to 1 mm. length) at a distance of 30 cms. from the prism and pointed along the axis of the reflected image of the tube.

[1] *Proc. Roy. Soc., 106*, 679, 1924; *Phil. Mag., 50*, 774, 1925; *Phil. Mag., 4*, 466, 1927.

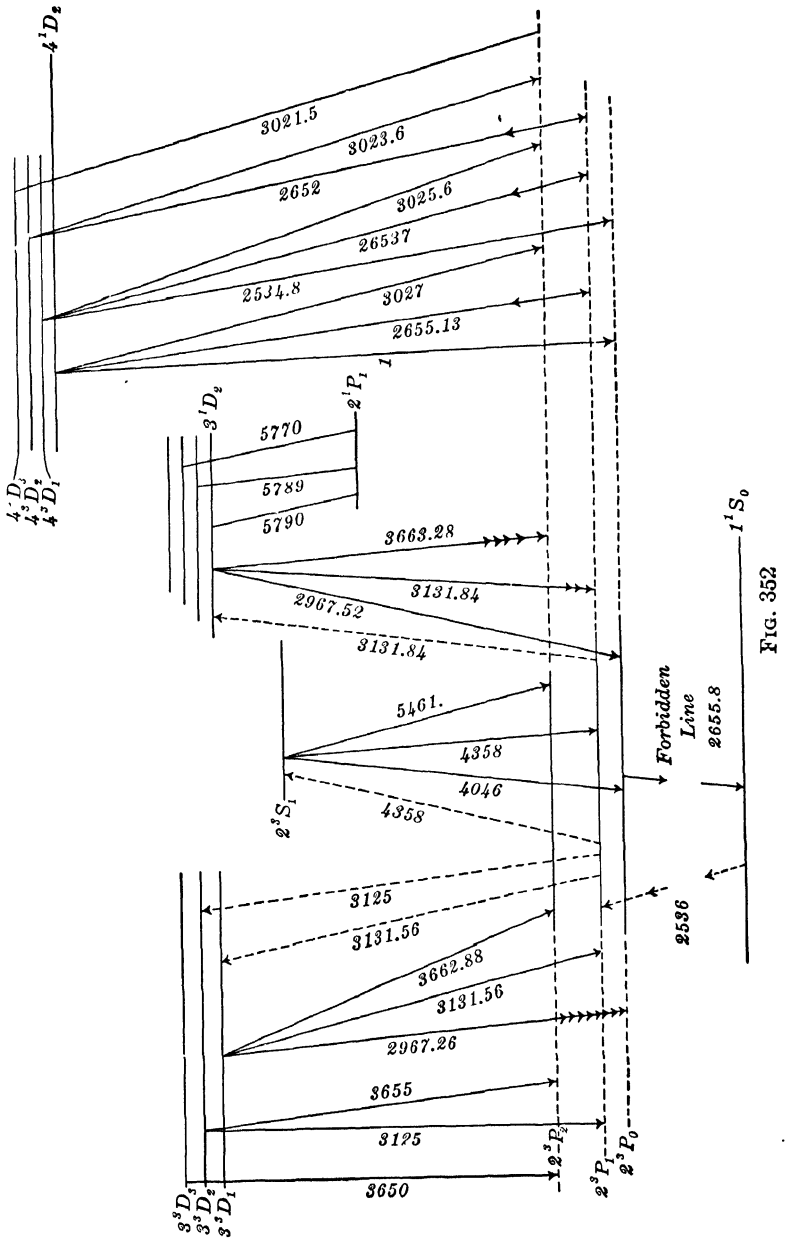

4^1D_2
4^3D_3
4^3D_2
4^3D_1
3021.5
3023.6
2652
3025.6
26537
2534.8
3027
2655.13
I
3^1D_2
2^1P_1
5770
5789
5790
3663.28
3131.84
2967.52
3131.84
2^3S_1
5461.
4358
4046
4358
Forbidden Line
2655.8
2536
1^1S_0
3125
3131.56
3662.88
3131.56
2967.26
3655
3125
3650
3^3D_3
3^3D_2
3^3D_1
2^3P_2
2^3P_1
2^3P_0

FIG. 352

The arc is now started and the circular diaphragm of the spectrograph contracted until it cuts off the light diffused by the walls of the tube, this adjustment being made by bringing the eye up to the focal plane of the spectrograph, when we see at once the glowing green vapor in the tube surrounded by the bright ring of light diffused by the walls. The green luminosity disappears at once if a sheet of glass is interposed between the tube and the lamp for it absorbs the 2536 light, and prevents the excitation of the mercury atoms.

The first experiment was made in the following way: A filter consisting of a large cell of fused quartz 2 cms. thick, filled with bromine vapor at room temperature tension was interposed between the quartz arc and the resonance tube. This filter transmits 2536, 2967, and 3125, 3131, but is opaque to the light of the two violet lines. The luminosity in the resonance tube is now of a pure yellow color, for no visible lines with the exception of the yellow mercury lines are emitted, under this stimulation, as we shall see presently.

A second mercury arc operated under normal conditions is now brought close to the tube. The light of this arc alone produces no radiation of the mercury vapor, since, not being water cooled, its 2536 line is reversed and incapable of bringing the atom into the excited condition. But the vapor is under excitation by the 2536 radiation of the water-cooled arc, and in this condition is able to absorb the violet line 4358 of the second arc, electrons passing out to the 2^3S_1 orbit. Falling back from here 2^3P_2 they emit the green line 5461, as we can show by interposing a sheet of dense cobalt glass between the resonance tube and the high temperature lamp. The green radiation was found to persist, though only the violet light of the lamp entered the tube. We will now follow in detail some of the more interesting cases of radiation with optically controlled orbital transfers of electrons. The energy diagram for mercury is given (Fig. 352) in which the *S.P.D.* levels are arranged in the usual manner and the transitions associated with the line emissions indicated by the oblique lines. It will be remembered from the Chapter on Origin of Spectra that short arrows mean long wave-lengths and *vice versa*. On this diagram absorption processes are represented by dotted, emission by solid arrows. Neglecting at first the upper *D* levels, we see that atoms brought from the ground level to the 2^3P_1 state by the absorption of the 2536 light have five possibilities of behavior. Most of them will revert to the ground level from which they were raised, with emission of resonance radiation, or they may be raised to the 2^3S_1 level by absorption of 4358, or to the *D* levels by absorption of 3131.56−3131.84

and 3125, since excited atoms can absorb all radiations of frequencies represented by downward transitions terminating on the excited level.

From the 2^3S_1 level they may revert to either of the 2^3P levels with emission of the mercury green line 5461, or 4358–4046 in the violet. We have here a very simple type of fluorescence, the emission of green light caused by the absorption of violet light by excited atoms.

The absorption of 3125, 3131.56 and 3131.84 causes the emission of all lines originated by the downward transfers from the levels to which the electrons have been raised by absorption. For example, the yellow line 5790 results from the absorption of 3125. It appears also that the lines 3650 and 3021.5 are emitted only as a result of absorption of their own frequencies present in the exciting light by atoms in the 2^3P_2 state, involving three successive absorptions instead of two, as in the case of the other lines. This results in the curious circumstance that their intensities will vary with the cube of the intensity of the exciting light, while the other lines vary as the square, as was pointed out by the author in one of the earlier papers and more thoroughly investigated in collaboration with E. Gaviola. This matter will be fully discussed presently.

Optically Controlled Radiating States. — It is clear that we have here a very perfect method of proving that transitions take place in the manner indicated, for by placing filters between the lamp and the tube we can restrict the number of exciting frequencies and observe the resulting emission, a more precise method than the electrical one, in which the impressed voltage is gradually raised.

A large number of combinations were tried, of which only a few can be cited as typical cases. Using, between the lamp and resonance tube, a filter of bromine which cuts off 4046 and 4358 and reduces 3650 to one-third of its value, it was found that the color of the light emitted by the mercury vapor changed from green to yellow and the spectroscope showed that the 3650 line was absent while 3654 and 3663, as well as both yellow lines were strong. The absence of 3650 is due in part to the great reduction in the number of atoms in the 2^3P_2 state, by removal of the 5461 emission, and in part to the weakening of 3650 in the exciting radiation. The yellow lines and the two 3663 lines result from the absorption of 3131.84, 3131.56 and 3125 which are transmitted by the bromine.

With a filter made of a dilute solution of salicylic acid flowing continuously through a quartz cell (to avoid photochemical change) the lines 3125 and 3131 are removed from the exciting light, and the ultra-violet triplet 3650–54–63 now appears as a

result of absorption of these same frequencies by atoms in the 2^3P_2, state caused by absorption of 4358 and the resulting emission of 5461. In this case the intensities of the three lines are as 5, 1, 3, about their value in the arc which is what we should expect. In this latter case we have emission as a result of three successive absorptions, while in the former case two absorptions only are involved. This difference in the manner of excitation results in the interesting variation in the relative intensities of the lines with the actual intensity of the exciting lamp, which we will now consider.

Power Relation of Intensities. — That the intensity of certain lines in the optically excited spectrum would increase with the square, and of others with the third power of the intensity of the exciting lamp, was shown by the author in one of the earlier papers. In the case of emission involving one absorption process only (resonance radiation for example) the intensity varies as the first power: Where two absorptions are involved, as in the case of the emission of 5461 the intensity varies as the square of the intensity of the exciting light, since by doubling the intensity of the latter we excite double the number of atoms and consequently have double the number capable of absorbing 4358, which itself has been doubled in intensity giving us four times as many atoms in condition to emit the green line. An analogous case which is easier perhaps to visualize would be the following. There are cases in which light, by passing through a vapor, causes a chemical change which results in the precipitation of a fog of scattering particles. If we double the intensity of the light, we obtain a fog twice as dense illuminated by a light twice as bright, hence four times brighter than formerly.

For lines which are emitted as a result of three successive absorptions, say 3650, we have an eightfold increment of intensity, the variation being as the cube. This relation was discovered qualitatively in the earlier work as a result of the observation that the relative intensities of the lines were greatly altered by applying a weak magnetic field to the arc, which, as we have seen, amounts to increasing the intensity of all lines weakened by self-reversal, which was found to be the case for all of the lines involved.

A quantitative study of the relation was made [1] which fully established the power relation. It is obvious that a rotating sector cannot be used for reducing the intensity of the illumination, since the momentary flashes will be of full intensity, and moving the lamp away from the tube is equally undesirable, owing to the uncertainty about the change of illumination in the case a long narrow source in close proximity to the tube. A wire gauze is,

[1] Wood and Gaviola, *Phil. Mag.*, August, 1928.

however, non-selective, and admirably adapted to the purpose, as mentioned by the author in an earlier paper. A gauze of fine copper wire, which reduced the light to $\frac{1}{5}$ of its normal value, when inserted between the lamp and mercury tube, reduced the intensity of the visible radiations to $\frac{1}{25}$ and the 3650 and 3021 lines to $\frac{1}{125}$ of their values before the insertion of the gauze. Holding the gauze between the eye and the mercury tube we see without difficulty the green light, while with the same gauze inserted between the lamp and tube nothing is to be seen, owing to the 25-fold reduction of intensity.

Formation of Metastable Atoms by Nitrogen. — Remarkable modifications in the orbital transfers of the electrons were obtained by the introduction of other gases into the resonance tube, the most remarkable result being that found with nitrogen which at a pressure of 2 millimetres caused such an immense increase in the intensity of the light that a green glow was visible laterally through the walls of the tube in spite of the proximity of the mercury arc. It was found that some lines were not enhanced at all by the nitrogen, while the increase in intensity of the green light was as great as thirtyfold. It was of the utmost importance to ascertain, if possible, the action of the nitrogen in producing this enormous increase in the radiation of certain frequencies. The hypothesis was made that it produced an increase in the absorption of the light, and the hypothesis was tested in the following way:

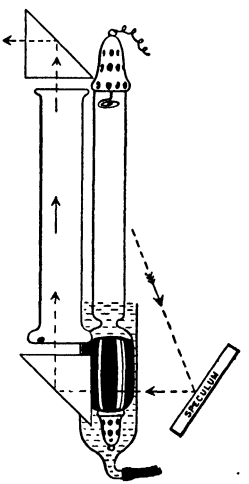


FIG. 353

A quartz tube ten centimetres long, closed with flat end plates fused on, was mounted as close as possible to the water-cooled mercury arc (Fig. 353). Two right-angle prisms of quartz were arranged as shown in the figure, and a small plane mirror of speculum metal mounted at such an angle as to reflect the light of the arc into the lower prism, from which it eventually reached the spectroscope after traversing the vertical tube. The tube contained a drop of mercury and was connected to the pump and nitrogen bulb (not shown in the figure). By interposing a sheet of glass between the lamp and the absorption tube, the mercury vapor could be changed from the excited to the normal condition. A series of spectra of the mercury arc light after transmission through the absorption tube were made with different exposure times,

both with and without the glass plate. The most casual examination showed that the 4046 line was very powerfully absorbed by excited mercury vapor in 2 mms. of nitrogen. The same was true of the lines 2752 and 2967, though none of these lines should be absorbed at all according to the diagram of orbits.

Comparison of the spectra showed that the arc lines had been reduced to less than 25% of their normal value by the passage of the light through a layer of excited mercury vapor at room temperature 6 cms. in thickness in an atmosphere of nitrogen at 2 mms. pressure.

The absorption was next investigated with a very fine Lummer-Gehrcke plate of quartz. Even with the high resolving power of this instrument no trace of reversal could be detected on the spectrograms for any of the lines except 2536 in the case of mercury in vacuo, but the presence of a fraction of a millimetre of nitrogen caused the reversal of the 4046 line, while with nitrogen at 2 mms. the reversal was so strong as nearly to obliterate the line. Collisions of excited mercury atoms with nitrogen molecules brings about a transfer of electrons from the 2^3P_1 to the 2^3P_0 orbit in the case of a large number of atoms. This orbit as well as 2^3P_2 is metastable, that is the electrons are unable to fall back to the 1^1S_0 orbit, a "forbidden" transition in consequence of which all of them are available for absorption.

Measurements showed that the introduction of nitrogen enhanced the intensity of 5461 about twentyfold while 4358 and 4046 were enhanced about tenfold, and fourfold respectively. This was very puzzling at first as the three lines are formed by transitions from the same upper level. It was at first suspected that certain transitions were favored by the presence of the nitrogen, but as this seemed unlikely a search was made for the cause. It seemed likely that absorption of the light coming from the bottom of the tube by the vapor in the upper part, which would be very strong for 4046 less for 4358 and zero for 5461, was responsible for the smaller enhancement of the violet lines, as no contribution would be made by the lower layers while the green radiations would come from the entire column. This was tested by limiting the illumination to a narrow sheet of light passed by a horizontal slit mounted between the lamp and the tube, in which case the enhancement of all three lines was the same.

As an interesting case of the effects of the introduction of nitrogen the following may be cited: It was found that the yellow lines had the same intensity with mercury *in vacuo,* observed visually by widening the slit of the spectroscope until the two lines just made contact.

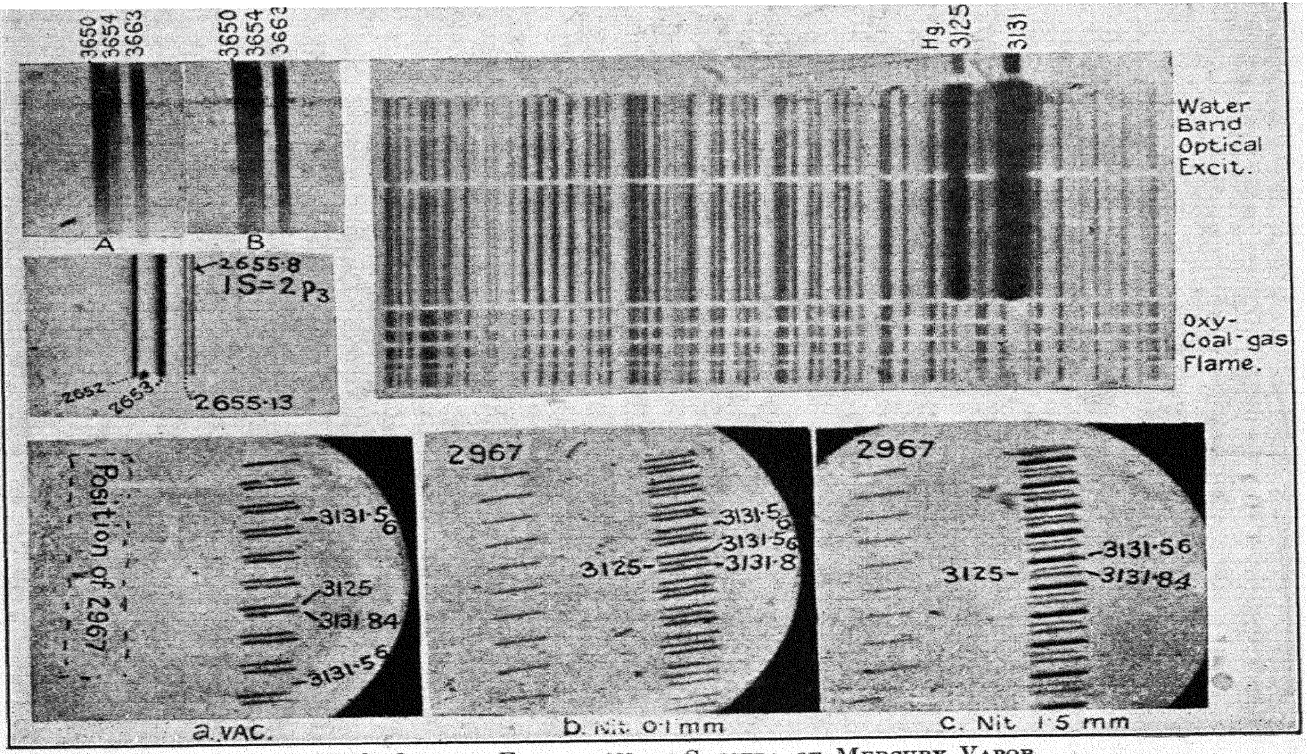


PLATE 9. OPTICAL EXCITATION OF SPECTRA OF MERCURY VAPOR.

With nitrogen present 5790 was much stronger than 5770, and the intensity difference was still more marked with the bromine filter. The most probable explanation of this effect is that collisions with nitrogen transfer electrons from 3^3D_1 (to which level they have been brought by 2967.28 absorption) to the lower 3^1D_2 level, which would weaken 5770 and strengthen 5790. The increased enhancement of 5790 with the bromine filter may be explained as the result of a greater number of atoms in the metastable state due to the removal of the 4046 absorption, which ordinarily raises them to a higher state. There are thus more available for the absorption of 2967.28. Plate 9, Fig. *a*, the 3650 group *in vacuo*, Fig. *b* in N_2.

On Plate 9, Fig. *c* the increment in intensity of certain members of the 3131 group of lines by the addition of nitrogen is strikingly shown. These photographs were made with a Lummer-Gehrcke plate. The reader will find it of interest perhaps to account for these changes by the energy diagram, Fig. 352.

The Forbidden Line 2655.8. — A transition from the 2^3P_0 level to the ground level 1^1S_0 would cause the emission of a line of wavelength 2655.8, but this line is forbidden by the exchange principle, and it never appears in the spectrum of mercury under ordinary conditions. Some anomalies in the intensity ratios of the lines in the close group 2652, 2654, 2655 caused by the presence of nitrogen (2655 abnormally bright) which could not be explained by the energy diagram were finally found to be due to the fact that this "forbidden" line appeared in the optically excited spectrum of mercury in nitrogen. It was clearly separated from its close neighbor 2655 by employing a more powerful quartz spectrograph than the one used in the earlier work. Its appearance is probably due to the great concentration of atoms with electrons on the lower metastable level, giving a much greater probability for this improbable transition to occur. ("Improbable" being more a correct term than "forbidden.") The line is shown on Plate 9, Fig. *b*.

Resonance Radiation of Electrically Excited Atoms. — As we have seen, the optical excitation of a gas or vapor, by raising electrons from the lower to higher levels, makes possible the absorption of other frequencies for which the gas is normally perfectly transparent. The same thing can be accomplished by subjecting the gas to an electrical discharge, in which case the atoms are raised to higher energy states by electron impact.

The discovery of the absorption of light by an electrically excited gas dates back to the work of Pflüger in 1907 and Ladenburg and Loria in 1908. These investigators obtained the reversal of

the bright hydrogen line by passing the light of a powerfully excited hydrogen tube through a second tube containing hydrogen under feeble electrical excitation, but the nature of the processes involved was but imperfectly understood.

The first case of resonance radiation by an electrically excited gas under optical excitation was made by Paschen in 1914 in the course of an investigation of the properties of helium gas in the infra-red region.[1] He found that helium gas at low pressure excited by a current of small intensity absorbed powerfully the light corresponding to that of the two infra-red lines of wave-lengths 10,830 and 20,582.

Helium has the remarkable property of existing in two systems of excited states between which transitions never occur, a peculiarity which lead spectroscopists to believe for a long time that there were two kinds of helium, which were named parhelium and orthohelium, one giving a system of triplets and the other doublets. Helium in the ground state 1^1S_0 can be excited electrically to the 2^3S_1 state (a metastable level as with mercury) constituting the ground level of orthohelium. This transition corresponds to light of wave-lengths never observed, however, owing to the metastability of the level. It has never been produced by absorption, but the atom in this electrically excited state is now able to function in the same manner as the optically excited mercury which we have just studied, and give rise to resonance radiation by the absorption of two infra-red frequencies of wave-length 10,830 and 10,829 as shown in Fig. 354. The excited state 2^1S_0, the ground level of parhelium, is also metastable and atoms in this state are capable of absorbing 20,582 and emitting it as resonance radiation. This was actually accomplished by Paschen in spite of the difficulty that all observations had to be made with the thermo-electric cell.

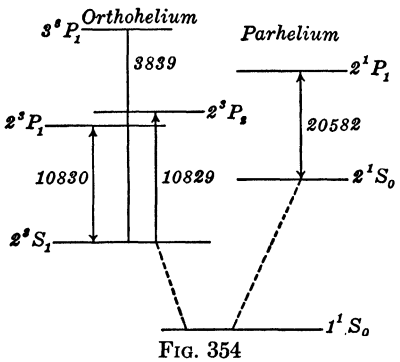


FIG. 354

Pure helium gas, when carrying a weak current of electricity of density equal to about 10^{-7} amperes per square centimetre, absorbs powerfully these two radiations, though showing only a very small emission either of these or the visible lines. Similar absorption for the light of the visible lines occurs only when the

[1] *Ann. der Phys.*, *45*, 625.

density of the current is increased to a value ten thousand times as great as that mentioned above. Paschen's experiments showed that practically all of the energy abstracted from the primary beam was reëmitted as resonance radiation. As we have seen a similar conclusion was reached by Wood and Dunoyer in the case of sodium vapor. In the opinion of the author these experiments stand out as perhaps the most remarkable ever performed in the field of radiation, on account of the difficulties which stood in the way. More recently McCurdy has obtained the resonance radiation of the 3839 line with excited helium [1] the transition involved being shown in Fig. 354.

Resonance Radiation of Neon. — One of the most beautiful cases of resonance radiation which we have is that exhibited by electrically excited neon gas. It was discovered and studied by de Groot in 1926, and can be shown with one of the commercial neon glow lamps manufactured by the Philips Lamp Company of Eindhoven, Holland. These lamps are of two types, one operating on 110 volts containing a small trace of argon; the other, operating on 220 volts, is filled with pure neon. The latter alone shows the resonance radiation.

If we examine one of the 220-volt lamps we shall find that the region inside of the spiral (which glows with the yellow negative glow of neon) is filled with a pure crimson luminosity quite different from the color of the negative glow. This luminosity de Groot attributed to resonance radiation excited by the light coming from the glow around the filament. To test this he focussed the light of a powerfully illuminated neon tube by means of a large condensing lens at a point within the neon lamp just outside of the spiral electrode, where the luminosity of the gas was very low. The path of the exciting light within the bulb was marked by a brilliant crimson luminosity, the spectrum of which was found to be identical with that of the crimson glow inside of the spiral. This spectrum is characterized by the relative greater intensity of the lines corresponding to electron falls on the metastable orbits, the bright yellow neon line being notably weak in the spectrum of the resonance radiation.

The author has observed the resonance radiation in the neon lamp by means of concentrated sunlight. The experiment should be performed in a dark room, the bulb being viewed against a black background.

The following simple method of showing absorption by excited neon will be found useful for demonstration purposes. The light of a small end-on neon tube is rendered parallel by a lens and

[1] *Phil. Mag.*, *2*, 259, 1926.

passed through the region occupied by the spiral electrodes of a Philips neon lamp, or some similar lamp and then focussed by a lens (placed close to the lamp) on the slit of a spectroscope. It will then be found that certain neon lines are weakened when the neon lamp is turned on. By placing the lens close to the lamp, no image of the latter is formed on the slit, consequently its contribution is small.

Duration of the Metastable States. — The investigations of K. W. Meissner on the absorption in excited neon clearly showed that the S_3 and S_5 states of neon had a longer duration than the S_2 and S_4 states. His work was repeated and extended by Dorgelo,[1] who ascertained by experiment how long after the cessation of the discharge which brought the neon atoms into the excited state the absorption of the light of a second neon tube could be observed. His apparatus consisted of an iron disk perforated by an aperture, and mounted on the shaft of an electric motor. The disk was mounted between the emission and absorption tube, and served also as a commutator for switching momentary flashes of current into the tubes, the time interval between the flashes being determined by the distance between two contact points. It was found possible to choose this distance so that with a definite velocity of revolution of the disk the light of the emission tube was still absorbed by the atoms which remained in the state of excitation, whereas no trace of absorption was perceived with smaller times of revolution. With the neon line $\lambda 6402$ this time was found to be about $\frac{1}{230}$ of a second, which may be considered the time during which an electron remains on the S_5 orbit after the cessation of the electrical discharge. For the S_3 orbit the time was much shorter, namely $\frac{1}{2000}$ of a second. A very pronounced temperature effect was found, the duration at 174° being $\frac{1}{2000}$ sec. at room temperature $\frac{1}{250}$ while at −196° it increased to $\frac{1}{10}$ sec.

Fluorescence of Atomic Vapors. — Fluorescence is a term applied by Stokes to the emission by certain substances when illuminated, of a radiation having wave-lengths greater than that of the light falling on them. We have the simplest type in the case of certain monatomic metal vapors. An attempt to excite the D lines by the radiations from a quartz-zinc arc, which emits a line 3303 in coincidence with the second member of the principal series of sodium in the ultra-violet, was made by the author [2] but no trace of the D lines was observed, though a fluorescent band spectrum was photographed. The vapor of sodium was formed in the large steel tube used for the study of the resonance fluorescence,

[1] *Zeit. für Phys.*, *34*, 766, 1925.
[2] *Phil. Mag.*, Oct., 1909.

which will be described in the next chapter, which accounts for the failure to observe the phenomenon, for any emission of yellow light in the fluorescence chamber at the centre of the tube would be absorbed by the less dense vapor in the cooler part of the tube. The technique of employing glass bulbs, developed by Dunoyer, came later and in 1915, R. J. Strutt now Lord Rayleigh, observed the emission of both D lines in a bulb containing sodium vapor, illuminated by the zinc line 3303, a very beautiful case of fluorescence, explainable on the Bohr theory as resulting from the fall of electrons from the upper level, to which they have been raised by the absorption of 3303, to the next lower level, from which they then return to the ground-level with the emission of the D lines. The intermediate transition, if accompanied by radiation, should give a red line, but Strutt found no trace of this.

It was shown by Franck and Cario [1] in their paper on sensitized fluorescence, which will be treated presently, that the failure of the red line in the emission spectrum was due to collisions of the second kind, in which case the energy liberated by the passage of an electron from an upper to a lower level is spent in increasing the velocity of translation of the rebounding atoms, in other words it is transformed into heat. They found that the emission of the D lines by sodium vapor excited by the 3303 line of zinc is much more copious when argon at a pressure of two or three millimetres is introduced into the sodium bulb: this shows that collisions are responsible for the transfer of the electrons from the $3p_1$ $3p_2$ levels to the $2p_1$ $2p_2$ levels from which they fall to 1, 5 s with the emission of the D lines. The difference of energy between the $3p_1$ and $2p_1$ levels corresponds to 1.5 volts, an amount sufficient to give the rebounding sodium atom a velocity considerably above the mean velocity of the other atoms. We should therefore have a Döppler effect in the case of the light emitted in this way, and Cario and Franck found that this was the case for the yellow light emitted by the bulb was found to pass readily through a second bulb containing sodium vapor, whereas resonance radiation excited by the D lines was completely absorbed by the second bulb. Another very interesting case was discovered in 1926 by Terenin which settled the much disputed question about the absorption of the thallium green line by thallium vapor.

A few of the energy levels of this element are shown in Fig. 355, with one Hg level upward arrows indicating absorption, downward arrows emission. (Old notation is used for convenience.) In the lowest energy state the vapor is capable of absorbing 2768 and 3776, the electrons being carried to the two upper levels, from each

[1] *Franck and Cario, Zeit. für Phys., 17*, 202, 1923.

of which they may return to either of the lower levels with the emission of radiation of longer wave-lengths 3526 and 5350, the simplest case of fluorescence known. Pringsheim and Orthmann found, however, that 5350 was absorbed by the vapor when its temperature was raised to 800, the electrons being brought to the $2p_1$ level by the thermic energy. An emission of 5350 and 3776 resulted, violating Stokes's law.

Analogous results were also found by Terenin with lead, bismuth and antimony, his observations aiding in establishing the energy levels of these elements, which were imperfectly understood at the time. A spectrum of the exciting arc was projected on the quartz tube containing the metal vapor by means of a quartz monochromator with a horizontal slit. The vapor was thus excited by the horizontal bands of monochromatic light incident on the tube, and an image of these bands was projected on the vertical slit of a quartz spectrograph which thus recorded, as superposed spectra, the fluorescent spectra excited by the various lines of the arc.

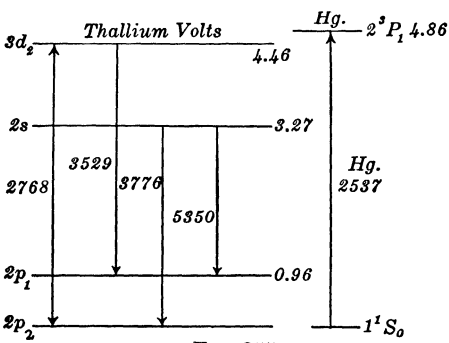


FIG. 355

Sensitized Fluorescence. — The very interesting observation was made by Cario [1] that a mixture of mercury vapor and the vapor of thallium, in a quartz tube, when illuminated by the 2536 mercury line, radiated a light which the spectroscope showed contained the radiation peculiar to thallium vapor. The mercury atoms, thrown into the excited state by the absorption of the 2536 radiation, on colliding with the thallium atoms, gave up their absorbed energy throwing the thallium atoms into the excited state and causing them to radiate their characteristic lines. This type of radiation was named "sensitized fluorescence" since thallium vapor by itself is unaffected by the light, the presence of the mercury vapor rendering it sensitive to the radiations of the mercury arc.

If silver was used in place of the thallium the characteristic arc lines of silver could be obtained in the same manner. The condition essential to this type of radiation by vapors at comparatively low temperature is that the quantum of energy absorbed by the sensitizing vapor shall be greater than the quantum of energy to be

[1] *Zeit. für Phys.*, *10*, 185, 1922.

radiated by the other vapor. This means that the lines radiated will have wave-lengths longer than that of the absorption line of the sensitizing vapor. Exceptions may occur in the case of vapors at high temperatures.

A very complete investigation of the phenomenon was immediately carried out by Cario and Franck [1] who stated the problem as follows.

Suppose that we have a mixture of two gases, gas A of atomic weight m, of which the frequency of the longest λ absorption line is ν and gas B of atomic weight m_1, and longest λ absorbed of frequency ν_1. The energies of excitation in the two cases are $h\nu$ and $h\nu_1$ in the ratio of $1 : \frac{3}{4}$ (say) as shown by Fig. 356. If we radiate the mixture with frequency ν then both ν and ν_1 frequencies appear in the emitted light, the gas atom A, excited by the frequency ν (absorbing energy $h\nu$) may either reëmit this frequency or it may, on collision with atom B give up $\frac{3}{4}\,h\nu$ $(=h\nu_1)$ and excite B, the other $\frac{1}{4}\,h\nu$ of energy being spent in increasing the velocity of translation of the rebounding atoms. If the temperature of the gas mixture is so low that the kinetic energy of the heat motion of the atoms is small in comparison to $h\nu$, the atom B receives in addition to its energy of excitation a kinetic energy

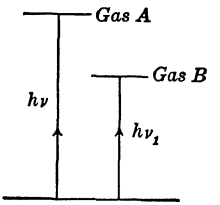


FIG. 356

$$\frac{1}{2}m^1V^2=\frac{h\nu}{4}\cdot\frac{1}{1+\frac{m_1}{m_2}}$$

and a Döppler effect frequency ν_2 represented by

$$\nu^2=\nu_1\left(1+\frac{V_1}{C}\cos\Phi\right).$$

In other words, the emitted line is broadened and the radiation will not be absorbed by the unexcited B atoms which may be regarded as practically at rest, consequently lines excited in this way will be relatively intense with respect to lines excited in such a way that no Döppler effect results.

If, however, we radiate the gas mixture with frequency ν_1, an emission of ν can result only if the gas temperature is so high that the failing $h\nu/4$ is supplied on collision. Both effects were observed.

The arrangement of Cario's apparatus is shown in Fig. 357. The quartz resonance bulb was mounted in an electric furnace

[1] *Zeit. fur Phys.*, *17*, 202, 1923.

which could be raised to a temperature of 800°. Two lateral tubes, one containing thallium and the other mercury were heated by smaller furnaces to any desired temperature, the densities of the two vapors corresponding to the temperatures of the two furnaces, and the mixed vapors could be superheated in the larger furnace if desired. The vapors were excited by the radiations of a water-cooled quartz-mercury arc, the exciting radiation being the 2536 line with an energy represented by $h\nu = 4.9$ volts. The mercury tube was heated to 100°, the thallium tube to 750°, while the larger furnace was kept at 800°. Under these conditions thallium lines appeared which required 5.5 volts for their excitation, 2709 for

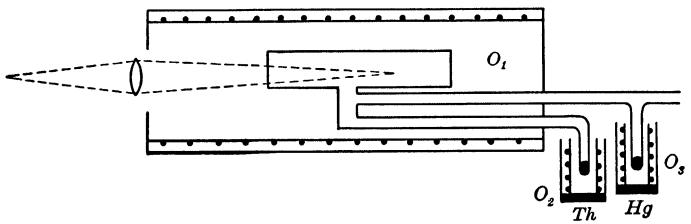


FIG. 357

example $2p_1 - 5d$ (see Fig. 355), so that thermic energy must be considered as coöperating with the energy of the excited mercury atoms to bring about a transfer of the electron from the ground level $2p_2$ to $5d$. This does not prove, however, that the two energy transfers take place simultaneously, for the transfer from $2p_2$ to $5d$ may have occurred in two steps. The energy of the thermic motion in thallium vapor at 800° is sufficient to raise electrons from the $2p_2$ to the $2p_1$ level, as is shown by Grotrian's observation that the vapor at even 600° will absorb the light of the green line 5350, which terminates on the $2p_1$ level, which differs from the $2p_2$ level by one volt. To pass from $2p_1$ to $5d$ requires less than 4.9 volts, the amount of energy which can be supplied by the excited mercury atom.

That the two processes do occur simultaneously, however, can be inferred from the intensity ratios of the lines. The line 3776 ($2p_2 - 2s$) is very intense while 2768 ($2p_2 - 3d_2$) is very weak. Normally both are strongly absorbed by unexcited thallium vapor, since they terminate on the ground level $2p_2$. That the $3d_2$ level is reached as often as $2s$ by energy transfers from excited mercury atoms is shown by the presence of the very strong 3530 line ($2p_1 - 3d_2$). The weakness of the 2768 line results from absorption, since it takes but 4.6 volts of energy from the mercury atom, the remaining energy $4.9 - 4.6 = 0.3$ volt not being sufficient to produce an

appreciable Döppler effect. The line 3776 requires but 3.6 volts for its excitation, leaving 1.3 volts for the Döppler effect, consequently its wave-length is altered and it is therefore not absorbed by the vapor: this accounts for its greater intensity.

Indium has a spectrum similar to that of thallium, the corresponding lines being shifted towards the red, however. This means that they require less energy for their excitation. The line 3039, which corresponds to 2768 of thallium, is in this case strong in the spectrum, as it is broadened by the Döppler principle, and hence not absorbed by the unexcited indium vapor.

Similar results were obtained with a mixture of cadmium and mercury vapors, the triplet at 3261 appearing only at high temperatures, with no trace of it at 400°. To excite the triplet requires 6.3 volts, that is 1.4 volts must be supplied by the thermic motions.

An interesting case of sensitized fluorescence was found by the author in the study of the "step-up" excitation of mercury, the so-called water-band appearing in the spectrum. It is shown on Plate 9, Fig. *d*, in coincidence with the spectrum of the oxy-coal-gas flame which shows the same band.

Effects of Foreign Gas on Sensitized Fluorescence. — The very interesting observation was made by Donat[1] that the addition of argon or nitrogen to the mixture of mercury and thallium vapors increased the intensity of the thallium lines and decreased the intensity of the mercury resonance radiation. He attributed this to a transfer of electrons from the 2^3P_1 level of mercury to the metastable 2^3P_0 level, atoms in this state being able to excite thallium, but unable to radiate directly.

The concentration of atoms in the 2^3P_2 state produced by addition of nitrogen was shown experimentally by the author as stated in the section on "step-up" excitation of mercury.

[1] *Zeit. für Phys.*, *29*, 345, 1924.

CHAPTER XIX

THE RESONANCE AND FLUORESCENCE SPECTRA OF MOLECULES

Absorption Spectrum of Sodium Molecules. — The relatively simple type of optical excitation which we have considered in the case of monatomic gases and vapors becomes more complicated in the case of diatomic molecules. Curiously enough we can begin the study of this subject with the same substances which we have already investigated, namely, the vapors of metallic sodium and mercury, which under certain conditions form diatomic molecules. Wiedemann and Schmidt discovered many years ago that the vapor of sodium obtained by heating the metal in exhausted glass bulbs to a temperature of 250°, when illuminated by an intense beam of white light, emitted light of a brilliant green color quite different from the yellow resonance radiation which we have already studied. We now know that this fluorescence or green radiation is not produced by the sodium atom, but by molecules,

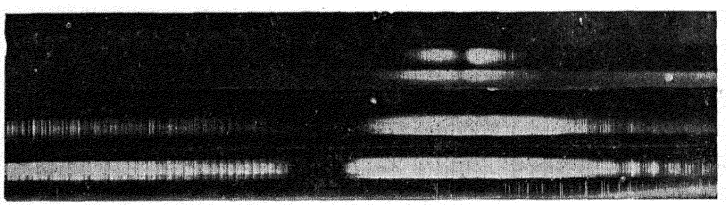


FIG. 358

Na_2, formed in small numbers by the union of two sodium atoms, in rather dense vapor. The absorption spectrum which is responsible for this fluorescence can be seen by heating a small amount of the metal in a long steel tube, the ends of which are closed by plate glass windows cemented with sealing wax. The tube is well exhausted with an oil pump and then heated, the exhaustion continuing until most of the hydrogen has boiled off, when the stopcork between the tube and the pump is closed. Under these conditions the absorption spectrum is found to consist of channelled or fluted bands, one group appearing in the red and extending down into the yellow region considerably beyond the *D* lines, the other

in the blue-green region extending up well into the yellow-green and almost meeting the orange-red band. Photographs of this spectrum are reproduced in Fig. 358. Spectrum *d* taken with the tube at a temperature of about 350 and spectrum *a* with the tube red hot. It is seen that as the temperature and density increase the red-orange portion spreads over and below the *D* lines while the blue-green one pushes up in the other direction, the two eventually meeting at about wave-length 5500, where a broad hazy absorption band appears, which is just appearing in spectrum *b* and is very distinct in *a* taken with very dense vapor. With a spectroscope of low dispersion we see merely a hazy double green line at wave-length 5500 and the extreme violet end of the spectrum, everything else being cut off. Colored frontispiece, Fig. 11.

This band spectrum was photographed by the author in 1900 with a 10-foot concave grating [1] and later, in collaboration with Clinkscales, in the second order spectrum of a 21-foot instrument.

The spectrum showed on the average from 60 to 70 absorption lines within a space only 12 Ångström units in width, in other words, as many as 30 lines in a region not wider than the distance between the *D* lines. This means that in the blue-green channelled absorption spectrum, which is about 1200 Ångström units in width, there are roughly speaking about 6000 absorption lines. Taken collectively, these lines form themselves into a number of groups or bands, which resemble the groups seen in the absorption spectra of iodine and bromine and in certain banded emission spectra, such as the spectrum of nitrogen. If the vapor is very dense the absorption lines become so numerous and close together that they block out completely certain regions of the spectrum, these dark regions, however, being separated by bright lines, quite as narrow as the emission lines of the iron arc, the spectrum resembling an emission rather than an absorption spectrum.

The spectrum of the yellowish-green fluorescent light which the vapor emits when excited by white light, as a result of the absorption which we have just considered, was photographed by Wood and Moore in 1903 [2] and appeared to be very nearly complementary to the absorption spectrum; that is, the dark bands of one replaced by bright bands in the other. The work was carried on by the author during the following year, the apparatus being gradually improved and the dispersion of the spectroscope increased.[3] Very remarkable relations between the nature of the

[1] "Anomalous Dispersion of Sodium Vapor," *Proc. Roy. Soc.*, *69*, 169, 1901.
[2] *Astrophysical Journal*, July, 1903.
[3] *Phil. Mag.*, Nov., 1905, May, 1908.

exciting light and the fluorescent light were found, which we will now consider.

The apparatus in its final form consisted of a seamless tube of thin steel 3 inches in diameter and 30 inches long, with a steel retort at its centre in which a large amount of sodium could be stored. The retort was made by fitting two circular disks of steel to a short piece of tubing, just large enough to slip snugly into the larger tube. The circular ends of the retort were provided with oval apertures as shown in Fig. 359*a*. The retort was half filled with sodium, the molten metal being poured in through one of the apertures. It was then introduced into the tube and pushed down to the centre, after which the plate glass windows were cemented to the ends of the tube with sealing wax, as shown in the figure. This arrangement prevented the rapid diffusion of the vapor, and enabled a large supply of metal to be kept at the centre of the tube. The tubes used in the earlier work required recharging after two hours' continuous operation, while the retort tube could be operated for several hundred hours on a single charge. The ends are kept cool by wet cotton or better by coils of lead tubing in which water circulates.

If a steel tube cannot be procured, a large brass tube will answer the purpose. The retort is not necessary except for long-continued photographic study, a lump of sodium being placed at the centre of the tube. The conical beam of light can be thrown in in an oblique direction, and the fluorescence observed through the opposite end of the tube, or by means of a mirror. The spectrum excited by white light is shown on Plate 10, Fig. 1.

After exhausting the tube to a pressure of a millimetre or two a Bunsen burner is placed under the retort and a cone of sun or arc light focussed on the oval aperture by means of a lens of long focus placed to one side of the tube, and a small piece of mirror-glass. The arrangement of the apparatus used in the earlier work for exciting the vapor with monochromatic light is shown in Fig. 359 as soon as the sodium vapor begins to form a brilliant spot of green fluorescent light at the aperture of the retort. The spectrum of the light is made up of an enormous number of fine lines, which in the yellow and yellow-green regions are arranged in groups or bands, which lie close together in the vicinity of the D lines, widening, however, as the blue region is approached. This spectrum is shown in Fig. 360*a*. Coincident with the D lines there appears a hazy band (the surrounding region being nearly devoid of light), which, if the vapor is not too dense, can be resolved into a double line, the components of which coincide with D_1 and D_2. These lines appear, not only when the vapor is stimu-

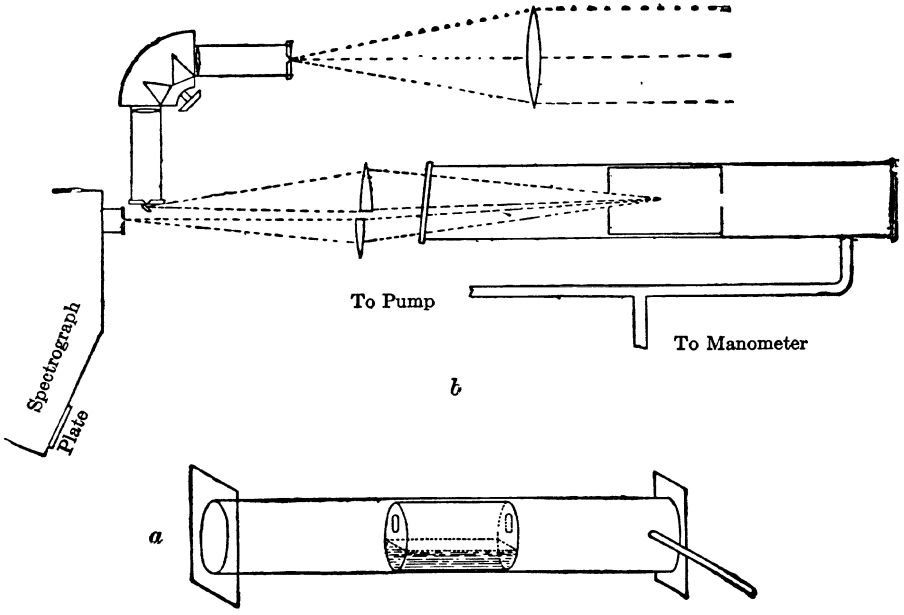


FIG. 359

lated with light of the wave-length of the sodium lines, but also when a powerful beam of blue-green light is thrown into the vapor. This showed that the mechanism which gives rise to the *D* lines is connected in some way with the one which produces the complicated channelled absorption spectrum. The nature of this mechanism will be explained later on.

The rest of the fluorescent spectrum behaves in a very different manner. Stimulating the vapor with light of a deep violet color from the spectroscope produces no effect; as the wave-length is

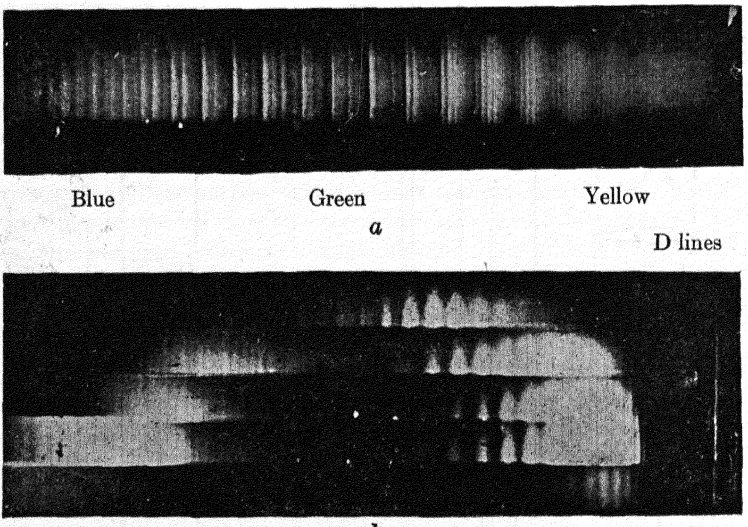


FIG. 360

gradually increased a yellowish fluorescence appears, which spectroscopic examination shows to be made up of two parts, a reëmission of the same wave-lengths as those absorbed (blue) and the extreme yellow end of the fluorescent spectrum, comprised between wave-length 571 and 505. This portion consists of bands, which a high resolving power resolves into fine lines. As the wave-length of the exciting light is further increased, the point of maximum fluorescence moves down the spectrum, the first bands or groups of lines disappearing. In other words, as the exciting light moves up, the fluorescent light moves down the spectrum. There is, however, in all cases an emission of light of the same wave-length as that of the exciting light. These relations will be better understood by referring to Fig. 360*b*, where a number of photographed spectra are shown one above the other. The region of the spectrum

excited is represented by the luminous band on the left-hand side of the spectra. This shift of the region of maximum intensity will be explained later on. The same phenomenon is observed with monochromatic excitation, as shown by the spectrum on the frontispiece (Fig. 9).

Stokes's law is violated in a most flagrant manner, especially when the exciting light is near the middle of the fluorescent spectrum.

Discovery of Resonance Spectra. — The most remarkable phenomenon of all appeared, however, when the slits of the monochromatic illuminator were narrowed to the width of a hair. The stimulating light was now limited in range to a width not much greater than that of one of the sodium absorption lines. It was at once apparent that the character of the spectrum was much altered, and as the wave-length of the exciting light was slowly changed, the lines of the fluorescent spectrum appeared to move about in the liveliest manner. The whole spectrum appeared in motion, the luminous bands moving in a rippling manner, like moonlight on water. The motion was soon seen to be an illusion due to the continual disappearance and reappearance of the bright lines, the phenomenon reminding one forcibly of the scintillations produced on a zinc sulphide screen by the radium bombardment. This phenomenon led at once to the discovery of resonance spectra, photographs showing that a different series of nearly equidistant lines was produced for each wave-length change of the exciting light. The next step, of course, was to substitute more homogeneous radiation for the light from the monochromator.

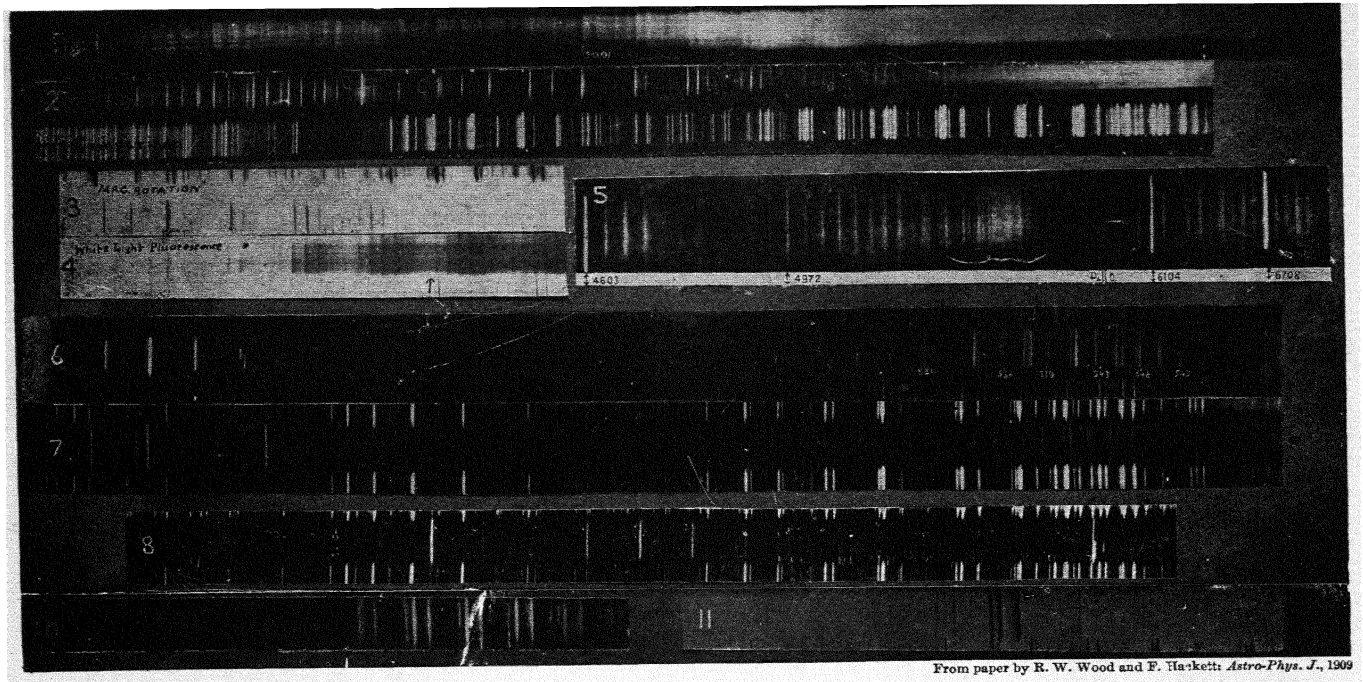
Resonance Spectra of Sodium Vapor. — In the course of an extended study of these spectra, commenced in 1904, the very interesting observation was made that if the vapor was illuminated by monochromatic light of a single line picked out from the spectrum of a metallic arc, it emitted a spectrum consisting of a series of narrow bright lines spaced at very nearly equal intervals along a normal spectrum and separated by distances equal to about 37 Ångström units. Various series of lines with varying distribution of intensity could be brought out by changing the wave-length of the exciting light. In every case light of the same wave-length as that of the exciting light was emitted by the vapor, and in addition a large number of other frequencies which bore definite relations to each other. Discontinuous bright line spectra obtained in this way were named resonance spectra to distinguish them from resonance radiation and fluorescent spectra of the more usual types. A number of photographs of the resonance spectra of so-

dium excited by different monochromatic radiations are reproduced on Plate 8 and also on the colored frontispiece.

A chart (Plate 11) has been made of the principal resonance spectra that have been obtained thus far, and a study of this will give an idea of the complexity of the phenomena. The positions of the exciting lines are indicated by arrows.

In certain cases the resonance spectrum consists of a series of *doublets*, as in the case of excitation by the cadmium 5086 line and the blue-green line of lead. This is a very important point as we shall see when we come to the theory of these spectra. The bismuth arc was found to give the best and most typical resonance spectrum. It has but a single line which is operative, and this line is located in the remote blue. The resonance spectrum consists of a series in the blue which, with long exposures can be traced well up into the green, and a host of lines at the upper end of the spectrum in the yellow-green region, in agreement with the earlier observations. Plate 10, Fig. 6; also colored frontispiece, Fig. 9. The spectrum excited by the triple green line of magnesium is very remarkable for we find the triplet reproduced more or less perfectly at regular intervals along the spectrum. Plate 10, Figs. 10 and 11; also colored frontispiece, Fig. 8. Stokes's law of fluorescence, that only light of longer wave-length can be emitted is violated in a conspicuous manner in all of the spectra being especially noticeable with magnesium excitation. If the spacing of the individual lines forming the series excited by each one of the three components of the triple line was the same for each series, it is clear that we should have the triplet reproduced over and over again along the spectrum. That this is so only to an imperfect degree is due to the difference in the spacing of the lines in the different series. The light of the lithium arc stimulates the most remarkable group of resonance spectra. It was the only means found capable of exciting definite series of lines in the red region. There are four exciting lines in all 4603, 4972, 6104 and 6708, but they are so widely separated that the resonance spectra overlap to only a very slight degree. A photograph of this spectrum is reproduced on Plate 10, Fig. 5, the points of excitation being indicated by arrows (and also on the colored frontispiece Fig. 7). The 4603 lines excited a group in its immediate vicinity and a host of lines in extreme yellow; 4972 excites a series in the green region, the phenomenon being the same as shown in Fig. 360*b*.

The early study of the resonance spectra of sodium was attended with great experimental difficulties, since the intensity of the light emitted under monochromatic stimulation was so faint that exposures varying from 8 to 20 hours were necessary. For this length



From paper by R. W. Wood and F. Haskett: *Astro-Phys. J.*, 1909

PLATE 10. RESONANCE AND MAGNETIC ROTATION SPECTRA OF SODIUM VAPOR.

1. Resonance Radiation Excited by White Light.
2. Magnetic Rotation Spectrum.
3 and 4. Small Portion of Same at $\lambda = 5000$.
5. Resonance Radiation Excited by Lithium Arc.
6. Resonance Radiation Excited by Bi Arc.
7. Resonance Radiation Excited by Zinc Arc.
8. Resonance Radiation Excited by Barium Arc.
10 and 11. Resonance Radiation Excited by Magnesium Arc.

of time it was necessary to keep the metallic arc, which requires constant attention, burning steadily. Some years later it was found [1] that iodine vapor at room temperature in a highly exhausted glass tube emitted a brilliant resonance spectrum when excited by the radiations of the mercury arc, operating in a glass or quartz tube, and at once the study of these spectra became very simple, since these lamps can be operated continuously without attention, and the iodine vapor can be enclosed in glass bulbs or tubes which require no heating, as the most favorable density of the vapor turns out to be that which obtains at room temperature. At the time of the discovery of these resonance spectra it was hoped that they might furnish a clue to the nature of the mechanism giving rise to band spectra. The optical theory of the time was totally unable to account for them, however, and it was not until the advent of the quantum theory that their explanation was given by Lenz who developed a very beautiful theoretical treatment as applied to the special cases of iodine resonance spectra reported by the author [2] which we will now take up in some detail, on account of their importance in connection with the quantum theory of the formation of band spectra.

Resonance Spectra of Iodine.[3] — The band absorption spectrum of iodine extends from the extreme red into the blue-green region and is made up of a large number of fluted bands resembling in its general appearance the channelled absorption of sodium vapor, for in both cases the bands at the long-wave-length end are quite regular in their appearance while at the short-wave-length end they become more or less confused owing to the overlapping of different band groups.

Preliminary to the more complete investigation of the resonance spectra of iodine, this absorption spectrum was exhaustively studied with a plane grating spectrograph of 42-foot focus in the fifth-order spectrum. Sunlight from a heliostat was passed through a large exhausted glass bulb containing a few small crystals of iodine and focussed on the slit of the instrument. The absorption spectrum seen under these conditions presented a most wonderful appearance nearly the entire visible spectrum being filled with thousands of lines. By photographing this spectrum in coincidence with the green emission line of the quartz-mercury arc operated under usual conditions it was found that seven sharp and beautifully resolved absorption lines occurred within the region covered

[1] R. W. Wood, *Phil. Mag.*, Feb. and Oct., 1911.
[2] *Phil. Mag.*, Oct., 1911.
[3] R. W. Wood, *Phys. Zeit.*, *11*, 1195, 1910; *Phil. Mag.*, *22*, 469; *24*, 673; *26*, 828; *35*, 236, 1912–18.

by the green emission line of mercury, the total width of which is about 0.4 A.U. in the case of an arc operated at high temperature. At this rate we have about 18 lines to the Ångström unit or some 36,000 in all. Over 100 absorption lines were counted between the *D* lines of sodium, *i.e.* in a region 6 A.U. in width. A portion of the absorption spectrum in the region of the green mercury line is reproduced on Plate 12, Fig. *a*, the seven lines enumerated above being enclosed between the two arrows. Just below these lines and in coincidence is the spectrum of the green mercury line of the Cooper-Hewitt arc. The resonance spectrum excited by illuminating the iodine vapor with a lamp of this type, in which the green line is so narrow that it covers a single absorption line only, is much simpler in structure than the spectra excited by the quartz-mercury arc, in which the green line has a width sufficient to cover several absorption lines.

The form of tube finally adopted for the study of these spectra is prepared in the following way (Fig. 361). A piece of thin-walled soft glass tubing about 3 cms. in diameter and 40 or 50 cms. long is blown out at one end to the form of a thin bulb 4 cms. in diameter. It is quite important to have this bulb of good optical quality and free from striae, which is not the case with bulbs blown in the usual manner. The important thing is to avoid having the thick drop of glass which is formed at the point where the tube is drawn down come at the centre of the bulb where it forms a lens. This can be accomplished by drawing off the tube sharply in a lateral direction instead of straight out as is usually done. The flame is then directed against the end of the tube and when the glass is well softened the bulb is blown. The other end of the tube is drawn off obliquely, as shown in the figure, and a short lateral branch inserted in the side, which is used for controlling the density of the iodine vapor, as it is often advantageous to work in densities less than that which obtains at room temperature. After the tube has been carefully cleaned and dried a few small flakes of iodine are introduced and shaken down into the lateral branch. The tube must now be exhausted through the "drawn-down" end to a high vacuum with a good oil pump, a U-tube immersed in liquid air being placed in circuit between the tube and the pump to prevent the passage of iodine vapor. The tube should be well heated with a Bunsen burner during the exhaustion to remove adsorbed water vapor, and it is a good plan to keep the end of

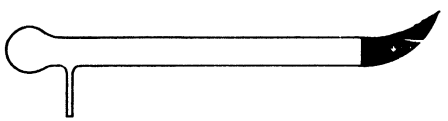
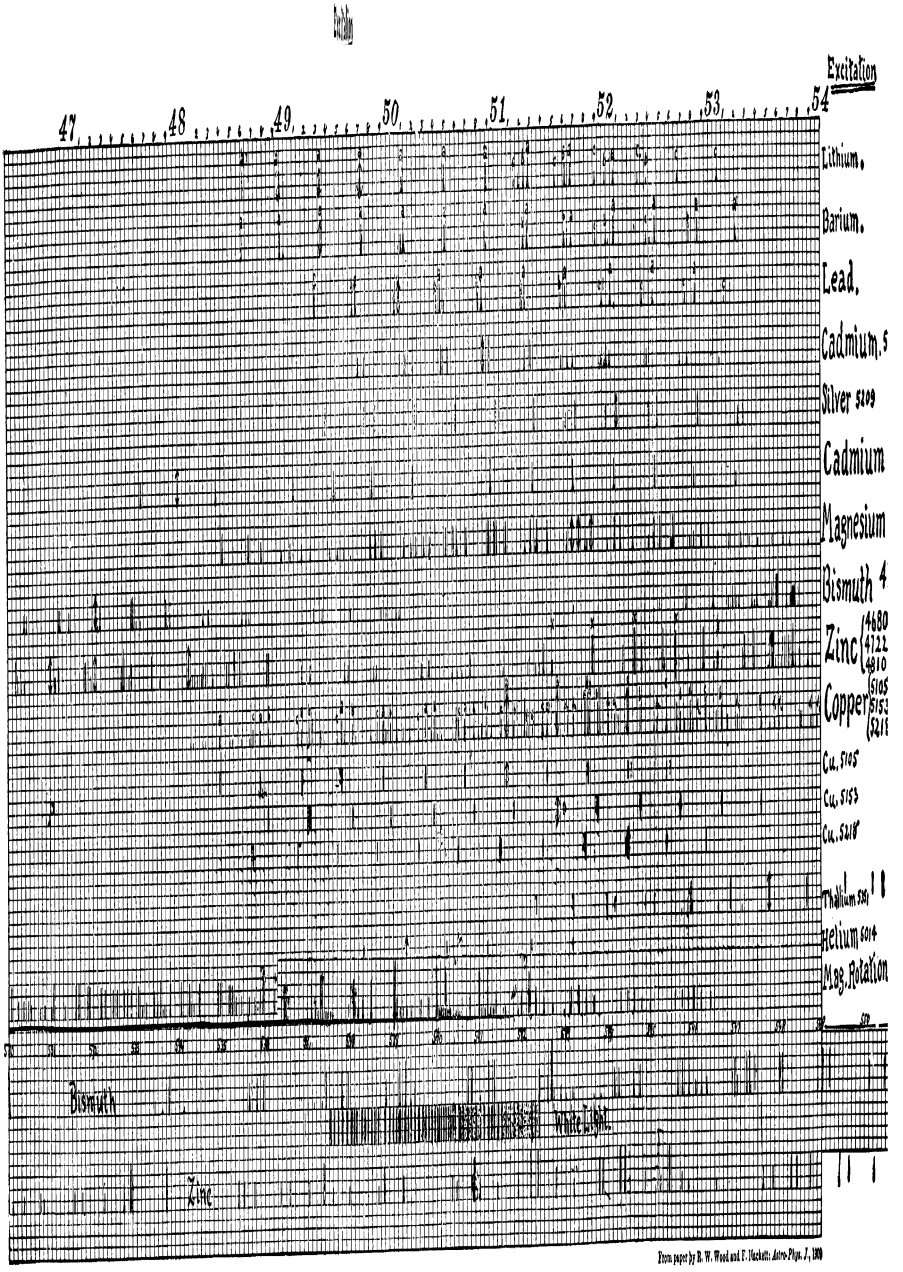


FIG. 361



From paper by R. W. Wood and F. Mackett: *Astro-Phys. J.*, 1909

PLATE II. CURVE OF RESONANCE AND MAGNETIC ROTATION SPECTRA.

the lateral branch at a low temperature to prevent evaporation of the iodine during the process of exhaustion which will require from half an hour to an hour. The iodine is now distilled into the large tube by heating the small tube rather rapidly with a flame, when the tube is quickly sealed off from the pump. This must be done at once, otherwise the iodine will pass out of the tube. It is a good plan to chill a small spot just above the side tube with ice, during the distillation. This will trap most of the iodine and lessen the risk of loss.

For excitation of the resonance spectrum by a Cooper-Hewitt mercury arc it is necessary only to mount the iodine tube parallel to and almost in contact with the glass tube of the lamp. The illumination can be materially increased by surrounding the two tubes with a cylindrical reflector of thin sheet aluminum, which is now obtainable with a high optical polish. The green line is, in this case, so narrow that it covers only a single absorption line, while seven lines are covered by the green line emitted by the quartz arc operating at high temperature. Referring to Plate 12, Fig. *a*, we find between the two arrows the seven iodine absorption lines 1 to 7 above referred to. It is the stimulation of line No. 3 with which we are now concerned. The scale of wave-lengths for the absorption spectrum is immediately below it. The green mercury line $\lambda = 5460.74$ lies on the right-hand edge of the absorption line No. 3 and is nearly equal in width to the distance between lines 3 and 4. The resonance spectrum resulting from the excitation of this absorption line consists of a series of doublets, of which 27 members have been photographed. The doublet which arises at the point of excitation we shall speak of as the group of zero order, and the succeeding doublets of increasing λ's as groups of the first, second, third, etc., order. Groups to the left (short-wave-length side) of the excited line will be designated as -1, -2, etc. On Plate 12, Fig. *b*, will be found a reproduction of the group from the 0 to 5, inclusive, made in the first-order spectrum of a 7″ plane grating combined with a lens of 3-metre focus. The comparison spectrum is that of the molybdenum arc. The doublet of zero order (retouched) is made up of the line resulting from the reëmission without change of wave-length of the light absorbed by line 3 of iodine. This is the line at the extreme left of the spectrogram. Close to this line on the right is the companion line, the two forming the doublet of zero order. On Plate 12, Fig. *a*, this companion line $\lambda = 5462.22$ would come approximately in the position indicated above the spectrogram, the doublet of 0 order having been drawn in its proper position. Some of the emission lines are in coincidence with absorption lines and are weakened in

consequence. For example, if we reduce the density of the iodine vapor by cooling a portion of the tube to zero, the two members of the first-order doublet are of equal intensity; if we warm the tube to 30° or pass the light from the cooler tube through a large globe containing iodine vapor, the right-hand member of the doublet disappears almost completely; in the photograph reproduced its intensity is only about one-third of that of the left-hand member.

Doublets from the 6th to the 22nd order are shown in Figs. *g* and *h*, and the doublets of the 1st and 3rd order in coincidence with the electrically excited band spectrum in Fig. *i*. Employing a quartz-mercury arc for the excitation of the iodine vapor, we find complicated groups of lines in place of the simple doublets. This is due to the fact that the green mercury line has broadened to such a degree that it covers a number of the iodine absorption lines. We may call this Multiplex Excitation.

Groups of strong lines are built up around the strong doublets, weak groups around weak doublets, and only a few very faint lines at the points where the doublets are missing. The complexity of the group depends upon the width of the green line which increases with the potential drop across the terminals of the quartz arc. If sufficient resistance is put in circuit with the arc to keep the potential down to 35 volts, the iodine emits the doublets with

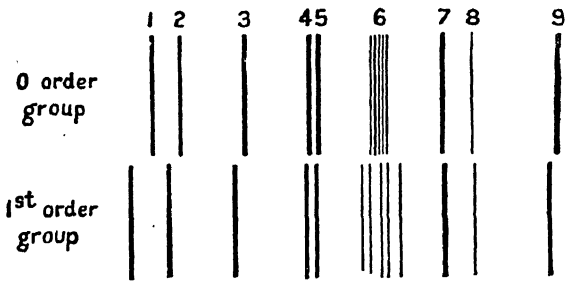


FIG. 362

a fainter companion excited by the short λ component of the green line (Plate 12, Fig. *b*). With the potential at 60 volts we have two new lines to the left of the doublet, Fig. *c*, while with a potential difference of 110 volts we have the complicated groups shown in Fig. *d*. Figures *e* and *f* show the groups between the 4th and 8th orders, the former excited by a Cooper-Hewitt glass lamp, the latter by a quartz arc at 115 volts, the yellow radiations having been removed by a neodymium filter in the latter case.

These groups originate in the following way. The seven absorption lines which are covered by the broad green mercury line are

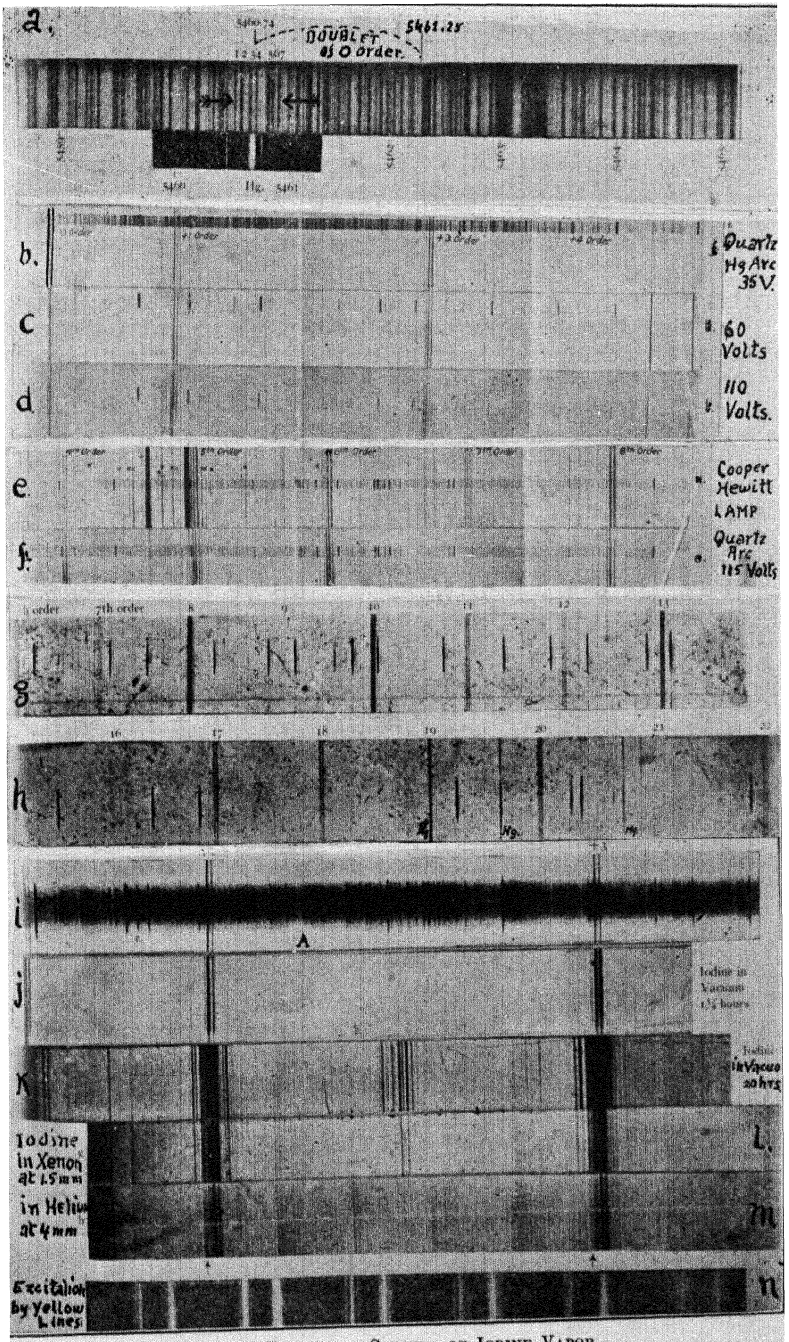


PLATE 12. RESONANCE SPECTRA OF IODINE VAPOR.

simultaneously excited and emit these seven wave-lengths without change. These lines we may call the resonance radiation lines (marked 6 in Fig. 362). Each one of these is, moreover, the first member of a series such as is expressed by the formula which we shall give presently. The resonance radiation lines are not resolved by the spectrograph employed in photographing the resonance spectra, and consequently appear superposed, but each one is accompanied by one or more companion lines lying to the right or left, and it is these companion lines which form the group of zero order. The actual width of the group of seven resonance radiation lines is only about $\frac{1}{30}$ of the width of the group formed by the companion lines. Measurements of the wave-lengths of the doublets made in collaboration with M. Kimura [1] showed that the frequency differences between the components was very nearly a constant, while the wave-length difference increased from 1.51 A.U. at zero order to 2.5 A.U. at the 27th order. For the zero and 20th order we have:

	λ	$\frac{1}{\lambda}$	Freq. Dif.
0.	5462.23	183,075	50
	5463.74	183,025	
20.	6998.96	142,878	
	7001.39	142,828	50

The series as a whole was well represented by the formula

$$\frac{1}{\lambda_m} = 183{,}075 - 2131.414m - 12.734\,\frac{m(m-1)}{2}$$

in which λ_m is the wave-length of one member of the doublet of mth order, 2131.4 the frequency difference between order 0 and $+1$, 12.7 the constant second difference of frequency and m the order of the doublet. Groups of higher order are built up in the following way. Suppose each of the seven resonance radiation lines covered by the green mercury line to be the first member of the series, such as was represented by our formula, and suppose that for each one we have the same values of the constant. Suppose, moreover, that each member of any given series is accompanied by a companion line. In this case the group of zero order will be exactly duplicated at intervals along the spectrum. The centre of each group will be composed of seven superposed lines in reality separated by the same small intervals as the resonance radiation lines, each one of which is accompanied by a companion line to the right or left as the case may be. As a matter of fact, the spac-

[1] *Phil. Mag.*, *35*, 236, 1918.

ing is not exactly the same for the seven series of main lines, consequently as we ascend to higher group orders they begin to separate, even with the resolving-power employed in photographing the resonance spectra, as indicated in 1st order, group 6 of Fig. 363. This accounts for the fact that the groups of higher order differ in appearance from those of lower.

Theory of Resonance Spectra. — Numerous attempts were made to account theoretically for these remarkable spectra, but no progress was made until the advent of the quantum theory, when it was shown by Lenz [1] that they constituted a clear example of Bohr's selection principle as applied to rotational quantum numbers.

The tremendously complicated absorption spectrum of iodine vapor, with its 40,000 or more lines described in the last section, has been analyzed by Mecke,[2] F. W. Loomis [3] and others, the resonance spectra obtained by monochromatic excitation furnishing valuable clues in the unravelling of the tangle of lines formed by the multitude of overlapping bands. The absorption spectra of iodine and sodium vapor are electronic band spectra of the type considered in the Chapter on Origin of Spectra. We will first account for the series of doublets excited by the Cooper-Hewitt mercury arc, in which case the green line is so highly monochromatic that it covers but one of the iodine absorption lines, and all of the iodine molecules which take part in the emission of the resonance spectrum of doublets must be in the same initial state, as regards electron configuration, e, vibration, n, and rotation, j, and they must all be raised to the same higher state by the absorption of quanta of the green radiation.

Since this resonance series shows no anti-Stokes terms (*i.e.* members of shorter wave-length than that of the exciting light) it is most probable that the initial state has zero vibrational energy. The reason for this will appear presently. From the state of excitation represented by the upper energy level the molecules will return to states represented by the various lower levels, and in so doing they will radiate the doublets. We will now see how this comes about, considering first only electronic and vibrational changes, *i.e.* neglecting rotational transitions.

The vibrational motion is anharmonic, consequently all values of n'', the vibrational states of the lower electronic level are possible, and the reversion of the molecules is, in every case, to the same electronic state, since there is but a single fluorescent series,

[1] Lenz, *Phys. Zeit.*, *21*, 691, 1920.
[2] Mecke, *Ann. der Phys.*, *71*, 104, 1923.
[3] Loomis, *Phys. Rev.*, *29*, 112, 1927.

and it contains a line in coincidence with the green exciting line. Denoting the lower electronic state by e'' (Fig. 363), and the upper by e', and the corresponding vibrational states by n'' and n', the state of the molecule before it is optically excited is represented, say, by the dot below the 0 vibrational level of the lower electronic level. This means that the molecule is not in vibration when struck by the light quantum. The energy of the quantum (of wave-length corresponding to the exciting line) raised the molecule to the upper excited state, as indicated by the upward pointing arrow (electronic state e', vibration state $n'=7$). Part of the energy of the exciting light is spent in changing the electronic configuration to e' and part in giving to the molecule 7 vibrational quanta. We will suppose that we have a large number of molecules excited in this way. They will radiate light on their return to the lower state, but though all return to the lower electronic state, they may end up with all possible vibrational states, the transitions being indicated by the downward pointing arrows. The first of these, from the left, causes an emission of the same wave-length as that of the exciting light (termed the resonance radiation line in the earlier papers of the author), the next arrow represents a transition in which the molecule ends up with one quantum of vibrational energy, radiating a line of less energy and, therefore, of longer wave-length, and the other arrows represent transitions giving each one a line of longer λ than the preceding, forming a series of lines very nearly equidistant on the frequency scale. In the case of molecules in which no change of vibrational energy take place in the transition (drop from 7th level of e' to 7th level of e'') the energy of the line radiated represents approximately the energy of the change in the electronic configuration, the excess energy, corresponding to the difference of length of arrows 0–7 (up) and 7–7

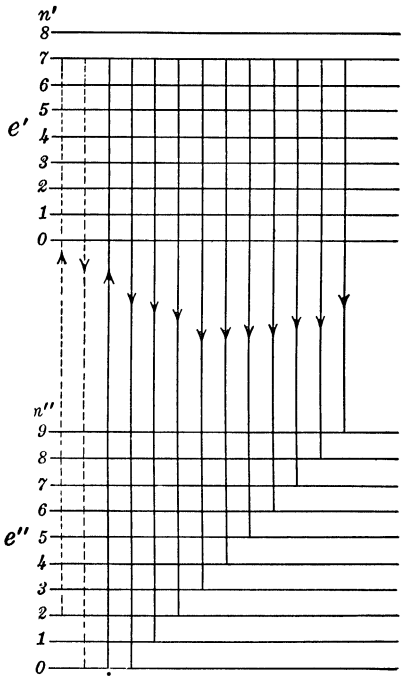


FIG. 363

(down) being expended in giving to the molecule 7 quanta of vibration.

There is nothing, however, to prevent the radiation of lines of still longer wave-length than this line, as by the transitions shown by the two right-hand arrows. In this case the molecule has *more* vibrational energy after the transition than it had in the excited state. In these cases a part only of the energy liberated by the change of electronic configuration is given off as a monochromatic radiation, the remainder being spent in still further increasing the vibrational energy (to levels 8 and 9 of e'' in our diagram).

Now in the case of the iodine doublets, analysis of the complete absorption spectrum has shown that the vibrational transition involved by absorption of the green mercury line is from $n''=0$ to $n'=26$, and it is in the manner described above that there results the extension of the series observed by Oldenberg beyond the 27th term (the last observed in the earlier work) to the 37th term, and still later by the author to the 44th.

If the unexcited molecule in its initial state had two quanta of vibrational energy, a slightly different wave-length would be absorbed and some of the molecules, after the transition back, would end up with vibrational states of 0 and 1. There would thus be radiated two lines of higher energy than that of the exciting line (*i.e.* shorter wave-length) the excess energy being contributed by the energy of vibration originally present in the molecule. A transition of this type is shown by the dotted arrows of Fig. 363.

Thus far we have accounted for the radiation of a series of nearly equidistant spectrum lines extending from the green well into the infra-red region, by iodine vapor absorbing the monochromatic green light of the mercury arc. We will now take up the explanation of the circumstance that we have a series of close doublets instead of single lines, which involves the rotational energy of the molecule.

The principle was established by Bohr, that in transitions of this type the rotational quantum number can change either not at all, or by ± 1, the former case giving a Q branch and the latter the P and R branches of the Fortrat diagram, as explained in the Chapter on Origin of Spectra. This selection principle holds both for absorption and emission. In the case of iodine we have no Q branches, so that the rotational changes are of the ± 1 type.

In the treatment which is being given the excitation is supposed to give a lower vibrational state than is actually the case with iodine excited by the green mercury line. This is for the purpose of simplifying the energy diagram. As a matter of fact the excited state really corresponds to the 26th vibrational level and the

34th rotational level, but we study the phenomenon as if only lower states were involved, as otherwise the diagrams would be too complicated.

In Fig. 364 we have the vibration level $n'=5$ (subdivided into rotation levels j' 0−5) of the upper electronic level, and three vibration levels of the lower electronic level n'', each similarly subdivided. Assume now that the molecules absorbing the light are in the state $n''=2$ and $j''=2$ (shown by black dot). By the absorption of suitable quanta of radiation all are raised to the upper level $n'=5$, $j'=3$, j increasing in this case by +1 (indicated by large arrow head). On reversion to the normal state $j''=j'\pm 1$ by the selection principle, but n'' can take any value. If the return is to the initial state of vibration ($n''=2$) we have the transitions represented by the two middle arrows, the left-hand one, where j changes from 3 to 2 giving rise to the emission of a wave-length equal to that of the absorbed light, the right-hand one, where j changes from 3 to 4, representing a lesser energy change and consequently a slightly longer wave-length giving rise to the companion line, the two forming the doublet of zero order. Other molecules revert to the $n''=3$ vibrational level, giving the +1 order doublet (two right-hand arrows) and others to the $n''=1$ state (two left-hand arrows); these latter give the −1 order doublet (anti-Stokes term).

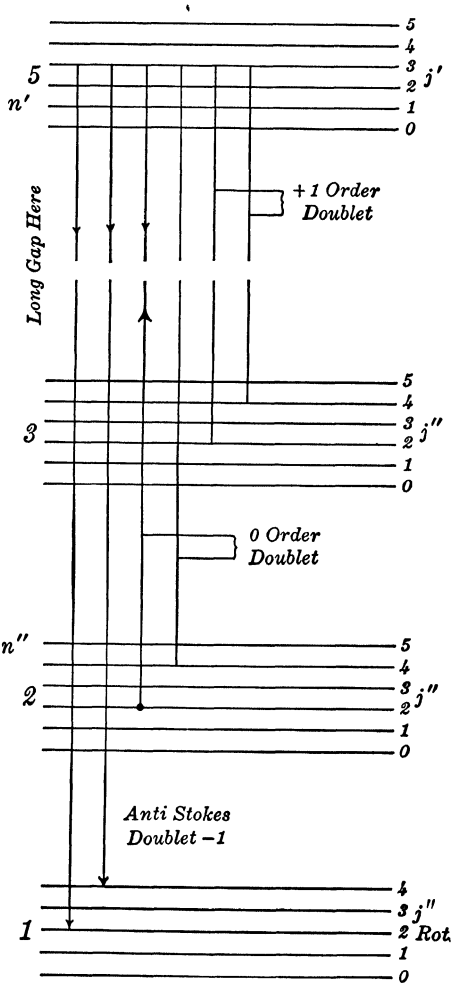


FIG. 364

The molecules in the same initial state might have been raised to the upper rotational level $j'=1$ (j decreasing by 1 instead of increasing as in the previous case). This would require the absorption of quanta of slightly less energy (that is of a wave-length slightly longer than the one just considered), corresponding to another absorption line of the iodine. By the transitions back we should now have j changing from 1 to 2, giving the emission of the line of the same wave-length as that of the absorbed light, and j changing from 1 to 0, giving rise to a line of greater energy or shorter wave-length. In this case the companion line lies on the short-wave-length side of the main line.

The complicated groups of lines, which appear in place of the doublets when the vapor is excited by the light of a high temperature quartz-mercury arc, in which the green line has widened to such an extent that it covers a number of adjacent absorption lines, originate in this way, that is, there are a number of superposed (*i.e.* not resolved) main lines, each one accompanied by a companion line, some of the lines having companions on the long, others on the short-wave-length side. This was pointed out by the author in the original communications on resonance spectra, but the cause was a complete mystery until the matter was cleared up by the very brilliant explanation given by Lenz.

If the absorption line is on an R branch, the companion will be on the long-wave-length side, the reverse being the case if it is on a P branch. This matter will be made clearer if we employ the Fortrat diagram, which was described earlier in the chapter, which will show as well why the distance of the companion line from the main line varies.

We will first consider the excitation of the group of lines of zero order. Figure 365 is a Fortrat diagram by F. W. Loomis for the portion of the iodine absorption spectrum in the vicinity of the green mercury line. The complete parabolas for the band 26, 0 (transitions from 0 vibrational level to the 26th vibrational level, the rotational quanta being shown by small dots) are represented while a portion only of the parabolas for 27, 0, is given: they are to be thought of as curving down to the right and meeting on the lower line. More complete diagrams will be found in the *Report on Molecular Band Spectra* published by National Research Council.

The absorption lines excited by the broadened green mercury line of the quartz arc (the width of which is indicated by the two vertical lines) are represented by black dots between the vertical lines: they are numbered to correspond to the numbers given to the absorption lines by the author in the earlier papers, and shown on

Plate 10, *a*. When the arc is operated at a low temperature, as with the Cooper-Hewitt glass lamp, absorption line 3 (on the *R* branch of band 26, 0) is the only one excited, and the companion line on the *P* branch is indicated also by a black dot, the width of the doublet being indicated by the horizontal line. This is the doublet of zero order with the companion line on the long-wave-length side.

If the excitation is by a quartz-mercury arc operated at 120 volts the green line is much wider and absorption line number 6 is covered. This line is on the *P* branch of the band 27, 0. Since the portions of parabolae for this band are flatter in the region of the

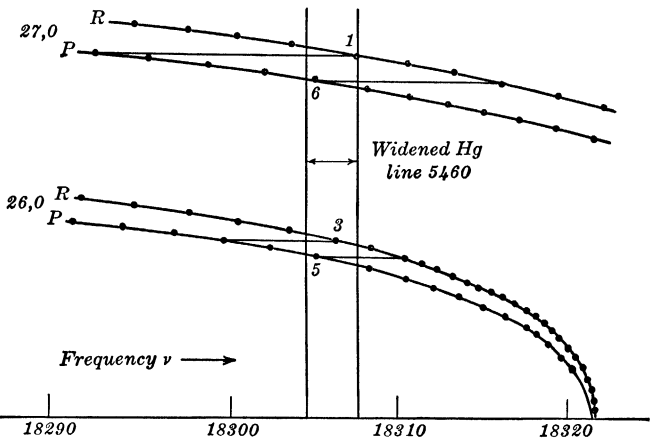


FIG. 365

mercury line, the companion line on the *R* branch will be more widely separated from the main line, the width of the doublet being shown as before by a horizontal line. The companion line in this case is on the short-wave-length side of the main line, as was found in the earlier experiments. As the voltage operating on the arc is increased, other absorption lines are covered, and finally lines on both branches of the parabolae are excited, as shown in the diagram.

If the excitation of the vapor were by a line very close to the vertex of the parabolae, the widths of the fundamental doublets would be much less, and it will be recalled that in the case of sodium, resonance spectra of single lines were found. These were probably unresolved doublets. No anti-Stokes term can occur as the molecule was initially in the zero vibrational state. By excitation of another molecule having one quantum of vibration to start with we have the fundamental doublet on the band 29, 1,

the +1 order doublet on band 29, 2, and the −1 or anti-Stokes doublet on band 29, 0, the excitation in this case being by another absorption line, very close to the one previously considered. Anti-Stokes terms appear in greater number as the temperature of the vapor is raised, a larger percentage of the molecules acquiring vibrational energy. They were nearly always present in the resonance spectra of sodium vapor investigated by the author. The vapor in this case was at a temperature corresponding to a red heat, and recently Pringsheim has observed that the formation of these terms is facilitated in the case of iodine, by super-heating the vapor at 300°, at which temperature molecules having one or more quanta of vibrational energy become relatively more numerous, and the series of doublets which have anti-Stokes terms become relatively brighter, while those which do not, become weaker.

Emission of Polarized Light by Fluorescent Vapors. — It was supposed for many years that polarization was never shown by fluorescence. Schmidt searched for it employing only a Nicol prism, and reported negative results in all cases.

The phenomenon was first detected by the author [1] in the case of potassium vapor by means of a Savart plate which is capable of showing 2% of polarization. It was immediately picked up in the fluorescence of sodium and iodine, and it is probable that other vapors will be found to show it as well.

The apparatus used in the work consisted of a steel tube with a lateral branch brazed to its centre for the observation of the fluorescent light at an angle of 90° with the exciting beam. This tube was used in the earlier work on fluorescence, before the expedient of "end-on" examination had been adopted. The metal was contained in a small retort also brazed to the large tube, immediately below the lateral observation-tube. The light of the arc was focussed by a large lens at the centre of the large tube immediately in front of the lateral branch. Owing to the sensitiveness of the Savart plate it was necessary to eliminate all other possible sources of polarization, such as fog or mist in the tube. Exhaustion with a Gaede mercurial pump while the tube was heated was sufficient to completely banish all trace of fog.

The complete absence of fog can be told by illuminating the vapor with an intense beam of light embracing the spectrum region comprised between wave-lengths 5400 and 5700, obtained by prismatic dispersion. Such a beam is incapable of exciting any fluorescence, and if fog is absent the cone of light is absolutely invisible when viewed through the lateral tube. The Savart

[1] *Phil. Mag.*, July, 1908.

fringes were very distinct, and the percentage of polarization was determined by compensating it with a pair of glass plates which could be turned about on a vertical axis furnished with a graduated circle. The first measurements were made with the exciting light polarized with its electric vector vertical. The polarization was strongest with low vapor density, 30% in the case of the vapor at the lowest density consistent with an observable fluorescence. With unpolarized excitation, the percentage fell to twenty as was to be expected. Strong polarization was also observed in the lines constituting the resonance spectrum of iodine. An image of a quartz arc was formed at the centre of a bulb containing the vapor. The field observed was limited by a vertical slit, in front of which a Fresnel double quartz wedge was mounted, which showed dark bands when viewed through a Nicol and direct vision prism, the latter yielding the resonance spectrum.

Collision Transformation of Resonance to Band Spectrum. — Shortly after the discovery of the resonance spectrum of iodine it was observed by the author and J. Franck that the introduction of helium at a pressure of a few millimetres reduced the intensity of the doublets and developed a complete band spectrum, the total amount of light emitted being about the same.

Subsequent experiments by the author showed that very long exposure also brought out the bands as shown on Plate 10, Fig. *k* a 20-hour exposure with iodine in vacuo in contrast to *j* showing only the doublets with a one-hour exposure. Figs. *l* and *m* show the bands developed by xenon at 1.5 mms. and helium at 4 mms. These experiments showed also that this band system was much simpler than the one excited by white light, and a careful examination of the spectra taken under high dispersion showed that the alternate lines appeared to be missing. A detailed study of this matter was immediately taken up in collaboration with F. W. Loomis.[1]

The tube containing the iodine had been surrounded by a group of four Cooper-Hewitt mercury arcs, which excite such a bright fluorescence that it was possible to photograph the spectrum with much higher resolution than heretofore, making use of a 7-inch plane grating in the second order, with a lens of 3-metre focus. As has been shown the doublets excited by the green mercury line are picked out from the complex band spectrum of iodine by the stimulation of one particular iodine absorption line. The absorption of this line leaves all the excited molecules in a single state ($n'=26$, $m'=34$). Excited molecules which suffer collisions have their rotational and vibrational numbers n' and j' changed.

[1] R. W. Wood and F. W. Loomis, *Phil. Mag.*, vi, 231, 1928.

There will no longer be a single initial state for the emission of the fluorescent light, but a group of vibrational and rotational states clustering about the original one. Consequently the fluorescent spectrum will correspond to transitions from all these states and bands will be developed around the fundamental doublets ($n'=26$), as well as neighboring bands with slightly changed values of n' (24, 25, 27, 28).

A number of these spectra are reproduced on Plate 13 on which *a* is contact print of the resonance band spectrum from the 0 to the fourth order, and *b*, *c* and *d* enlargements of the band heads near the +1 doublet, while *f* is the band associated with the third order doublet. A more complete discussion is found in the paper referred to.

It was evident that the fluorescent spectrum contained only lines starting from states with even values of j', the alternate lines being missing. Now, the original excited molecules, just after absorbing the light from the mercury arc, and before colliding with helium atoms, had $j'=34$. It follows that, *during collisions in which the electronic quantum number is unchanged, the rotational quantum number of an iodine molecule can change only by an even number*, a circumstance that is referred to different orientations of the spin of the two nuclei of the molecule. For a fuller treatment the reader is referred to the original paper or to Jevons, *Report on Band Spectra*.

The Fluorescence of Diatomic Mercury. — The fluorescence of mercury vapor was first observed by Hartley and is easily shown by boiling a little of the metal in a quartz flask or a small quartz bulb blown on the end of a tube and illuminated by an aluminum spark. The bluish green fluorescent light appears only after the boiling mercury has expelled all of the air and is condensing on the tube wall. An investigation of the nature of the spectrum in its relation to the wave-length of the exciting light was commenced by the author in 1907 [1] and continued intermittently until 1928. A large amount of work has been done by other investigators more recently, the phenomena becoming more and more complicated, and resisting until very recently all attempts at an explanation.

In the first of the two papers, the absorption bands which appear, and broaden with increasing density, were described, a continuous one headed at the 2536 absorption line was then referred to as an unsymmetrical broadening of this line on the red side, a second fluted band headed at 2346 and shaded towards shorter wave-lengths, and a continuous one headed below 1850 (Fig. 366). The fluorescence was found to be destroyed by the presence of

[1] *Astrophysical Journal*, 1907; *Phil. Mag.*, Aug., 1909.

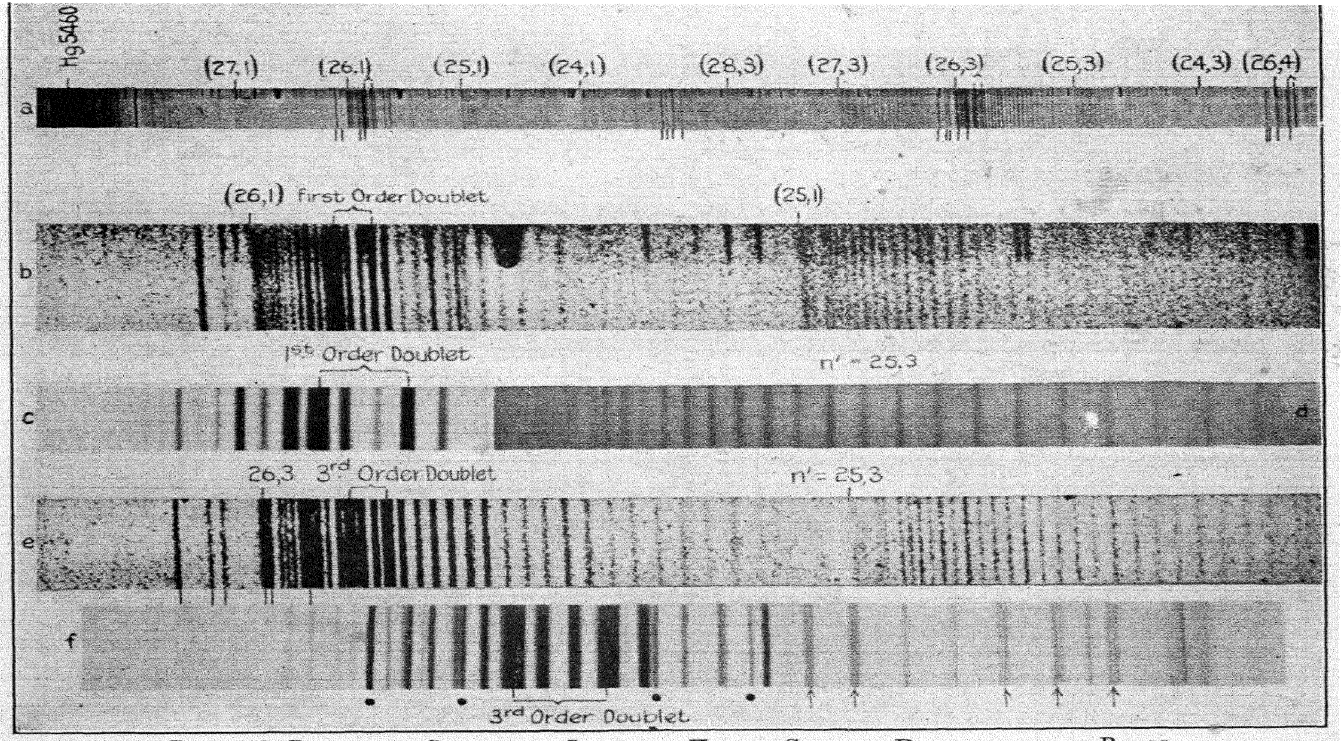


PLATE 13. RESONANCE SPECTRA OF IODINE IN HELIUM SHOWING DEVELOPMENT OF BANDS WITH ALTERNATE MISSING LINES.

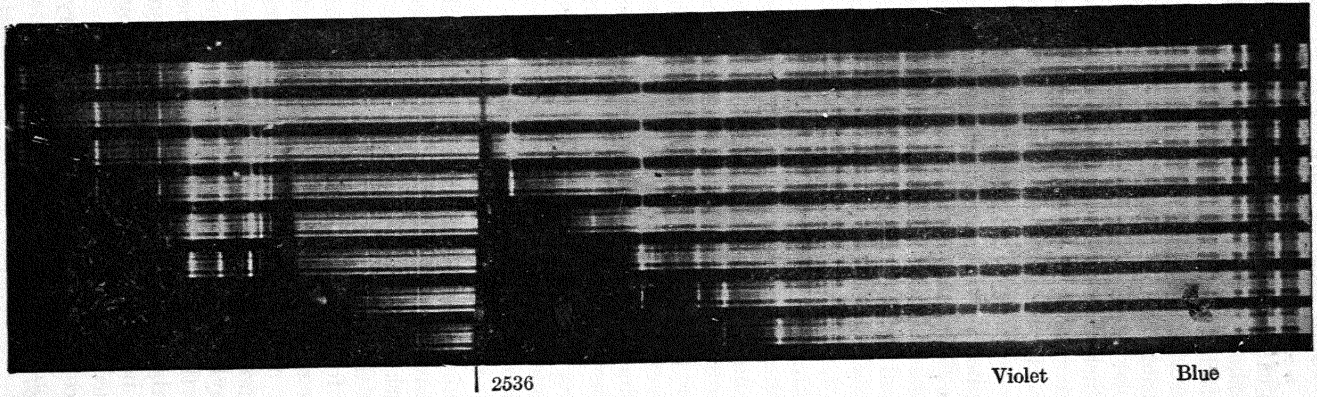


FIG. 366

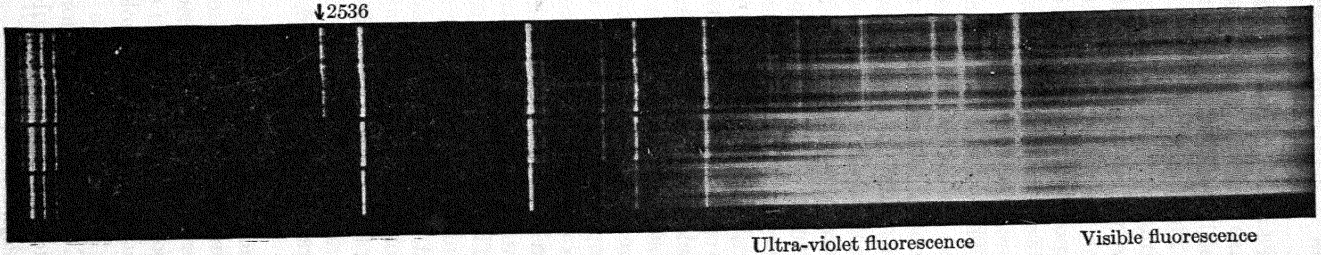


FIG. 367

air and other gases which also caused the appearance of a narrow hazy absorption band on the short-wave-length side of the 2536 line, a phenomenon recently studied by Oldenburg, and others.

The fluorescent spectrum excited by the cadmium spark was found to be made up of a visible band extending from 3700 to the yellow-green, an ultra-violet band, 3000–3500, and in addition the 2536 (atomic) line, which, however appeared only at low vapor density. Monochromatic radiations separated from spectra of Zn, Cd, and Al sparks gave both bands with very little difference in relative intensity.

The same bands were described by Steubing two years later,[1] who obtained in addition the 2346 band in fluorescence excited by an aluminum spark as a fluted band of over twenty members. This band has never been found by the author, and evidently is obtained only under very precise conditions. A continuous band in this same region was, however, obtained in some later work done in collaboration with van der Lingen.

One of these early photographs showing the line at 2536 and the continuous fluorescence spectrum of the vapor at higher density is reproduced in Fig. 367. The bright lines are cadmium. As the work progressed it became more and more evident that something other than atoms were involved in the phenomenon, an early observation being that the fluorescence is destroyed by high temperature. To show this the vapor can be formed by boiling the metal in a quartz test-tube, illuminating it with a cadmium spark and then directing the small pointed flame of a blast lamp against the wall of the tube where the fluorescence is brightest.

The appearance of the 2536 atomic line in the fluorescence excited by a cadmium or aluminum spark was one of the earliest and most puzzling problems investigated, for it appeared only when the direct light of the spark fell on the bulb, and never when the fluorescence was excited by the individual lines from a quartz monochromator. It was finally found that the effective radiations were at the extreme end of the spark spectrum in the vicinity of the 1849 mercury line, where mercury has also an absorption band. These rays were not transmitted by the monochromator. More recent work by Lord Rayleigh has shown that it also appears when the excitation is by the 2539 iron line 2.8 A.U. on the long-wave-length side, *i.e.* nearly at the head of the absorption band of the mercury molecule, but not when the excitation is by the 2544 line.

Evidence that the fluorescence was not due to normal mercury atoms, but probably resulted from aggregates of atoms or diatomic

[1] *Phys. Zeit.*, *10*, 790. 1909.

molecules which developed in rapidly increasing number with small temperature increments appeared to be obvious in the early stages of this investigation. A sealed quartz tube exhausted and containing a drop of mercury, heated in proximity to a spark in an oven showed traces of fluorescence at 170° and strong fluorescence at 180°, certainly a tenfold increase, though the pressure of the mercury vapor had risen only from 8 mms. to 11 mms.

A further study was made in 1921 in collaboration with J. S. van der Lingen,[1] the excitation being by single lines of spark spectra isolated by a quartz monochromator. The green band, 4850, alone appeared when the excitation was by the last aluminum line 1854, the ultra-violet band appearing also when the next aluminum line was used. With excitation by lines in the region of absorption bands of longer wave-length the two fluorescent bands were of nearly equal intensity. The most interesting observation, however, was the phenomenon which occurred in the case of excitation by the group of zinc lines, 2024, 2061, 2109 and 2138 which when employed together for excitation gave the complete fluorescent band 2100–2349; but if used singly gave a portion only of the band, the fluorescent spot moving down the spectrum as the exciting wave-length moved up, a phenomenon precisely similar to the one observed previously with sodium vapor. The explanation of this shift will be given presently. Many circumstances seemed to indicate that only freshly formed vapor was capable of fluorescing, the most convincing experiment being the following:

Two small exhausted quartz bulbs joined by a short narrow tube, and containing a drop of mercury, were heated in a current of hot air rising from an asbestos chimney surrounding a nest of Bunsen burners, and illuminated by the light of an aluminum spark. With both bulbs at the same temperature (210°) no fluorescence was seen, but if a blast of cold air was directed against the bulb in which no metal was visible, the other immediately burst into brilliant green fluorescence, which continued until all of the mercury had condensed in the cooled bulb, when it immediately faded away.

Niewodniczanski, however, obtained contrary results, claiming that he always observed fluorescence with uniformly heated tubes, and that cooling a part of the system invariably decreased the fluorescence, and the matter was further investigated in collaboration with V. Voss.[2]

After many preliminary experiments a double bulb was pre-

[1] *Phys. Rev.*, 1921.

[2] R. W. Wood and V. Voss, *Nature*, February, 1928; *Proc. Roy. Soc.*, *119*, 689. 1928.

pared which could be made to fluoresce brightly at constant temperature, or to show fluorescence only when condensation and distillation was taking place. The bulb contained a trace of water vapor, which could be driven back into the quartz by a high-frequency electrodeless discharge, the bulb becoming almost non-conducting. After this treatment it showed brilliant fluorescence at constant temperature. A small flame was now played over the surface, liberating the occluded vapor, and fluorescence now occurred only when distillation took place, *i.e.* with unequal heating or local cooling. The bulb was carried back and forth from one condition to the other for a number of times. The mystery now disappeared entirely. Water vapor at very low pressure destroys the fluorescence, but the distilling mercury vapor quickly pumps this gas into the other bulb, the pure vapor freshly formed showing fluorescence. The result was confirmed by a similar experiment independently and at about the same time by Pringsheim and Terenin [1] who also made the very important observation that the intensity of fluorescence excited by the aluminum spark or quartz arc, plotted against mercury vapor pressures, gave a line which cut the pressure axis at 4 mms. instead of 0. With an inert gas also in the tube the line cut the axis at 0. This shows that the intensity is a function of something other than the number of mercury molecules, namely the total pressure and that fluorescence results from a collision. Without collision it is probable that the 1849 band is emitted, since the effective aluminum lines are covered by this band. This cleared up in a most satisfactory manner the mystery of the complete absence of fluorescence below a certain critical pressure, and its sudden appearance with a very small pressure increment.

Time Interval between Excitation and Fluorescence of Mercury. — Phillips [2] showed that if mercury vapor, distilling in an inverted U-tube of quartz, was illuminated by the 2536 line of mercury at a point near the surface of the boiling metal, the fluorescent luminosity was carried over the bend by the moving vapor stream.

By a modification of this experiment the author showed that when the illumination was produced by an aluminum spark, the vapor remained dark for a measurable time.[3] A tube of fused quartz of the form shown in Fig. 368 (with a total length of 15 cms. and a bore of 4 mms.) containing about 1 c.c. of mercury, is exhausted and sealed. The tube should be heated during the ex-

[1] *Zeit. für Phys.*, *47*, 330.
[2] *Proc. Roy. Soc.*, *89*, 39.
[3] *Proc. Roy. Soc.*, *99*, 362, 1928.

haustion and the metal boiled. It is mounted in a clamp-stand and the mercury bulb heated by a Bunsen flame about 1 cm. in height. An aluminum spark (condensed discharge) is arranged as shown in the figure, the light passing through a hole 2 mms. in diameter in a thin metal plate mounted close to the tube. On lighting the burner, as soon as active distillation of the mercury commences, we observe a faint greenish light filling the entire tube. This is the phenomenon observed by Phillips [1] with a tube of similar form, excited, however, by the 2536 line of the mercury arc.

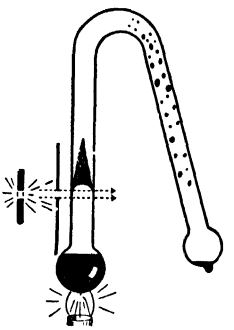


FIG. 368

As the temperature of the "condenser" end of the tube rises, the velocity of the vapor becomes less, the luminosity draws down towards the illuminated region, and presently appears in the form of a beautiful green flame, concave on the under side as shown. The flame form is due obviously to the high velocity of the vapor along the axis of the tube, and the low velocity close to the wall. Absolutely no sign of luminosity is seen in the region traversed by the exciting beam, except a trace on the inner wall of the tube where the velocity of the vapor is very low. At the centre of the tube the dark region extends 2 or 3 mms. above the beam. A photograph of the fluorescent "flame" is reproduced on Plate 14, Fig. 5. Only very rough determinations of the time of the dark interval could be made, owing to the impossibility of calculating the vapor velocity with any accuracy. A new type of phosphoroscope was accordingly designed by which the time interval between the excitation and the emission of the first fluorescent light was found to be 1/15,600 second. With this instrument it was possible to detect phosphorescence of a duration of less than 1/400,000 second, which is about ten times as brief as with previous instruments. This phenomena of the dark interval can be much better shown with the arrangement employed by Lord Rayleigh for studying the fluorescence of moving vapor under controllable conditions.

Effects of Temperature on Mercury Fluorescence. — The early work showed also that the green fluorescence of the vapor was destroyed by superheating. Figure 369 is a photograph of an exhausted quartz bulb containing a drop of mercury and heated by the hot air above a small Bunsen burner fluorescing by the condensed cone of rays from a spark. The application of the blast-lamp flame has caused the fluorescent cone to retreat from the

[1] *Proc. Roy. Soc.*, *89*, 39.

wall. The appearance when the illumination is by direct spark light is shown by Fig. 370.

Lord Rayleigh [1] found that the luminous vapor distilling from a mercury arc in an exhausted quartz tube showed the same lu-

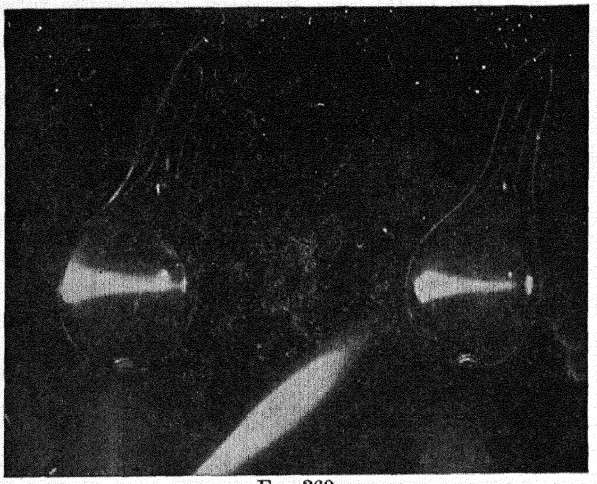


FIG. 369

minous bands as fluorescent mercury vapor and that while the blue-green band was destroyed by superheating, the 3300 band was

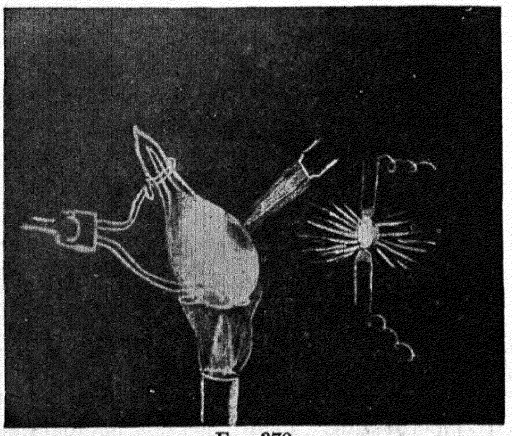


FIG. 370

unaffected. These bands were obviously identical with the bands observed in fluorescence, and this observation, together with Niewodniczanski's report that he had observed the 4358 atomic

[1] *Proc. Roy. Soc.*, 632, 1927.

line in the fluorescence excited by the aluminum spark but not 4046 or 5461 prompted the further investigation of the subject by Wood and Voss referred to above. This investigation brought out very clearly that not only was the green 4800 fluorescent band destroyed by superheating, but that the 3300 band was enormously enhanced. Practically all of the strong atomic arc lines were also obtained with excitation by the aluminum spark.

The intensity of all of the arc lines (except 2536) was increased by superheating at densities corresponding to temperatures below 250°, but for higher densities 4046 and 4358 were decreased in intensity by superheating. It was found, moreover, that the intensity of the arc lines increased with the square of the intensity of the exciting light, the wire gauze method employed in earlier work by the author on the step-up optical excitation of the vapor being used.

In the present case excitation by two successive absorption processes is out of the question, but a similar power relation would obtain if the emission process required a collision between two optically excited molecules since doubling the intensity doubles the number of excited molecules and doubles also the collision frequency of each.

The destruction of the fluorescence by superheating, when first observed, was interpreted as due to the dissociation of mercury molecules formed in small numbers in dense mercury vapor. Inasmuch as the green fluorescent band is destroyed and the ultra-violet band enormously enhanced, it now seems more probable that these effects are related to energy levels, the ultra-violet band emission involving molecules in higher energy states than those concerned in the emission of the green band. It seems possible that the former are the ones which by collision give rise to the atomic arc lines as these lines are enhanced by superheating. The more complete theory will be given presently.

Much additional light has been thrown on this very complicated type of vapor fluorescence by the quite recent experiments of Lord Rayleigh [1] who obtained very beautiful and striking results in his investigation of the luminosity carried along by a vapor current rendered fluorescent in a restricted region, as in the experiments of Phillips and of the author. Working with a more complicated and far more satisfactory apparatus, in which the velocity of the mercury vapor jet as well as its density could be controlled, he made a long series of experiments on the excitation of the moving vapor by radiations of different frequencies.

[1] *Proc. Roy. Soc.*, *125*, 1, 1929; *132*, 650, 1931; *135*, 617, 1932; *137*, 101, 1932; *139*, 507, 1933.

Using the resonance line 2537 for the excitation, he found that, by focussing light of wave-length 4046 or 5461 from a second mercury arc on the green fluorescence carried up by the vapor stream, he obtained a stationary spot of green light, showing the presence of 2^3P_0 and 2^3P_2 atoms. The resonance radiation does not move up with the vapor stream, but diffuses uniformly up and down stream, except at very low pressures when the velocity of the vapor is more rapid than the handing on of 2536 quanta from atom to atom (imprisonment or radiation). He regarded the long duration of the fluorescence as due to the presence of metastable excited atoms. He confirmed the earlier observation of Wood and Voss that the atomic arc lines appeared with an intensity depending on the square of the intensity of the exciting radiation (in this case by Hg 2536) which he calls "core" excitation, finding in addition that the lines appeared in the "persistent" luminosity carried up by the vapor stream.

Of still greater importance was his discovery of faint fluted bands associated with the more intense continuous bands, a circumstance which has aided greatly the development of a theory of the processes which are operating.

He used two chief types of excitation, the monochromatic 2536 radiation, of a water-cooled, magnetically deflected mercury arc, which he terms "core" excitation, and "wing" excitation, which may be the wing of the 2536 line from a hot arc, which encroaches on the absorption band headed at 2539, or radiations from the iron arc in this same region. The complete fluorescent spectrum, obtained with core excitation is reproduced on Plate 14, Fig. 1, which is a composite made from a number of his photographs. Beginning at the left, we find first a fluted band at 2345 coinciding with the fluted band found in absorption by the author in the earlier work. This is a striking violation of Stokes's law, if the 1849 mercury line was completely eliminated. Next comes a strong band headed at 2539 near the exciting line, a second strong band centred at the mercury triple line 2656, then the very weak fluted band (core series) and finally the strong continuous bands at 3300 and 4850. The short spectrum above this one is the emission of electrically excited vapor, showing the fluted band to better advantage. Figure 2 shows the effect of superheating, which destroys the 4850 band completely and enhances 3300 and the "core" series.

With "wing" excitation by a group of iron lines between 2536 and 2625, the same bands were found as with core excitation, but measurements showed that the fluted band in this case was not identical with the one found by core excitation. This band he

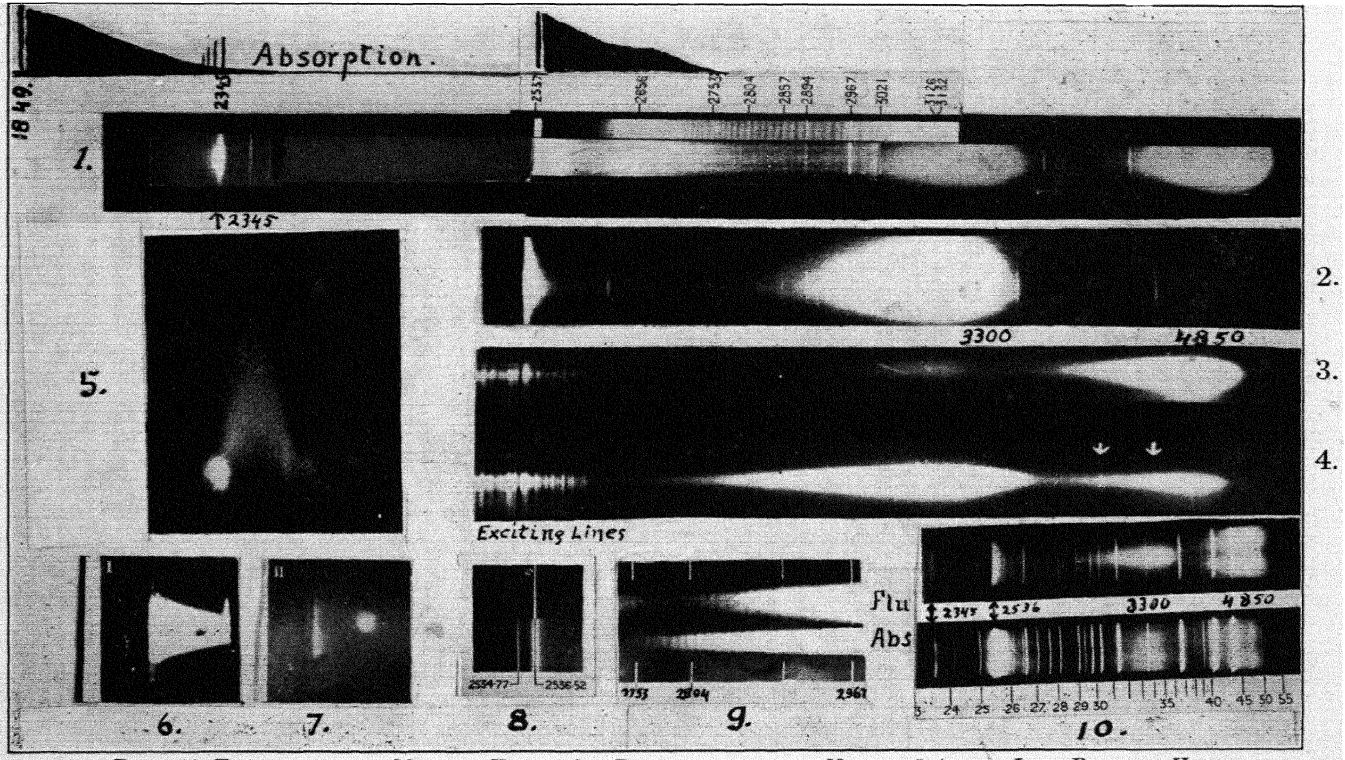


PLATE 14. FLUORESCENCE OF MERCURY VAPOR. ALL PHOTOGRAPHS EXCEPT NUMBER 5 ARE BY LORD RAYLEIGH II.

called the "wing" series, and he found it also in the absorption spectrum. The spectrum with wing excitation is shown by Figs. 3 and 4, the latter with superheating of the vapor showing the weakening of the 4850 and the enhancement of the 3300 band and the "wing" series. The 4047 and 4358 arc lines of Hg are indicated by arrows. The wing series in fluorescence and in absorption are shown in Fig. 9.

Lord Rayleigh also employed an electrically heated wire for demonstrating in a striking manner the effect of superheating the vapor. Photographs of the region near the wire seen "end-on" by the light of the 4850 band, in which a dark spot appears around the wire, and the 3300 band in which case the spot appears bright (Figs. 6 and 7). Fig. 8 shows the 2536 line of a cooled and hot mercury arc. This method of heating the vapor was employed in the study of the superheating effect on the fainter bands. The upper spectrum in Fig. 10 was made with the vapor at 20 mms., the "end-on" image of the wire being focussed on the slit. The weakening of a band is shown by a dark horizontal region, which appears in band 4850, while the luminosity of the 3300 band is confined chiefly to the vicinity of the hot wire. With the vapor at 2 mms. (lower spectrum) bands 2345, 3300 and 4850 are all weakened by super-heating.

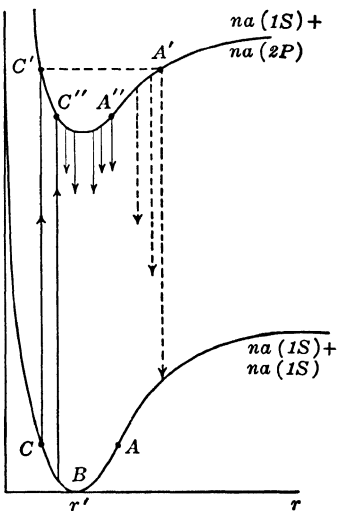


FIG. 371

Theory of Sodium and Mercury Fluorescence. The Franck and Condon Principle. — The simple explanation of resonance spectra by energy level diagrams, requires elaboration to account for such phenomena as the shift in the region of fluorescence down the spectrum as the exciting frequency moves up (as found in the case of sodium, and in one of the mercury bands) and the very complicated effects found with mercury. The extension of the theory we owe to Franck and Condon, and its application to sodium to F. W. Loomis. In Fig. 371 the potential energy of a sodium molecule is represented by the ordinates of curves of which the abscissae represent the distances between the nuclei, the lower curve for a normal the upper for an excited molecule, the physical interpretation being as follows. As the two atoms approach there come into play two forces, one of attraction and

another of repulsion, which vary with different powers of r, the internuclear distance. For a certain distance r', the forces neutralize each other, this being the position of equilibrium. If a to-and-fro vibration occurs the curve is traversed back and forth, say from A (maximum potential energy) through B, the equilibrium position (zero energy), to C (again maximum energy). Notice that the left branch of the curve is much steeper than the right branch, owing to the more rapid increase of the repulsive force with decreasing r. The absorption of a frequency corresponding to a line of the band absorption spectrum is represented by the upward arrow from C. This gives us an excited molecule, one atom in the normal, the other in the $2P$ state. The Franck and Condon principle, which governs transitions of this type, considers that the electronic change occupies a time so small in comparison to the time of vibration, that the internuclear distance remains constant during the transition, which means that transitions between two potential curves can take place only along vertical lines. Transitions back to the lower state will be more frequent from points on the curve near C' and A', since these points represent maximum amplitudes of the vibration (where the velocity changes sign passing through zero value), and the molecule is on these parts of the curve for a much longer time than when near the middle points which are quickly traversed, owing to the high velocity. The upper vibration curve is quantized and downward transitions can occur from other points in the vicinity of C' and A', represented by dots. The emission will therefore consist of a series of lines in the short-wave-length region (where the excitation took place) and another group at the long-wave-length end of the band spectrum. If the excitation is by a lower frequency, giving less vibrational energy we have emission corresponding to transitions from points in the vicinity of C'' and A'', in other words, the spectral regions of fluorescence have approached, as was observed in the early work of the author on sodium fluorescence, and later on in the case of one of the mercury bands by Wood and van der Lingen. The application of the Franck and Condon principle to the fluorescence of mercury was made by Kuhn.[1]

In this case the lower or unexcited state is represented by a repulsion curve as shown in Fig. 372. This means that stable molecules are not formed by two atoms in the normal state, though two of them, separated by a small distance, may act as a unit in an absorption process, this process giving rise to a stable molecule with one atom in an excited state, as represented by either one of the curves 2 and 3. Lord Rayleigh proved the presence of 2^3P_0

[1] *Zeit. für Phys.*, *72*, 462, 1931.

and 2^3P_2 atoms in the moving stream of fluorescent vapor excited by the 2536 mercury line. The initial excitation gives a 2^3P_1 atom, but in dense vapor or in Hg+ nitrogen the greater part are transformed into 2^3P_0 metastable atoms. If one of these collides with a normal atom, the 2^3P_0 atom traverses curve 2 from right to left, and back, and the atoms separate without forming a stable molecule. If, however, a third atom collides with them, it may carry away some of the energy, leaving a molecule in the state represented by the lower part of curve 2. To transitions from 2 to 1 Kuhn attributes Lord Rayleigh's "core" bands and the green band, the latter resulting from molecules which have lost most of their vibrational energy (minimum of the potential curve). Both are decreased by superheating. The wing bands, however, are associated with the 3300 band, and both are increased by superheating, the increased kinetic energy in this case favoring the formation of molecules in the state shown by curve 3, one atom being in the 2^3P_1 state. This is in agreement with the observation of Cario and Franck that heating prevents the formation of metastable atoms. In the opinion of the author the high intensity of the 3300 band and the low intensity of the wing bands is not well accounted for. The dark period of 10^{-4} sec. following excitation before luminosity appears, results from the rarity of the triple collision required.

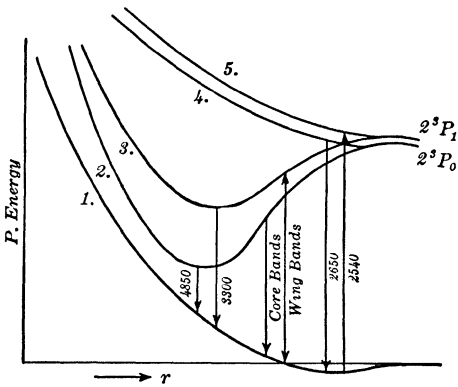


FIG. 372

The excitation of the 4850 and 3300 band by aluminum spark lines of very short wave-length results from the formation of molecules in higher states, associated with the 1849 mercury line, instead of the 2536. That these bands are without structure results from the circumstance that the upper and lower potential curves slope the same way and the lengths of the vertical transitions change very gradually. Where fluted bands occur the slope is in opposite directions. The bands 2540 and 2650 are ascribed by Kuhn to transitions between the repulsion curves 1, and 4 or 5, as shown in the figure. Lord Rayleigh obtained the forbidden line 2270 (transition from 2^3P_2 to $1S$) whenever the 2345 band appeared. This suggests that the 2^3P_2 atoms are concerned in its production.

CHAPTER XX

THE FLUORESCENCE AND PHOSPHORESCENCE OF SOLIDS AND LIQUIDS

Resonance Radiation of Solids and Liquids. — There are some solutions and a number of crystallized solids which exhibit absorption lines almost as narrow as the absorption lines shown by vapors. No liquid or solution has, however, been found to exhibit the slightest trace of resonance radiation. This is perhaps to be expected for we must remember that the resonance radiation of mercury vapor and sodium disappears entirely as a result of collisions of the atoms with other gases at low pressure. This mutual influence of the atoms or molecules in destroying their resonance radiation will prevail to an even greater degree in the case of a solution where the molecules are in closer proximity. There appears to be one case, however, of a crystalline solid which exhibits resonance radiation or something very similar to it. The ruby, either natural or artificial, has a pair of very narrow absorption lines in the remote red region of the spectrum. These lines become still narrower if the crystal is cooled with liquid air. On illuminating such a crystal with an intense beam of white light it scatters light of a deep red color which the spectroscope shows as a double line in the red coinciding with the position of the two absorption lines just referred to. This seemingly unique case was found in the course of an investigation on the optical properties of absorbing crystals made by the author in collaboration with Professor Mendenhall. It had been previously observed by Ebert.

The crystal examined was a synthetic ruby cut in the form of a cube. It was cooled with liquid air and illuminated with light from a prismatic monochromator. On gradually changing the wavelength of the exciting light the crystal was found to glow only when illuminated by the narrow region of the spectrum corresponding to the position of the two absorption lines.

Absorption of Light in Solids. — Very little work has been done on the actual processes which occur when light is absorbed by a solid, as the phenomena are, in general, very complicated. We have, at the present time, a fairly definite picture of what takes place when light is absorbed by a gas, and a start has been made

in the case of certain solids of simple constitution by Pohl and his collaborators.[1]

We will take as an example the case of potassium bromide, of which large single crystals have been prepared in Pohl's laboratory. This substance has two well-marked absorption bands in the ultra-violet, but is quite transparent to the rest of the spectrum. The bands were located by measuring the intensity of the light transmitted through very thin flakes by means of a photo-electric cell. This absorption process gives rise to a photo-chemical reaction, an electron passing from an anion to a kation, the product of the reaction having an absorption band in the visible spectrum, with its maximum at 6500. The concentration of this material is, however, too small to give rise to an observable absorption (*i.e.* to color the substance) but it can be recognized by an electrical method.

The first process cannot be detected, for the electron goes from an anion to the next kation transforming it into a neutral potassium atom, which can absorb light in the region covered by band 2. This process releases the electron which goes back to the bromine atom, producing KBr with absorption band 1 (Fig. 373).

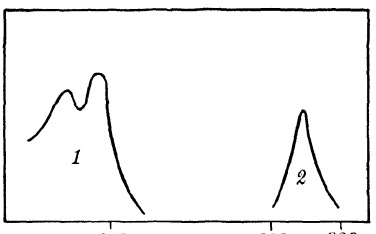


FIG. 373

Their apparatus is shown in Fig. 374. A plate of KBr 5×15 ×50 mms. is gilded on both sides by evaporating gold in a high vacuum. A "guard ring" 5 mms. wide is cut through the lower film and an electrometer and battery connected as shown. Ultra-violet light from a small spark is first allowed to enter the plate from the left for one-fifth of a second. The electrometer remains at rest during this process. Light from an arc filtered free from radiations below 550 is next passed through the plate from the right. The electrometer now shows a deflection which rises to a maximum in a fraction of a second.

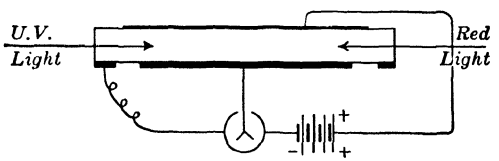


FIG. 374

If the plate is illuminated with the total radiation of the arc, a

[1] Hilsch and Pohl. *Phys. Zeit.*, *32*. 734. 1931.

steady current of 2.10^{-11} amperes can be shown to flow. This is due to the fact that the ultra-violet absorption band does not really terminate at 200 as shown in the figure but runs asymptotically to the zero line and reaches the region emitted by the arc. The two processes thus occur simultaneously.

The authors state that the absorption in this region is so small that a layer of KBr 5 kilometres thick would be required to reduce violet light to half value. The action of the KBr is analogous to the action of the photographic plate, the latent image corresponding to the KBr crystal after illumination with ultra-violet light.

Fluorescence of Organic Compounds. — The earlier studies of fluorescence were made with organic compounds of complicated molecular structure, usually in solution in water or some other medium. The observation of the phenomenon dates back to the 16th century, but the first real study of the subject was made by Herschel in 1845 and Sir David Brewster in 1846.

Herschel found that light which had traversed the solution was incapable of exciting any further emission, and that the blue luminosity was confined to the surface. He termed the phenomenon epipolic dispersion, believing that he was dealing with a new type of polarization. The light on entering the solution became "epipolarized," a lateral emission or dispersion resulting from the process, and this epipolarized light was incapable of exciting further fluorescence. Brewster found, however, that by employing an intense beam of light, the blue emission marked the entire path of the beam, and he accordingly changed the name to "internal dispersion."

It occurred to Stokes that the blue light "dispersed" by the quinine solution might not be the blue light of the illuminating beam, but a new creation due to the absorption of more refrangible radiations. This would explain the inability of the light to excite further fluorescence after it had already traversed a sufficient thickness of the solution, the rays effective in provoking the emission being removed by absorption. Experiments verified this surmise, establishing the general law that the fluorescent radiations are always of longer wave-lengths than those of the light which excites them. This change in the wave-length can be very easily observed by interposing colored glass in the path of the illuminating beam. A beam of sunlight, from which the orange-yellow and green has been removed by means of a sheet of dense cobalt glass, is concentrated with a lens upon a few crystals of uranium nitrate. Although the illuminating beam is of a deep blue violet color the crystals shine with a brilliant green light. A piece of the ordinary canary glass, which is colored with oxide of uranium, can be sub-

stituted for the nitrate crystals. Small vases of this glass, which can now be found in almost any glass-store, are suitable for the experiment. It can be easily recognized by the greenish color which it assumes when held in sunlight, which is in marked contrast to the yellow color of the transmitted light. One of the best substances for the exhibition of fluorescence is an extremely dilute aqueous solution of uranin, an alkaline salt of fluorescein.

Aesculin, which can be prepared by soaking the bark of the horse-chestnut in hot water, and sulphate of quinine in dilute sulphuric acid shine with a blue light. The red fluorescence of the green solution of chlorophyl is one of the earliest known cases. Green leaves, previously dried in a warm oven, are soaked in alcohol, and the solution filtered. Prepared in this way the solution remains in good condition indefinitely. If the leaves are not dried, the solution decomposes in a short time. Sunlight, concentrated by a lens, causes a red fluorescence of the green solution.

Under powerful illumination by invisible ultra-violet rays, as with the lamp developed by the author in which the light of a high temperature quartz-mercury arc is filtered through a plate of glass colored with oxide of nickel, practically all organic materials are more or less fluorescent. The lens and cornea of the human eye, the teeth, hair and skin, shine with a bluish light, while artificial teeth appear jet black. The fluorescence of the lens of the eye illuminates the retina uniformly, giving rise to the illusion of the "lavender fog" first described by the author in 1918.[1]

The absorption and fluorescence of benzene is of especial interest, as the phenomena are displayed by the substance in the vapor, liquid and solid state, and the structure of the molecule is well known. In some respects its behavior is not unlike that of iodine, for it exhibits resonance doublets under monochromatic excitation when at extremely low pressure, and vibration bands when at higher pressure. The earlier studies of the absorption and fluorescence by V. Henri [2] showed that the ultra-violet bands discovered by Hartley were resolvable into lines which conformed to a characteristic band formula of the usual type. The fluorescence spectrum was also investigated by McVicker and March [3] and by Stewart.[4] It resembles the absorption spectrum in structure, but the bands can be traced much farther into the region of longer wave-lengths.

[1] *Jour. d. Phys.*, 1919. Photographs made by ultra-violet light by the author will be found in the *Encycl. Brit.* (14th ed.), article on "Fluorescence."
[2] *Jour. d. Phys. et le Radium*, *3*, 181, 1922.
[3] *Jour. of Chem. Soc.*, *123*, 820, 1923.
[4] *Phil. Mag.*, *48*, 268, 1924.

A few of the energy levels and transitions involved are shown in Fig. 375 though transitions involving level differences as great as ± 6 occur.

Excitation by monochromatic radiation from the mercury arc and the Zn, Sn and Al spark give the complete fluorescence spectrum, which indicates that molecular collisions change the sharply defined excited states resulting from monochromatic excitations, to all possible excited states in much the same manner as the presence of a few millimetres of helium transforms the resonance spectrum of iodine doublets into a complete series of band spectra. A. Reimann [1] found that at a vapor pressure of less than 1 mm. at $-20°$, the fluorescence spectrum excited by the 2536 mercury line is much simplified, many of the stronger bands being absent, while some of the weaker bands are relatively strong. The admission of a foreign gas to the tube caused the complete band spectrum to appear.

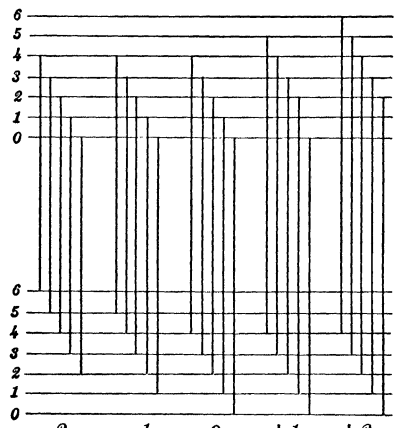


FIG. 375

The band heads corresponding to energy level transfers from $+6$ to -6 are as follows:

	+6	+5	+4	+3	+2	+1	0	−1	−2	−3	−4	−5	−6
Vapor Abs.	2275	2324	2363	2416	2471	2528	2589	2667					
Fluor.						2541	2602	2667	2739	2815	2895	2980	3065
Dil. Sol. Abs. in Alcohol	2290	2330	2378	2428	2485	2547	2598	2681					
Fluor.							2599	2679	2754	2827	2910	3005	
Pure Fluid Abs.	2297	2339	2385	2432	2488	2550	2608	2689					
Solid Fluor.								2686	2761	2837	2920		

In this table we see clearly the extension of the fluorescence spectrum towards increasing wave-lengths. The spectrum of the solid benzene at $-180°$ is much simpler than that of the vapor, many of the lines forming the vapor band being absent.

Resonance Spectrum of Benzene. — A very interesting discovery was made by Kistiakowski and Nelles [2] who illuminated the vapor at the very low pressure of .01 mm. with the light of a mercury arc and obtained a resonance spectrum of doublets very

[1] *Zeit. für Phys.*, *29*, 115, 1924.
[2] *Phys. Rev.*, *41*, 595, 1932.

similar to, and originating in the same way as, the resonance spectrum of iodine. The ordinary high-pressure fluorescence they attributed to transitions from the lowest vibrational level of the excited states. This level is reached by excited molecules as a result of inelastic collision.

Fluorescent Efficiency and Stokes's Law. — The ratio of the amount of light emitted in fluorescence to the amount absorbed, or the yield, is termed the fluorescent efficiency. It is important, however, to note that there is a difference according as we measure the two quantities in terms of energy, or by the number of light quanta or photons absorbed and emitted.

Stokes was of the opinion that the wave-length of the emitted light was always greater than that of the exciting light (the so-called Stokes's law) and disputes followed extending over many years, as to whether any exceptions existed. Nichols and Merritt [1] showed by very careful photometric measurements that the position of the maximum of the fluorescent spectrum was independent of the wave-length of the exciting light, and that fluorescence was still observable when the wave-length of the exciting radiations was greater than that at the centre of the fluorescent band.

We know now that the exceptions to Stokes's law occur when the atom or molecule possesses energy of vibration at the time of its excitation to a higher energy state, the return transition carrying it to a lower state than the original one, resulting in the emission of a frequency higher than that of the exciting radiation. The exceptions are not very striking in the case of fluorescent solutions, but are very marked with vapors as we have seen.

The highest fluorescent efficiency appears to be shown by uranin in water solution of proper concentration, which was determined by Wawilow.[2] Measured in radiated energy a value of 0.8 was found. The efficiency diminishes rapidly with increase of concentration, owing to disactivation by neighboring molecules. Owing to the change of frequency, the efficiency, expressed as a ratio of emitted to absorbed photons, is a little higher than the value expressed in energies.

To make Stokes's law conform to present knowledge it must be expressed as follows. The "centre of gravity" of the complete fluorescent band, is always of greater wave-length than that of the excitation band, the latter term meaning a superposition on the absorption spectrum of the energy curve of the exciting light. In a subsequent investigation Nichols and Merritt [3] investigated the

[1] *Phys. Rev.*, 1904.
[2] *Zeit. für Phys.*, *22*, 266, 1924.
[3] *Phys. Rev.*, *31*, 381, 1910.

relation between the absorbed and emitted energies in more precise form. Their results are shown in the curves of Fig. 376.

Curves 1 and 2 represent absorption and fluorescence spectra respectively; curve 3 the relative intensities of the fluorescence excited by various frequencies of radiation from a monochromator and curve 4 the specific excitation coefficient of various frequencies for equal amounts of absorption, *i.e.* the fluorescent efficiencies for the various excitation frequencies, the curve showing no marked maximum. Pushing the investigation over a wider range of frequencies Valentiner and Rössiger found that the efficiency falls off rapidly as the wave-length of the exciting radiation passes the maximum of the fluorescence curve. Wawilow [2] obtained a curve for the efficiency of uranin which rose rapidly with increase of wave-length of the exciting light, became nearly horizontal in the region of strongest absorption and then fell rapidly with further increment of wave-length. This rise is to be expected on the quantum theory, since for equal amounts of absorbed energy the number of quanta will increase with increasing wave-length, and a constant efficiency will be found only when it is referred to equal numbers of absorbed quanta and not to equal energies.

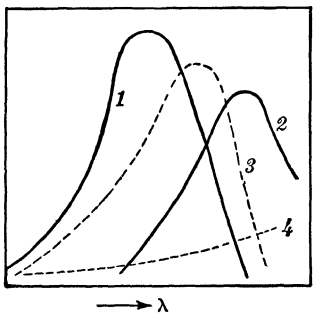


FIG. 376

The increment of energy necessary on the quantum theory for the emission of frequencies higher than that of the exciting radiation must come from the molecules which are in higher energy states as a result of temperature. An experiment made by the author apparently indicates that the fluorescence of fluorescein in a water solution, excited by the longest wave-lengths capable of stimulating a visible radiation, is perceptibly brighter with the solution at 100° than at 0°, but there is a slight temperature shift of the absorption band, which makes this result less precise. The fluid was illuminated in a test-tube by light from a monochromator, the lower portion at 0°, the upper at 100° (by quick local heating with a flame). On moving the tube up and down in the path of the light the change of intensity was very pronounced.

Methods of Investigating Fluorescence. — An exceedingly simple and ingenious method was devised by Stokes for detecting fluorescence and phosphorescence, which is applicable to cases

[1] *Zeit. für Phys.*, *22*, 266, 1924; *42*, 311, 1927.

where the emitted light is so feeble as to be overpowered by the irregularly reflected light. This method depends on the change of wave-length which accompanies both phenomena. Two screens are prepared, one of which transmits the violet and blue, absorbing the green, yellow and red, while the other absorbs the violet and blue, transmitting the rest of the spectrum. Dense cobalt glass combined with a thin sheet of signal green glass, or a solution of cuprammonium, makes a suitable blue-violet screen; while yellow glass or a solution of bichromate of potash will answer for the other. The two together should be practically opaque even to a fairly strong light.

If, now, a powerful beam of light is admitted to a dark room or box through the blue screen, objects illuminated by it will be invisible through the yellow screen unless they fluoresce or phosphoresce, that is, give out less refrangible radiations than those which fall upon them. Stokes succeeded in showing that ordinary paper, cotton, bones, ivory, leather, cork, horn, and many other substances exhibit the phenomenon.

This method, while admirably adapted to the detection of fluorescence, is not suited to the study of the relation between the wave-lengths of the fluorescent and incident light.

The method adopted by Stokes was analogous to Newton's method of crossed prisms; a very narrow and intense solar spectrum was thrown upon the surface of the liquid under investigation, the fluorescence resulting in this case from monochromatic light of varying wave-length. This spectrum was then viewed through a prism held in such a position as to deviate the spectrum in a direction perpendicular to its length, as shown in Fig. 377, in which AB is the undeviated and $A'B'$ the deviated spectrum, as it would appear if the surface of the liquid merely reflected or scattered light without fluorescence. This spectrum will always be visible to a greater or less degree, owing to the fact that even the surface of a transparent liquid scatters a certain amount of light without change of wave-length.

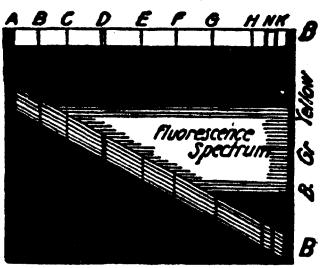


FIG. 377

Variation of the Intensity with the Angle of Emission. — The intensity of the light emitted by self-luminous solid and liquid substances varies as the cosine of the angle of emission (Lambert's law). On this account the intrinsic intensity is not increased by foreshortening of the source, *i.e.* by viewing it in an oblique direc-

tion. This is not, however, the case with fluorescent light, as is indicated by some experiments made by the author.

If a rectangular glass tank, or even a beaker glass, is partly filled with a solution of uranin (fluorescein) and a condenser discharge passed between cadmium electrodes close to the surface, the phenomenon can be very clearly seen. The surface is powerfully fluorescent, and if it be viewed from below, the intrinsic intensity will be found to increase rapidly as the surface is foreshortened, becoming of dazzling brilliancy at grazing emission. If a glass plate is interposed between the spark and the fluid, the effect of foreshortening becomes less marked or disappears entirely, for in this case the fluorescence is chiefly caused by the radiations which penetrate the body of the fluid, and the powerful surface fluorescence, excited by the ultra-violet rays, disappears. A still better method is to illuminate one face of a right-angle prism of crown glass with the light of the spark, which causes a blue fluorescence of the surface layer. The luminous surface is to be viewed through the other face of the prism. The intensity viewed in the normal direction is very slight, as can be seen by looking at the reflection of the luminous surface in the hypotenuse face of the prism. Seen edgewise the intensity is fully thirty times as great, as was found by measurement at an angle of 5 degrees with the surface.

Measurements of the change of intensity with the angle of emission were made in the following way.

A portion of the light from the spark fell upon the prism, while another portion, after traversing a sheet of ground glass and a

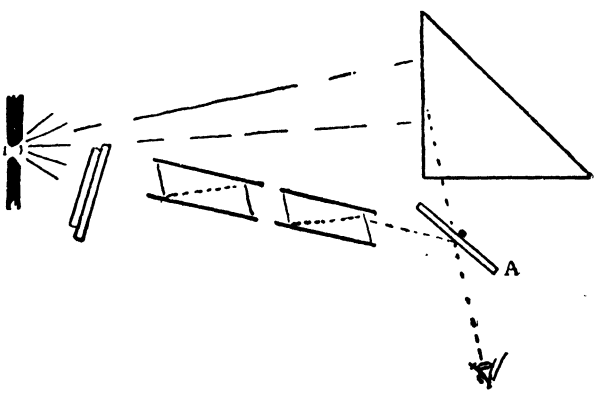


FIG. 378

blue screen, was passed through a pair of Nicol prisms and reflected to the eye by means of the narrow strip of silvered glass A, mounted on a pivot in front of the prism (Fig. 378). By turning one of

the Nicols the intensity of the light seen reflected in the strip could be balanced against the fluorescent background against which it was seen. The color was very accurately matched by means of a thin sheet of cobalt glass combined with a gelatine film stained with one of the blue aniline dyes. The intensity in the normal direction was measured by matching the reflected light against the image of the fluorescent surface seen by total reflection in the prism.

The values obtained are shown graphically in Fig. 379, intensities being plotted as ordinates, and angles as abscissae. The observed values, which have been corrected for the small loss due to internal reflection, are represented by circles, while the curve drawn through them is the theoretical curve, calculated on the assumption that the intensity of the radiation from each fluorescent molecule is independent of the direction within the glass. Under such conditions the intensity of the illumination of the surface would double each time its apparent area was halved by foreshortening.

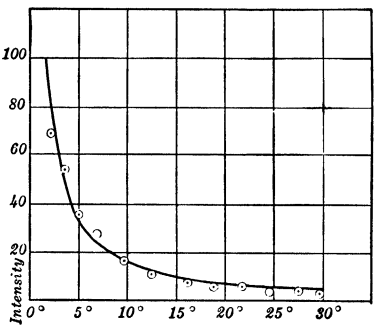


FIG. 379

Effect of the Solvent on the Intensity and Color of Fluorescence. — The solvent exercises a very marked influence upon the intensity of the fluorescence. This subject was carefully investigated by Knoblauch.[1] He found that the intensity of the fluorescent light was strictly proportional to the intensity of the exciting light, but that this intensity varied greatly when the substance was dissolved in different liquids, as is shown in the table on page 658.

The figures indicate the order of intensity, 10 indicating very strong fluorescence, 1 very feeble.

This table is of use in the preparation of fluorescent solutions, and shows us that some solvents are better adapted to certain substances, other solvents, however, to others.

Fluorescence of Inorganic Compounds. — Very few inorganic compounds exhibit fluorescence, though many have as sharply defined absorption spectra as the organic dyes. The salts of uranium, however, equal in brilliancy the most fluorescent of the organic compounds. Only the uranyl salts exhibit the phenomenon and some of these only in the solid state, or in solution in glass or

[1] *Ann. der Phys.*, *54*, 193, 1895.

concentrated sulphuric acid. The water solution is practically non-fluorescent. The absorption spectra of both classes of uranium salts are similar, but only the configurations containing the UO_2 radical are able to reëmit the absorbed energy.

Uranyl sulphate is fluorescent in a water solution, the intensity increasing as the temperature is lowered. It is still brighter in

	Water	Glycerine	Methyl Alcohol	Ethyl Alcohol	Acetone	Isobutyl Alcohol	Amyl Alcohol	Ethyl Ether	Gelatine	Xylol	Toluol	Benzol
Magdala red			4	4			3				1	1
Eosine (sodium)	1	2	6	5			4		3			
Phenosafranine	1	6	7	9	11	9	10			4	3	2
Fluorescein lithium	2	3	5	4			1					
Chrysolin	2	3	3	3		3	1					
Chrysaniline			1	2			3					
Curcumin			1	2		3	4					
Aesculin	3	3	3	3	1	3	2					
B. Phenylnaphthylamin			5	5	3	5	4	2		1	1	1
Phenanthrene				1								2
Anthracene				4	3	5	4	4		5	5	5
Petroleum				5	4			3		6	6	6
Fluorescein			1	2			3					

sulphuric acid, but is destroyed by a trace of HCl or alcohol according to F. Perrin.

Influence of Concentration. — The earliest studies of fluorescent solutions by Stokes showed that the intensity of the emitted radiation was a function of the concentration, rising to a maximum and then falling gradually to zero. The effect is easily shown with fluorescein dissolved in water slightly alkaline with NaOH. One drop of the non-fluorescent concentrated solution added to a litre of water renders the entire mass powerfully fluorescent.

About all that can be said in explanation of this is that while neighboring molecules of the solvent are without influence, close proximity of similar molecules prevent the return transitions from higher to lower energy states in a manner resulting in the emission of radiation, the transitions liberating energy in the form of heat. Jean Perrin observed that the addition of some other dye having an absorption band near the band of the fluorescent substance also reduced the intensity of the fluorescence. This "rayless"

disactivation can also be accomplished by certain uncolored substances as has been shown by Pinnow [1] and F. Perrin [2]. Iodide of potassium, added to a solution of fluorescein destroys the fluorescence, which return if the solution is diluted, showing that no chemical change results. The absorption band is unchanged by the iodide.

It was found by F. Perrin that in viscous solutions the disactivation of fluorescein by iodide of potassium was much less than in a water solution. This he explained as a result of a slowing down of diffusion, the chance of an excited molecule coming within the sphere of action of an iodide molecule being diminished.

Perrin considers that absorbing substances which show no fluorescence may be disactivated by some similar process occurring within the molecule. Erythrosine, which shows a comparatively feeble fluorescence, is a compound of iodine and fluorescein, and he suggested that in this case the iodine ions in the molecule disactivate it in the same way that potassium iodide destroys the emission of fluorescein. Potassium chloride is without effect.

Polarized Fluorescence. — The emission of polarized light by fluorescent solutions escaped detection for many years though it is a common enough phenomenon in the case of gases and vapors, as we have seen.

Weigert was the first to show that fluorescein (uranin) dissolved in glycerine when illuminated with polarized light shows fairly strong polarization which increases as the temperature is lowered. The percentage of polarization for uranin in glycerine and water mixtures of varying composition is shown by the following table:

Water percentage	3	20	33	50	66	83
Polarization per cent	42	29	19	10	6	3

In general powerfully fluorescent substances show appreciable polarization only when dissolved in a viscous liquid, but some of the feebly fluorescent ones, erythrosine for example, exhibit the effect in water solution. The viscosity of the liquid brings out the polarization by keeping the molecule in a fixed orientation during the interval between excitation and emission.

From this the inference follows that a greater degree of polarization in a non-viscous solvent is to be expected in the case of substances for which the excited state is of very short duration, such as erythrosine, with an excited life of only 10^{-9} sec. according to Gaviola's measurement. This is confirmed by experiment. Moreover F. Perrin found that uranin in water, exhibiting only

[1] *Jour. j. prak. chemie, 174*, 265, 1902.
[2] *Comptes Rendus, 184*, 1121, 1927.

2% of polarization, yields 40% of polarization when its fluorescent efficiency is reduced to $\frac{1}{150}$ by the addition of potassium iodide, the unpolarized radiations resulting from the molecules of longer excited life, which are the ones most likely to be disactivated by the iodide.

The fluorescent light emitted in an oblique direction from the surfaces of isotropic media was found by Millikan [1] to be polarized in a manner similar to that of light obliquely emitted by white-hot surfaces. This polarization results from refraction of the light as it passes out of the medium into the air.

Fluorescence of Crystals. — Certain crystals possess the remarkable property of emitting a polarized fluorescence. The most interesting case is that of magnesium platino-cyanide, which can be prepared by the addition of a solution of magnesium sulphate to a solution of barium platino-cyanide, until no further precipitation of barium sulphate takes place. The colorless solution is filtered, evaporated and crystallized.

A large crystal should be prepared by attaching a small one to a very fine glass fiber with wax or shellac and suspending it in the saturated solution in a constant temperature room. The crystals, which are of a deep red color, have most remarkable optical properties, showing a brilliant green surface color on the sides of the prisms, while the ends selectively reflect a deep violet light, which is polarized even at normal incidence. The crystal should be mounted in a small glass tube hermetically sealed, as otherwise it crumbles to a yellow powder. If examined at the polarizing angle with a Nicol, the violet metallic reflection shows to better advantage as the vitreous reflection is abolished. The fluorescence of the crystal can be observed best by exciting it in a concentrated beam of sunlight which has been passed through a piece of dense cobalt glass. We will suppose the crystal to be standing upright, upon one of its bases, and the incident beam horizontal. If the fluorescent light, which emerges from the sides of the prism, be observed through a Nicol prism, it will be found that the color is orange-yellow when the polarizing plane of the Nicol is perpendicular to the axis of the prism, and scarlet when the plane is parallel to the axis. If the exciting light be polarized horizontally, the color of the fluorescent light is yellow, changing to red as the plane of polarization is rotated through 90°. Examination with a second Nicol shows that the orange-yellow fluorescence is polarized perpendicular, the scarlet parallel to the axis of the crystals. If, now, the crystal be placed in a horizontal position, and the beam of unpolarized light directed against an end surface, the

[1] *Phys. Rev.*, September and November, 1895.

fluorescent light will be unpolarized and of a scarlet color. If the incident light be polarized in a vertical plane, and the crystal be turned on a vertical axis so as to vary the angle of incidence, the red color changes to yellow. If, on the other hand, the plane of polarization be horizontal, no change is observed on turning the crystal. This shows that the change from red to yellow takes place as the angle which the direction of vibration makes with the crystal's axis changes from 90° to 0°, the direction of vibration being perpendicular to the plane of polarization.

The results obtained in the two positions of the crystal are in perfect agreement, as will be seen by reference to Fig. 380, in which the exciting rays are indicated by solid arrows, the fluorescent rays by dotted arrows, the direction of vibration in each case being indicated by double-headed arrows.

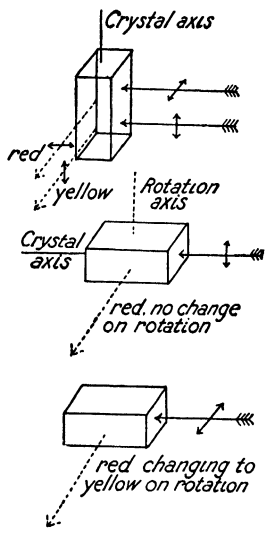


FIG. 380

Fluorescence of the Rare Earths. — The protection against the disturbing influences of neighboring molecules which makes fluorescence possible in the case of the organic colored media and uranium salts probably results from the manner in which the atoms are arranged in the molecule.

In the case of the rare earths which exhibit absorption lines nearly as narrow as those yielded by a gas, it is believed that the absorption results from the deeper electrons in the atom, which are shielded by the outer electron system. Though absorption of light by the commonly observed bands in the visible spectrum is unaccompanied by fluorescence, it was observed by Soret and later by J. Stark, that ultraviolet excitation by the cadmium spark produced emission of blue light in the case of solutions of salts of many of these elements.

Phosphorescence. — This term is applied in general to an emission of light which persists for an appreciable time after the exciting light has been cut off. The duration of the emission may be anything between a millionth of a second and several hours, and it has always been found difficult to draw any sharp line between fluorescence and phosphorescence on this account. Substances such as uranin, which in solution show no persistence of emission become phosphorescent when brought into a solid solution in gelatine.

The term phosphorescent is now restricted to a class of crystalline substances in which very minute traces of metallic impurities give them the power of emitting light for a longer or shorter time after exposure to a strong light. Such substances are called phosphors.

Persistent fluorescence, or fluorescence of long duration is the better term to apply in the case of other substances that remain luminous for a measurable time. Most remarkable is the prolonged emission of Balmain's luminous paint, a sulphide of calcium, which glows in the dark for many hours after exposure to a strong light. It appears probable that something analogous to a chemical change is produced by the action of the light, the condition being unstable and the process reversing as soon as the substance is screened from the action of light. This reversing of the process may be prolonged or rapid, and is accompanied by the emission of light. The energy of the absorbed light is stored in the substance in the form of potential energy of the atoms. At very low temperatures the condition brought about by the action of light may be more or less stable, as was shown by Dewar. A fragment of ammonium platino-cyanide was cooled by means of liquid hydrogen and exposed to a strong light. On removing it to a dark room no trace of phosphorescence was perceived, but on removing the crystal from the chilled tube and allowing it to warm up, it presently burst into a brilliant green phosphorescence.

A partial stability can be shown at ordinary temperatures with Balmain's luminous paint. If this be kept in absolute darkness for twenty-four hours it becomes non-luminous. A further emission of light may, however, be produced by concentrating invisible infra-red radiations upon it. The infra-red radiations have carried on the reverse process, farther than it is able to go spontaneously. Not until the powder has again been exposed to light and kept overnight in the dark can the experiment be repeated. The color of the phosphorescent light obtained in this way differs from that of the light by the spontaneous breaking down of the molecular condition produced by the light, being distinctly greenish instead of deep blue.

Duration of the Phosphorescence. Phosphoroscope. — While some phosphorescent substances remain luminous for a considerable time after their exposure to light, the majority cease to give out visible radiations in a few seconds after the exciting radiations cease to fall upon them. An instrument was devised by Becquerel for examining substances in complete darkness, a small fraction of a second after their exposure to a brilliant light. This instrument, which is known as the phosphoroscope, consists of two metal disks

side by side on the same axle. The disks are perforated with an equal number of apertures which are arranged out of step, and are driven at high speed by a train of cog-wheels. The substance to be examined is placed between the disks, and a strong beam of light directed upon it through the apertures of one of them. If the eye be brought close to the other disk the object will be seen only at the moments when the light beam is cut off, and it will be visible therefore only by phosphorescent light. The disks are mounted in a cylindrical metal box, to screen the substance from all light except that reaching it through the perforations.

With this instrument it is possible to observe an object one one-thousandth of a second or less after its illumination. Becquerel found that phosphorescence was much more common than had been supposed. The salts of the alkali metals, compounds of aluminum, and nearly all organic compounds were found to be phosphorescent. Compounds of the heavy metals for the most part showed no trace of luminosity, the salts of uranium and platinum being marked exceptions, however.

All solid fluorescent substances were found to be phosphorescent A simple phosphoroscope with a single revolving disk can be set up in a very few minutes, and gives excellent results with uranium glass, uranium nitrate, and other phosphorescent substances. The disk, which is 30 or 40 cms. in diameter, can be made of cardboard with holes about half a centimetre in diameter punched at regular intervals around its circumference. The distance between the holes should be about 2 cms., not less. The disk can be mounted on the shaft of a small electric motor, or on a whirling table. A beam of sunlight, reflected from a mirror, is focussed on one of the holes; the diverging cone is received by a second lens on the other side of the disk and again brought to a focus, this time upon the object under examination, *e.g.* a lump of nitrate of uranium. On setting the disk in motion, and viewing the object through the small holes, taking care not to get in the way of light, the phosphorescence can be easily observed. Obviously the eye must be moved about until the position is found in which the incident light is cut off from the object when it is exposed to view. By laying the crystals upon white paper the effect is more striking, for it is then apparent that we are seeing the crystals by their own light alone. A phosphoroscope designed by the author for measuring very brief phosphorescence is described in the paper on mercury fluorescence referred to in the last chapter.[1]

Nature of Phosphorescence: Solid Solutions. — It is now known that, in the case of most phosphorescent substances the

[1] *Proc. Roy. Soc.*, *99*, 362.

mechanism involved in the production of light is a crystal lattice, in which atoms of other elements, usually metallic, are imbedded. The crystal structure appears to be necessary in most cases, for there are many examples of cases in which material of exactly the same composition containing identical impurities can be prepared both in the crystalline and amorphous state, the former only exhibiting an emission of light. A striking example is the class of phosphorescent preparations of boron nitride prepared and investigated by Tiede and Tomaschek.[1] The material is prepared by heating boric acid and urea, and is crystalline or not according to the method of preparation, an excess of boric acid or the addition of sodium chloride promoting crystallization, this flux being subsequently dissolved with water. The phosphorescence in this case is due to carbon atoms imbedded in the nitride lattice, and X-ray photographs by the Debye-Scherer method show crystal structure only in the case of the luminous preparations.

Boric acid can be activated by crystallization with a trace of uranin. In the amorphous or fused condition only the fluorescent band is emitted, identical with that of a water solution, and it is extremely weak in the phosphorescent spectrum (*i.e.* after cessation of illumination). The strong phosphorescent bands appear as soon as crystallization occurs.[2] A similar dual type of photo-luminescence is shown by sugar crystals activated with fluorescent dyes.

If crystallized with a trace of aesculin, uranin, eosin or other fluorescent dye, the crystals exhibit both fluorescence and phosphorescence, the colors being quite different for the two phenomena, as different spectrum bands are concerned in their production. The crystals stained with aesculin are blue during their excitation, but shine with a green light as soon as the illumination ceases. The change is still more marked at $-180°$. With uranin the color changes from green to orange, and with eosin from yellow to red — if the crystals are at $-180°$.

The phosphorescent bands are always of greater wave-length than the fluorescent bands. The table on page 665 gives the wave-lengths for a number of cases, phosphorescent bands being underscored.

Fluorite Phosphorescence. — The word fluorescence is derived from fluor spar, a natural fluoride of calcium, which was one of the first substances studied in connection with optical excitation. Certain specimens, notably those from Weardale, England, were found by Morse to yield highly homogeneous radiations when excited by the light of a spark some of the lines in the visible spectrum being only a few Ångström units in width, while the 5736 line is as narrow

[1] *Zeit. f. Electrochem., 29*, 303, 1923.
[2] E. Tiede, *Phys. Zeit., 22*, 563; *Chem. Ber.. 55*. 588: *Ann. der Phys., 67*, 612.

as the sodium lines in a flame, and at liquid air temperature only 0.1 A.U. in width.

	Fluorescence at +20°	Phosphorescence at +20°	Phosphorescence at −180°
Aesculin			
	443 very strong	490 weak	490 weak
	490 strong	540 strong	540 very strong
	540 weak	600 weak	600 strong
	600 very weak		
Uranin			
	527 strong	527 strong	527 weak
	570 weak	——	570 weak
	640 very weak	——	640 strong
Eosin			
	578 strong	678 strong	
		650 weak	650 strong

We know now that the fluorescence of this mineral is due to small traces of impurities, and Urbain has prepared synthetic fluor spar sensitized with various metals, each one giving a characteristic line spectrum. Excitation is by ultra-violet light from a spark.

Quenching of Phosphorescence by Red and Infra-Red Rays. — Seebeck discovered that orange or red light falling upon a phosphorescent substance, excited to luminosity by violet or blue light, destroyed the phosphorescence completely. This discovery, though recorded in Goethe's *Farbenlehre*, remained comparatively unknown, and the phenomenon was rediscovered many years later by E. Becquerel, who investigated it more fully and found that the quenching of the phosphorescence by orange or red light was preceded by a momentary increase in the intensity of the luminosity, followed by complete darkness. It appeared as if the red rays squeezed out all of the stored energy in a few moments, and he attributed the effect to the heating action of the rays, for he had found that the same thing could be accomplished by warming the substance. He found further that the infra-red rays acted in the same manner and his son, H. Becquerel, continued the investigation, and in 1883 studied the invisible region of the solar spectrum, beyond the red as far as 1.5 μ, by projecting the spectrum on a luminous screen excited to phosphorescence, and observing the darkening produced by the invisible rays. J. W. Draper improved the process by laying the partially darkened phosphorescent screen upon a photographic plate, thus obtaining a permanent record. This method of phosphoro-photography of the infra-red was subsequently employed by Lommel, Dahms and others, but nothing of much importance has been accomplished by it.

Dahms in 1904 made the interesting observation that the luminosity of the zinc sulphide phosphors is quenched by red and infra-red radiations without any momentary increase, which shows that there are two distinct types of quenching. Balmain's luminous paint is an example of the first type. Under infra-red illumination, the phosphorescing material suddenly gives out a greenish light, quite different in color from the violet glow which it emitted before the infra-red rays played upon it. If warmed slightly with a heated glass rod there is also an increase in luminosity, but in this case the color is unaltered. This proves that the action of the infra-red rays is not merely a heating effect. The zinc-sulphide phosphors, which are now extensively used for stage effects, are examples of the second type of quenching. They darken rapidly under the action of infra-red rays without any preliminary increase of luminosity.

Lenard discovered that the absorption spectrum of an excited (*i.e.* luminous) phosphor was different from that of the unexcited. The absorption bands in the ultra-violet disappeared under powerful illumination, and were replaced by new bands in the longer wave-length region. (Compare with Pohl's experiment.) It is from absorption by these new bands that the quenching results. This change in the molecular state also manifests itself in other ways. The magnetic and dielectric properties of the phosphors are found to be different before and after illumination. Elster and Geitel found that both natural and artificial phosphors exhibited the photo-electric effect, *i.e.* they gave off electrons when illuminated, and Lenard, following the matter up, found that the photo-electric emission and the phosphorescence were caused by absorption of radiation of the same wave-lengths. His theory of phosphorescence supposed that the absorption of light resulted in the ejection of an electron from an atom, and its capture by a neighboring atom, the phosphorescence resulting from the gradual return of the captured electrons to their former places, the energy set free in this process being communicated to another electron and eventually radiated as light of longer wave-length.

The modern theory of the phenomenon rests chiefly on the recent work of Gudden and Pohl, who found that many non-phosphorescent (*i.e.* "pure") crystals conduct electricity when illuminated by light. This may be called an "internal photo-electric effect," the electrons set free by the light, travelling through the crystal lattice to the anode, under the influence of the applied electromotive force. It is only in a crystal, the space lattice of which is distorted by foreign atoms (impurities), that the return of the electrons ejected by the light causes an emission of radia-

tion. Experiments on the electrical conductivity of these crystals have shown that the electrons do not traverse the distorted lattice freely, as in "pure" crystals, but that many are captured on their way to the anode by atoms. It has recently been found by Rupp that the electric current, which flows in a phosphorescent crystal (under an applied electric force) during its illumination, and ceases as soon as the illuminating rays are cut off, starts up again if the crystal is warmed or illuminated by infra-red rays, and that the amount of electricity transported is proportional to the amount of light emitted.

If a "pure" crystal is illuminated in the absence of an electric field the electrons ejected from atoms by the light move about freely in the lattice, and the local fields, which result from the ejections, recapture them as soon as the illumination ceases, and no current will flow under an applied electric force when the crystal is warmed. In the distorted lattice of a phosphorescent crystal, on the contrary, the ejected electrons are captured by other atoms, and if such a crystal is illuminated for a definite time, and the photo-electric current, which flows during the illumination is measured, an equal amount of electricity will flow, on cutting off the illumination and warming the crystal.

Imprisonment of Fluorescent Radiation. — Cooper-Hewitt, in the course of experiments on the preparation of fluorescent reflectors to add red light to the illumination produced by the mercury arc, observed that the fluorescence was much more powerful when the celluloid lacquer, stained with rhodamine, was backed by white paper than when spread on a polished silver surface. This phenomenon was investigated by the author [1] and found to be due to the circumstance that in the latter case the fluorescent radiation was imprisoned to a large extent by total internal reflection, while its release was brought about by the diffuse reflection of the white paper. If we plot the paths of rays leaving a molecule at the centre of a transparent plate whether free or backed by silver, we find that all rays incident on the surface at an angle exceeding the critical angle will be multiply reflected in the medium. While in the case of the backing of white paper these rays, reflected from the upper surface, are diffused over an angle of 180° by the paper and by the repetitions of this process all of the light eventually escapes. Small flakes of celluloid or gelatine stained with fluorescent dyes, shine with greatly enhanced brilliancy along the edges, as the imprisoned radiation can escape at the edge. The solutions are poured on slightly greasy glass, and the dried film peeled off.

[1] *Phil. Mag.*, *25*, 449, 1913.

CHAPTER XXI

MAGNETO-OPTICS

Magneto-optical phenomena can be divided into two distinct classes, the first comprising those in which the source of light is acted upon by a magnetic field, which effects a change of wavelength accompanied by polarization (Zeeman effect) and those in which the velocity of light and its state of polarization are modified when it is passed through a magnetized medium (Faraday effect). These effects are very closely related and are best studied simultaneously, for the theoretical treatment of the Faraday effect involves the use of theorems developed for the explanation of the Zeeman effect, while the presence of the Faraday effect in some instances shows the existence of a Zeeman effect too small to be detected spectroscopically. The first observation of an effect of a magnetic field on optical phenomena was made by Michael Faraday in 1846. After many unsuccessful efforts to find a relation between magnetism and light he finally reached his goal. Placing a block of heavy lead glass between the poles of an electromagnet, and passing plane-polarized light through the block in a direction parallel to the lines of force he found that the plane of polarization was rotated by the magnetized medium.

Early investigations of the subject were made with solid and liquid substances, and it is only within recent years that the importance of the study of the effect with gases, especially those having narrow absorption bands, has been realized. From a theoretical point of view a new phenomenon is most profitably studied with matter in its simplest state, a condition best fulfilled by a monatomic gas such as the vapors of the alkali metals and mercury, the rare gases or atomic hydrogen. We shall, however, follow the historical development of the subject at first, reserving the more recently discovered cases, and their theoretical treatment until after the discussion of the Zeeman effect.

The Zeeman Effect. — A search for a possible effect of a powerful magnetic field upon a source of monochromatic radiations was first made by Faraday. Placing a sodium flame between the pole pieces of a large electromagnet, he examined the appearance of the *D* lines when the field was "on" and "off." He was unable, however, to convince himself that any change resulted in the ap-

pearance of the lines, a circumstance which we now know resulted from the insufficient power of his spectroscope. This was in 1862. Twenty-three years later the same experiment was tried by Fievez, who observed changes which appeared to be the result of the magnetic field. If the lines were single before the current was turned into the magnet, they appeared with dark reversals down their centres as soon as the magnet was excited. These observations do not appear to have attracted much attention, and they were not followed up. It is questionable even whether the phenomenon observed by Fievez was in reality the effect that we are about to study.

In 1896 the fact was established beyond all doubt by Zeeman, that the periodic time of vibration of a source of monochromatic radiations is altered when the source is placed in a magnetic field.

In his first announcement he described the effect as a simple broadening of the lines, but at the suggestion of Lorentz, whose theory predicted that the edges of the line should appear circularly polarized when the light was emitted in the direction of the lines of force, he employed a magnet with perforated pole pieces, and passed the light through a quarter-wave plate, which as we have seen transforms a circular vibration into a plane one. He found that the lines now shifted their position in the spectrum when the direction of the current in the magnet was reversed, which indicated that the edges of the line were circularly polarized in opposite directions, exactly as theory showed that they should be.

When the light was examined in a direction perpendicular to the lines of force, the edges were found to be plane-polarized, from which Zeeman drew the conclusion that, with a sufficiently strong field, the line would appear triple, a prediction which was speedily verified by experiment, the electric vector being perpendicular to the lines of force in outer components and parallel to the field in the inner component.

This phenomenon was easily explained on the electron theory, at least the simple case just described. A further study has shown, however, that many lines are broken up into quadruplets and sextuplets, the D lines of sodium belonging to this class. In a powerful field D_1 is seen to be a quadruple line, the inner components being polarized along the lines of force, the outer components perpendicular to them. D_2 is a sextuplet, with its four outer components polarized perpendicular to the field and the two inner parallel to the field.

The Zeeman Effect and Spectral Series. — As we shall see, spectrum lines are divided in all sorts of ways in a magnetic field.

the normal triplet being the exception rather than the rule; the normal triplet occurs however in hundreds of examples in the spectra of zirconium and titanium. A great number of lines of the iron spectrum also become triplets under the action of the magnetic field.

The measurements of Runge and Paschen and others have shown, however, that in the case of lines belonging to the same spectral series, the type of magnetic separation is the same, *i.e.* all lines of a given spectrum that can be represented by a formula are divided in the same way. As Lorentz says:

"In those series which consist of triplets or doublets, the mode of division of the lines is in general different for the lines of one and the same triplet or pair, but, according to the law just mentioned, the same mode of division repeats itself in every triplet or every doublet. Thus, in each triplet belonging to the second

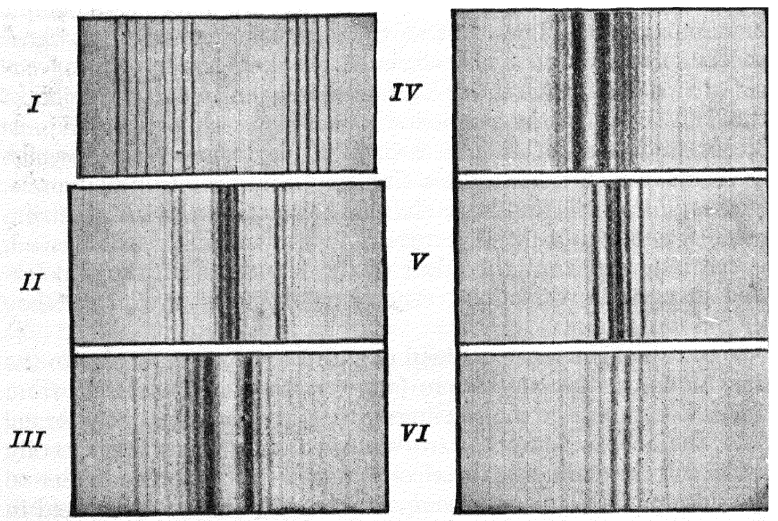


FIG. 381

subordinate series of mercury, the less refrangible line is split into nine components, the middle line into six, and the most refrangible line into three components.

"Equal modes of division are found not only in the different lines of one and the same series, but also in the corresponding series of different elements. For the lines D_1 and D_2 of sodium, which form the first member of the principal series, are changed into a quartet and a sextet, and the first terms in the principal series of copper and silver present exactly the same division."

A number of these interesting cases are shown in Fig. 381, which

is taken from Voigt's most interesting book on *Magneto and Electro Optics*. No. 1 is a photograph of the molybdenum line 4269.4 taken in the second order spectrum. On the left, we have the components, eight in number, in a direction perpendicular to the lines of force, and on the right, the five components seen parallel to the field. This photograph was taken in Professor Voigt's laboratory, while the six which follow were made by Dr. Lohmann in Halle with an echelon grating. No. 2 shows the neon line 6600 in a magnetic field of 10,800 Gauss. No. 3 is the 6718 line (field strength 5700). No. 4, $\lambda = 6507$, $H = 9950$. No. 5, $\lambda = 6143$, $H = 9300$. In No. 6, we have a decomposition into 12 components ($\lambda = 6335$, $H = 8420$).

Study of the Zeeman Effect without a Spectroscope. — A very convenient and simple method of showing the change in the wave-length is that employed by Cotton [1] and Koenig.[2] The spectroscope is dispensed with entirely, and an absorbing flame put in its place. As is well known, the sodium flame has the power of absorbing, strongly, radiations of the same wave-length as those which it emits. If we place a bright sodium flame between the poles of an electromagnet and in front of it a second sodium flame, preferably a less luminous flame, such as can be obtained by burning a jet of illuminating gas at the tip of a piece of soft glass tubing drawn down to a point, the light which is emitted by the first flame will be partially absorbed by the second, which will appear dark in consequence. If the second flame contains much sodium, it may happen that its edges only appear dark.

If now the magnet is excited, the wave-lengths emitted by the first flame are changed, and the second flame, no longer able to absorb them, brightens up in consequence, or, to be more exact, it no longer appears darker than the background. It is best to try the experiment first along the lines of force, using perforated pole pieces, and placing the second flame close to one end of the magnet. The phenomenon is less marked in a direction perpendicular to the lines of force, since in this case the emission line breaks up into a triplet, and the second flame is capable of absorbing completely the middle component, the period of which is unchanged. This central component is, however, plane-polarized and can be cut out by means of a Nicol prism, under which conditions the brightening up of the flame is quite as conspicuous as in the direction of the lines of force.

Magnitude of the Separation. — Spectroscopic recognition of the Zeeman phenomenon requires an instrument of high resolving-

[1] *Compt. Rend.*, cxxv, 865.
[2] *Wied. Ann.*, *63*, 268.

power. The most convenient form is undoubtedly the echelon grating devised by Michelson; it can be adjusted in a few minutes, gives a large amount of light and is fairly compact. The author has seen the effect without difficulty with an echelon improvised from four interferometer plates which were mounted on the table of a spectrometer with a step-width of 1 mm. A small screen of cardboard with a rectangular opening 5 mms. wide and 2 cms. high limited the beam coming from the collimator. This was mounted in such a position as to have a strip 1 mm. wide along the edge of the grating, five interfering beams, with high relative retardations, being obtained in this way. The spectrometer was illuminated with the light from a mercury vacuum tube placed between the poles of an electromagnet and a small direct-vision prism placed between the echelon and the telescope to separate the mercury lines. The splitting up of the green line on exciting the magnet was easily seen with this improvised apparatus. A helium tube is equally satisfactory and less troublesome, as it requires no heating. The sodium flame is unsatisfactory, owing to the small distance between the D lines.

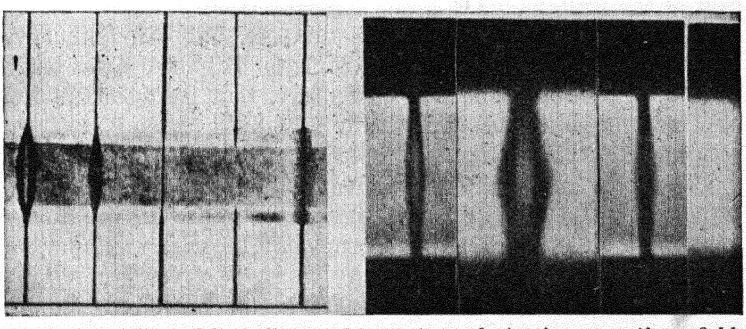
The results obtained with the echelon are not, however, easy to interpret, and the proximity of the spectra of other orders gives trouble, when anything more than a qualitative experiment is to be made. For accurate quantitative work the concave grating is undoubtedly the best type of instrument to use.

The magnitude of the separation, even for very intense fields, is very slight. Zeeman concluded from measurements of photographs that for a field-strength of 10,000 c.g.s. units the distance between the outer components of the D_1 quadruplet amounted to $\frac{1}{13}$ of the distance between D_1 and D_2. The same separation will of course be found in the case of the two circular components seen along the lines of force.

The Inverse Zeeman Effect. — The division of absorption lines into triplets and doublets when the absorbing medium is placed in a magnetic field is known as the inverse Zeeman effect. It results from the circumstance that the vibrating mechanism which gives rise to emission is also responsible for absorption, and the natural frequencies are affected by the magnetic field in the same way in the two cases. The amount of absorption produced by either component of a doublet or triplet, however, depends upon the state of polarization of the white light from which certain wave-lengths are abstracted by absorption. If the white light is circularly polarized, and traverses the magnetized absorbing medium parallel to the lines of force, one component of the doublet will be absolutely black, while the other will be invisible. If the

direction of rotation be reversed, the other component of the doublet appears and the first fades away. The same is true with the triplet and plane-polarized light. If the light is unpolarized to start with, all of the components appear, but they are not black, since only a portion (namely, that polarized in the proper way) is absorbed by the medium.

The Zeeman Effect on the Sun. — One of the most brilliant discoveries ever made in Astrophysics was G. E. Hale's observation of the Zeeman effect in the spectrum of sun-spots. His photographs, made with the spectro-heliograph, by means of which an image of the solar surface is made solely by light of a single wave-length, showed that in many cases the luminous matter surrounding the spot was rotating at high speed in the form of a vortex. If the solar gases contained free electrons, their rotation should produce a magnetic field, and the light emitted by the spot should be circularly polarized, or in the case of absorption lines we should have the polarized doublet described in the previous section. Observation of the spectrum through a circular analyzer



Sun-spot lines (Mitchell). Magnetic resolution in non-uniform field.

FIG. 382

showed that the absorption line was displaced by the rotation of the $\lambda/4$ plate, proving the existence of the Zeeman doublet. The triplet was discovered with a Nicol prism in the case of spots near the edge of the sun, in which case the lines of magnetic force were nearly perpendicular to the direction in which the light was emitted. This discovery of the existence of magnetic fields of vast size on the surface of the sun, produced by whirling vortices of electrically charged matter, comparable in intensity with those which we can produce in the laboratory by the aid of every instrumental refinement over an area of only a few millimetres, is one of the most sensational discoveries ever made.

Zeeman has published in his paper two photographs which illus-

trate in a striking manner the resemblance between photographs of the absorption lines in the spectra of sun-spots and the magnetically divided lines obtained when the absorbing vapor is placed in a non-uniform field. The intensity of the magnetic field is greatest at the centre of the spot, for it is here that the velocity of the whirl of charged electrons is greatest. These photographs are reproduced in Fig. 382. The left-hand one represents a portion of the solar spectrum, the dark band across the centre representing the spectrum of the spot. The widening of the Fraunhofer lines as they cut across the spot is strikingly analogous to the appearance of the lines obtained by Zeeman in a non-uniform magnetic field.

Classical Theory of the Zeeman Effect. — The classical theory of the Zeeman effect is based primarily on the deflecting force exerted by a magnetic field on a moving electron. This force acts always at right angles to the direction of the field and the path of the electron, and hence cannot alter the velocity of the latter.

Lorentz developed a theory which explained perfectly the division of the line into three plane-polarized components, for radiations perpendicular to the magnetic field, and two circularly polarized components parallel to the field. His theory was based on the conception, then held universally, that light consisted of electromagnetic waves emitted by elastically bound electrons rotating in orbits or vibrating to-and-fro along a line.

The theory of Lorentz is of great interest from an historical standpoint, and some of the conceptions which he introduced have been retained in the more modern theories. It may be summarized as follows:

In the case of orbits perpendicular to the field the magnetic force acting on the electron will be directed towards or away from the centre of rotation. Those which are revolving in such a direction that the force acts towards the centre are drawn in, and have their period of revolution accelerated, while those rotating in opposite directions have their centripetal force diminished and their period slowed down. In the direction of the magnetic field the two sets of oppositely rotating electrons will radiate circularly polarized light of higher and lower frequency than that of the light radiated in the absence of the field, giving the so-called Zeeman doublet seen along the lines of force.

Observations by Zeeman and König of the direction of rotation of the circularly polarized components showed that they were due to the action of the field on negatively charged particles or electrons in the manner indicated.

In a direction perpendicular to the field these same rotating

electrons give off plane-polarized radiations of the same frequency as that of the circularly polarized rays, the orbits being foreshortened to lines when viewed in this direction and the electrons appearing to travel to-and-fro along the line. These radiations form the outer components of the Zeeman triplet seen in directions perpendicular to the field, with directions of vibration perpendicular to the field. We must now account for the central component of the triplet which is polarized with its electric vector or direction of vibration parallel to the lines of force.

If we assume linear vibrations of electrons as sources of radiation, the component of unaltered frequency is at once accounted for by vibrators in which the electrons are moving to-and-fro parallel to the field. These will give the middle component of the triplet with proper polarization, while they will emit no light at all in the direction of their motion, *i.e.* parallel to the field.

If the electrons are vibrating to-and-fro along lines perpendicular to the field, the path ceases to be linear, for the electron is continually subjected to the deflecting force of the magnetic field, and its path is transformed into a rosette as shown by Fig. 383 when viewed along the lines of magnetic force. As a matter of fact there would be an enormous number of petals in the rosette, so that we could regard the motion as a plane-polarized vibration in rotation, which, as was shown in the Chapter on Polarization, can be regarded as the resultant of two oppositely polarized circular vibrations of different periods. These may be regarded as the origin of the circularly polarized doublets seen along the lines of force. In a direction perpendicular to the field from which the rosette appears foreshortened to a line we have plane-polarized radiations of rapidly varying intensity, for there are moments at which the electron is travelling to-and-fro on petals of the rosette directed along the line of radiation, and at these moments no radiation will be emitted in this direction, while at intermediate moments plane-polarized vibrations will be sent out.

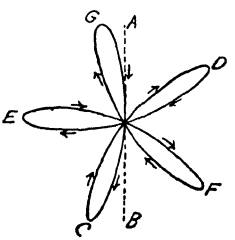


FIG. 383

Now a similar state of affairs results when two periodic disturbances of slightly different period pass through a point simultaneously. When the phases agree we have maximum illumination, and when they are opposed destructive interference results in darkness, and the spectroscope resolves the radiation of rapidly fluctuating intensity into two lines of frequency such as to give "beats" of the same period as that of the rosette.

In the strongest fields known the separation amounts to $\frac{1}{6}$ of the distance between D_1 and D_2, which means that each component moves $\frac{1}{12}$ of this distance from the original position of the line. Taking the frequency differences for the D lines, and dividing this by 12, gives us the frequency difference corresponding to the shift. If n is the normal frequency of the light, and the plane of polarization turns N times per second, the two component circular vibrations into which it can be decomposed have frequencies $(n+N)$ and $(n-N)$. The frequency difference in the above case turns out to be roughly 40,000,000,000, and if we divide this number into the original frequency, we shall obtain a number representing the number of to-and-fro excursions made by the electron while its path turns through one complete revolution. This number turns out to be 15,000, which gives us the number of loops in the star-shaped orbit previously figured.

Beats result when two continuous trains of uniform amplitude, and slightly different frequencies, pass through a point simultaneously. The components, therefore, into which the spectroscope analyzes the disturbance must be continuous in time, *i.e.* must show no fluctuations in intensity. How now is it possible to have continuous illumination in the spectroscope when there are moments at which the slit is in darkness?

An effect at the slit occupying an infinitesimal of time is by the action of the grating or prism spread out over a finite interval of time when it reaches the eye. In the case of the grating there is no difficulty, since the disturbances from the different grating elements, resulting from a *single* disturbance at the slit, reach the eye in succession. This being true it is obvious that what occurs at the eye at a given *instant* is the result of disturbances lasting for a finite time at the slit. There is thus no trouble about having a continuous disturbance at the eye, where there is a discontinuous disturbance at the slit.

In the case of circular orbits parallel to the field Lorentz considered the circular vibration as the resultant of two linear vibrations with a phase-difference of 90°, and then treated these separately in the same manner as the linear vibrations. These latter treatments are not as convincing as the one involving orbits perpendicular to the field. Moreover, they make both or all component lines of the triplet originate in a single atom, whereas the modern theories of radiation require that a single atom radiate a single frequency only. We will now consider an important theorem due to Larmor.

The Larmor Precession. — The action of a magnetic field on an electron in orbital motion was stated in a very complete and ele-

gant manner by Larmor, who showed that a uniform magnetic field of strength H, leaves the form and inclination of the orbit to the lines of magnetic force, unaltered, and the frequency and velocity of the electron in its orbit unchanged but produces a uniform rotation or "Precession" of the orbit about an axis parallel to the field.

The frequency of the precession, while extremely rapid (about a million rotations per second in the earth's magnetic field of 1.5 Gauss) is very slow in comparison to the electronic frequency.

In the strongest fields usually obtainable (40,000 Gauss) the precession frequency would correspond to that of radiation of a wave-length of 5.3 mms.

On the older theory of radiation the combination of the frequency of the precession n with the orbital frequency of the electron N, would give rise to two new frequencies $N+n$ and $N-n$, and two new lines would appear with wave-lengths corresponding to these frequencies. (See section on Light-Beats and Modulation of Light in the Chapter on Interference.) This interpretation of the effects of the precession is no longer held, however.

While Larmor's treatment proves that the precession will continue once it is established, it does not show very clearly how it originates. We cannot compare the electron in its orbit to a spinning top, which "precesses" about a cone, when gravity begins to pull it out of the vertical, for the top is a rigid body and the forces acting at any given points affect the motion of all other points. There is no difficulty in seeing how the precession originates and where the increased energy due to the precession comes from, in the case of an orbit already established when the magnetic field is created; we have only to take account of the electromotive force which is generated by the increasing field.

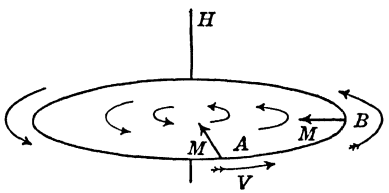


FIG. 384

We will begin with the case of an electron moving with a velocity V in a circular orbit placed at the centre of the magnetic field produced by a solenoid carrying a current, the plane of the orbit being perpendicular to the lines of force. We now introduce the new factor, namely, the electric force generated when the magnetic field is brought into existence. In this case (Fig. 384) the lines of electric force are circles surrounding the centre of the electron orbit, since we have assumed it placed at the centre of the field, as the calculation is thereby simplified. The electric force being every-

where parallel to the orbit will accelerate the electron giving it a velocity increment which we will assume to have reached the value v when the magnetic field is fully established and the electric force disappears. The original centrifugal force of the electron of mass m, moving in an orbit of radius r was mV^2/r and the present centrifugal force is

$$\frac{m(V+v)^2}{r}=\frac{mV^2}{r}+\frac{2mVv}{r}+\frac{mv^2}{r}$$

the increment of centrifugal force consisting of two parts mv^2/r, called the true centrifugal force and negligible in comparison to the second part, the compound centrifugal, or Coriolis force $2mVv/r$ which will be more fully defined presently. The reader should note that in this term the small increment in velocity v is multiplied by the very large value $2V$, in other words the increment in centrifugal force produced by increasing the velocity by a small fixed amount becomes greater and greater with increasing initial velocity V.

As a result of the increased electronic velocity the orbit would expand were it not for the force which the magnetic field exerts on the electron. This force is perpendicular to the field and the direction of flight of the electron, and is consequently directed towards (or away from) the centre of the orbit. It compensates the increased centrifugal force due to the velocity increment and holds the diameter of the orbit fixed at its original value. If originally the electron were travelling around the orbit in the opposite direction, the electric force would retard it, and the centrifugal force would be diminished, while the magnetic force would act radially outwards and prevent contraction of the orbit. We thus have the orbits of the same diameter as in the beginning, but increased and diminished electronic velocity, and consequently altered energy, as required by the Bohr theory of radiation. The "precession" is present in this case disguised as increased electronic velocity, the orbit rotating in its own plane so to speak.

If the orbit is not at the centre of the field, the electric force will accelerate or retard the electron according to its position in its orbit, but there will be an outstanding positive effect, due to the circumstance that the electric force is less on one side of the orbit than on the other. Or we may say that the increment in velocity of the electron is proportional to the magnetic flux through the orbit.

This is independent of the orbit's place in the field, but depends upon its inclination, for as it is turned out of the position in which we have considered it, it is cut by fewer and fewer lines of magnetic

force, until, when it is parallel to the field there is no flux through it at all and no considerable energy increment can be given to it by the electric force. There is, however, a small second order effect, as we shall see when we consider the case of inclined orbits.

This dependence of the energy absorbed from the field upon the orientation of the orbit is very important and forms the foundation of the Quantum theory of the Zeeman effect.

The Bohr theory of radiation considers that light is emitted by an atom only when an electron jumps from an outer to an inner orbit, the wave-length or frequency depending upon the difference of total electronic energy in the two cases. On this theory a given atom can emit at one time a single frequency only, *i.e.* give rise to a single spectrum line. The Lorentz theory of the Zeeman effect, as we have seen, predicts, in some cases, that two or even three lines, *i.e.* all of the components of the Zeeman triplet are emitted by a single atom (orbit parallel to the magnetic field). It is thus fundamentally at variance with the Bohr theory of radiation.

To explain the simple types of Zeeman effect on the Bohr theory we must show that the magnetic field can increase (or diminish) the energy of the electron in all of its stationary orbits.

We can deduce the action of the field on any one of the orbits by considering the energy increment or decrement contributed by the E.M.F. of the increasing field, which manifests itself as precession.

In this manner we calculate the energy levels. If now we excite the gas in a discharge tube in a magnetic field already established (the usual procedure) the electrons are driven from the lower orbit to the upper orbits *which must precess,* consequently more energy or less energy has to be applied by the impacting electron coming from the cathode than in the absence of the field. In this case there is no E.M.F. due to a magnetic flux, as the field was in existence initially, and we can consider that the energy of the precession is supplied by the impacting electron. In other words it requires more energy (or less energy for opposite precession) to raise an electron to an outer orbit *which must precess,* than to one which can remain stationary.

The angular velocity of the Larmor precession depends only on the field-strength, but the energy change depends upon the orientation of the orbit. As we shall see this results from the circumstance that the major part of the energy change is due to the component of precession velocity which is parallel to the orbit of the electron, for it is this component that supplies the term containing the product Vv in which V, the electron velocity, is very large in com-

parison to v. If the plane of the orbit is nearly parallel to H the component of v parallel to the orbit becomes small: if exactly parallel to the field, the orbit rotates about its diameter, but there is no change of energy (neglecting a second order effect).

To make clearer the matter of the altered energies of the outer orbits, to which the electron is to be raised in a constant magnetic field, devoid of an electromotive force, consider once more the case of the circular orbit perpendicular to the field, the atom supposed to be in the unexcited or normal state. We have seen that the E.M.F. due to the creation of the field, increases (or decreases) the electronic velocity, the diameter of the orbit remaining the same. We now excite the gas, say by electronic impact in a discharge tube, and drive the electron to an outer orbit, which we will assume to be circular and perpendicular to the field, *and to have the same diameter as in the absence of the field*, this latter condition being imposed by the theorem that the diameters of precessing orbits are not altered.

But the electron must be raised to this orbit *in a magnetic field*, and the force exerted by the field on the moving electron will alter the diameter of the orbit, unless we compensate the effect by giving it a greater (or less) orbital velocity, in other words establish a precession, which, as we have seen in this case, is present though disguised as rotation of the orbit in its own plane. The increment of energy necessary for the precession is supplied by the impacting electron.

Stating the thing briefly we can say, that the only orbits which can exist in a magnetic field of strength H are those calculated from normal orbits subjected to the E.M.F. of a field increasing from zero to the value H.

We will now consider the case of inclined orbits, considering more fully the Coriolis force, or the inertia resistance of the moving electron, which we shall make use of in the calculation of the angular velocity of the Larmor precession.

This is the force that comes into play when an attempt is made to alter the direction of the path of a moving particle. Its direction at any instant is perpendicular to the plane passing through the electron, and parallel to its direction of motion and the axis of rotation or precession. Its magnitude $y=2mv\omega \sin \alpha$ in which α is the angle between the direction of motion and the axis of rotation, v the velocity of the particle and ω the angular velocity of precession. y is perpendicular to v and ω.

In the case of the orbital electron, this force opposes the efforts of the electric field to rotate the orbit, but the force exerted on the moving electron by the magnetic field is always opposite and equal

to the Coriolis force, consequently the orbit rotates freely, and continues to rotate with uniform velocity, after the magnetic field has reached its steady state and the electric force has disappeared. The precessing orbit thus becomes a natural one in a magnetic field, but it must be clearly understood that the force exerted by a constant field on the moving electron is always perpendicular to its line of flight and hence cannot alter its velocity.

We must now calculate the angular velocity of the precession as we require this for the determination of the energy changes produced by the field.

Larmor gave this equation for the angular velocity

$$\omega = \frac{1}{2}\frac{e}{m'}\frac{H}{c}$$

in which e is the charge of the electron in electrostatic units, m' its mass, and H the field measured in Gauss, and c the velocity of light. This expression will be developed presently. It should be noted that the angular velocity is independent of the orientation of the orbit. If V is the linear velocity of the electron before the creation of the field, and R its distance from the axis of the Larmor precession, and ω the angular velocity of the precession, the effect of the precession results in a centrifugal force $m\omega^2 R$, and a compound centrifugal force (Coriolis force) proportional to the vector product $[V\omega]$ parallel to the direction of this vector, and of magnitude $2m[V\omega]$. This direction is mutually perpendicular to the vectors V and ω, V being the direction in which the electron is moving at a given instant, and ω the direction of the axis about which the precessional rotation is taking place.

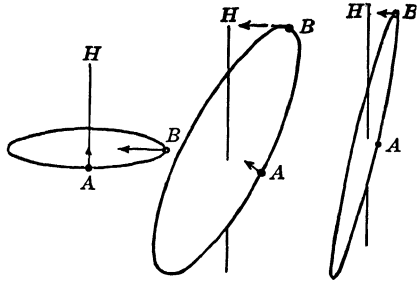


FIG. 385

The direction of the magnetic force exerted on the electron in two orbital positions A and B for three different orientations of the orbit is shown in Fig. 385. For the left-hand orbit, the force is directed towards the centre of the orbit and is equal for positions A and B.

For the middle figure the force at A is directed towards the centre as before, while at B the force is directed as shown, cutting H at some distance above the orbit's centre. In the right-hand figure the force at A has become very small since the electron at

this point is moving nearly parallel to the field, while at B it has the same value as before and is directed towards a point on H still higher up.

The centrifugal force $m\omega^2 R$ depends upon the distance of the electron from the axis of the Larmor precession and consequently varies from moment to moment as the electron traverses its orbit (except in the case of an orbit perpendicular to the field).

It is negligible in comparison with the Coriolis force as it depends upon the square of the small velocity of the precession, while the latter depends upon the product of ω with the large orbital velocity. The Coriolis force can be proved equal to, and opposite in direction from, that of the magnetic force exerted upon the moving electron, consequently the precession will take place with an angular velocity sufficient to make the two forces balance. The force exerted by a magnetic field of intensity H upon an electron moving with velocity V is

$$K = -\frac{e}{c}[VH]$$

the direction of the vector product being perpendicular to both V and H.

Equating the Coriolis force to K gives

$$2m[V\omega] = \frac{e}{c}[VH]$$

$$2m\omega = \frac{e}{c}H$$

$$\omega = \frac{e}{2mc}H$$

the angular velocity of the Larmor precession.

It is important to note that, while the angular velocity of the precession is independent of the orientation of the orbit, the energy acquired (or lost) by the electron depends upon the orientation being a maximum when the orbit is perpendicular to the field.

The Quantum Theory of the Normal Zeeman Effect. — We must now develop an expression for the energy added to or subtracted from the original energy of the electron by the magnetic field, required by the Bohr theory for the production of the new frequencies observed in the Zeeman effect. The original frequency ν of the radiation is given on Bohr's theory by the expression

$$h\nu = E_o - E_i \quad \text{or} \quad \nu = \frac{E_o}{h} - \frac{E_i}{h} \tag{1}$$

in which h is Planck's constant and E_o and E_i the energies of the electron in its outer and inner orbit respectively, a quantum of

radiation of frequency ν being emitted when the electron jumps from the outer to the inner orbit, $-E_o/h$ and $-E_i/h$ being defined as the spectral terms in frequencies sec.$^{-1}$. The Zeeman lines of altered wave-lengths must originate in the same way, and we must show how the magnetic field splits up these spectral terms. These new terms correspond to increased or decreased total energies of the electrons in their orbits.

The source of energy is the electric force developed when the magnetic field is increasing or decreasing as we have seen. The new frequencies

$$\nu+\Delta\nu=\frac{E_o+\Delta E_o}{h}-\frac{E_i+\Delta E_i}{h} \tag{2}$$

(in which $\Delta\nu$ can be positive or negative) and by subtraction of (1)

Δ_n (the frequency shift) $=\dfrac{\Delta E_o}{h}-\dfrac{\Delta E_1}{h}=\Delta\nu_o-\Delta\nu_i$ as indicated in Fig. 386, $\Delta\nu_o$ and $\Delta\nu_i$ being the Zeeman terms of the outer and inner orbits, measured from the level when no magnetic field is present. We shall now develop an expression giving ΔE and $\Delta\nu$ in terms of the angular velocity of the Larmor precession and the moment

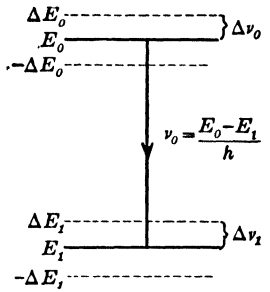


FIG. 386

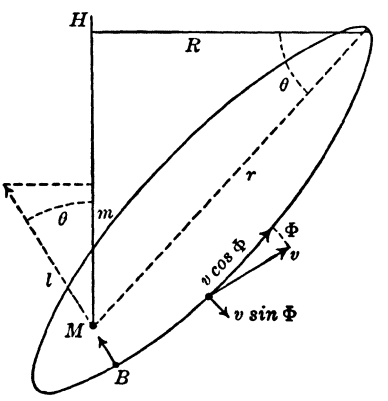


FIG. 387

of momentum of the electron in its orbit, or rather the projection of this vector on a line parallel to the field.

We shall consider the electron in this case as moving in an elliptical orbit around the nucleus M (which is at one focus of the ellipse) this being the more general case (Fig. 387). The quantum vector l, which must take integral values, and which, when multiplied by $h/2\pi$ gives the moment of momentum, is perpendicular to the orbital plane at M and describes the same Larmor precession as the orbit, which rotates about the line MH parallel to

the magnetic field. Let V be the velocity of the electron in its orbit. During the creation of the field, the electric force (assumed as before to be circular around MH) will start a rotation of the electron about MH, a motion which we can regard as independent of the orbital motion. This rotation would increase the kinetic energy of the electron, and its orbit would enlarge, were it not for the magnetic force, which balances the increased centrifugal force due to the superimposed rotation, and the orbital diameter remains fixed but the orbit as a whole rotates about MH. The energy of the precession comes from the electric field, while the magnetic field enables the precessional rotation to continue after the electric field has disappeared.

We will now derive an expression for the increase in energy which results from the field. Let V = orbital velocity, v = velocity of precession, and m' the mass of the electron. The total kinetic energy in the field is

$$E_{\text{kin}}=\frac{m'}{2}(V+v\cos\Phi)^2+\frac{m'}{2}(v\sin\Phi)^2=\frac{m'}{2}V^2+m'Vv\cos\Phi+\frac{m'}{2}v^2$$

in which v has been resolved into two components $v\cos\Phi$, and $v\sin\Phi$, parallel and perpendicular respectively to the direction of the orbital motion of the electron.

Now $(m'/2)V^2$ was the original kinetic energy, consequently the energy increment is $\Delta E=m'Vv\cos\Phi$ neglecting the last term which is small.

Here again we see that it is the product of the small velocity v and the large velocity V that gives the increment in kinetic energy.

Now $v=\omega R$ in which ω is the angular velocity of the precession and R the perpendicular distance of the electron from MH

$$\therefore \Delta E=m'V\omega R\cos\Phi.$$

Now $$R\cos\Phi=r\cos\theta.$$

$$\therefore \Delta E=m'V\omega r\cos\theta$$

and

$$\Delta E=l\frac{h}{2\pi}\omega\cos\theta \qquad \text{since } l\frac{h}{2\pi}=m'Vr.$$

The total energy $E+\Delta E$ will be a maximum if the electron is rotating in the same direction as the precession ($\cos\theta=1$) and a minimum if it rotates in the opposite direction ($\cos\theta=-1$) and vector l in direction opposite to that shown in Fig. 387.

Now $l\cos\theta=m$, the projection of l on the direction of the magnetic field

$$\therefore \Delta E=\omega m\frac{h}{2\pi}$$

and m will vary from $+l$ to $-l$ according to the angle θ, the inclination of the orbit. m is called the magnetic quantum number and can take only the values l, $(l-1)$, $(l-2)\cdots-(l-2)$, $-(l-1)$, $-l$ as is shown by experiment. Vectors l and m are measured in quantum units $h/2\pi$. Writing the equation in quantum numbers gives

$$\Delta E = \omega m \frac{h}{2\pi} \qquad \therefore \Delta\nu = \frac{\Delta E}{h} = \frac{\omega}{2\pi} m$$

$$\Delta\nu = \Delta\nu_0 - \Delta\nu_1 = \frac{\omega}{2\pi}(m_0 - m_1) = \frac{e}{m'}\frac{H}{4\pi c}(m_0 - m_1), \text{ since } \omega = \frac{e}{2m'c} H$$

$$\therefore \Delta\nu = m_0 - m_1 \text{ Lorentz units.}$$

The energy which can be given to, or abstracted from the electron by the electric field produced by the increasing magnetic field depends thus upon the inclination of the orbit to the lines of magnetic force. It will be a maximum or minimum according as the plane of the orbit is perpendicular or parallel to the field. In the former case the orbit is cut by the maximum number of lines of force, and $m=l$, in the latter no lines cut the orbit and $m=0$. If the orbits were oriented in random directions we should have all values of m between 0 and l and the energies added (or subtracted) of all intermediate values between the maximum and minimum mentioned above. The magnetic field would then simply add a wing on each side of the line, of a width defined by the values of the maximum energy contributed by the electric field. To account for the sharply defined Zeeman components it is necessary to assume that the orbits orient themselves at definite inclinations to the field. This is known as space-quantization, a brilliant conception due to Sommerfeld.

Magnetic Quantum Number "m" and Space-Quantization. — To account for the Zeeman effect we must assume that only certain values of m, the projection of the vector l of Fig. 387 on the line MH parallel to the magnetic field are possible. These values are determined by the orientation of the orbit, and are integers.

The orientation of atoms, which have a magnetic moment, in a magnetic field has been shown experimentally by the method originated by Stern and Gerlach, who projected a beam of atoms in a high vacuum through a non-homogeneous magnetic field, and showed that in certain cases deflections in two opposed directions resulted, proving that the magnetic axes of the atoms oriented themselves parallel or antiparallel to the magnetic field. The number of orientations possible depends upon the value of l.

If $l=3$ as in the F spectral term seven possible orientations are possible, giving m the values 3, 2, 1, 0, -1, -2, -3.

If $l=2$ as in the D term, and $l=1$ as in the P term, five and three orientations may occur respectively. If the radiation results in a transition from the upper F level to the lower D level, l decreases by 1, corresponding to an angular momentum decrease of $h/2\pi$.

For the S term $l=0$: This would correspond to an elliptical orbit degenerated to a line, in other words a to-and-fro oscillation requiring the passage of the electron through the nucleus. We cannot represent this condition by our model, but the difficulty disappears in the new wave-mechanics.

The orbits, with their appropriate values of l and m, the possible orientations in each case and the transitions possible are shown in Fig. 388*a*.

The selection rule governing the transitions is that m can change only by 0 or ± 1, as shown by arrows in Fig. 388*b*. Transitions corresponding to $\Delta m=0$ (dotted lines) give light of unaltered

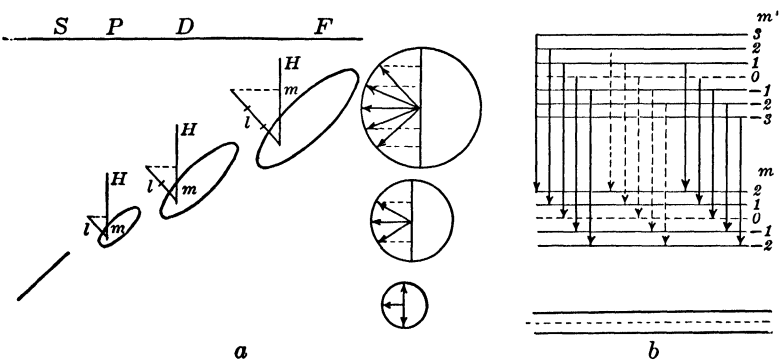


FIG. 388

wave-length linearly polarized parallel to the field, while transitions of ± 1 give light of altered wave-length linearly polarized viewed in a direction perpendicular to the field, or circularly polarized in the field direction. The allowable transitions are shown for $(l=3\rightarrow l=2)$. For the lowest level $l=1$. That a number of different transitions give a single Zeeman component results from the circumstance that the distance between the magnetic energy levels is the same for all values of l. We shall presently discuss examples in which this is not the case.

The Anomalous Zeeman Effect. — The normal Zeeman effect is exhibited by the hydrogen lines, and the singlet lines of other spectra. In general the spectrum of an element consists of series of lines which may be singlets, doublets, triplets or multiplets. A spectrum may consist of single lines only, as is the case with hydrogen, though under very high spectroscopic resolution the hy

drogen lines show close fine-structure. Or it may consist of a series of doublets, as in the case of the alkali metals, or of singlets and triplets as in the case of cadmium and mercury, or of multiplets.

An anomalous Zeeman effect is shown in the case of the components of all doublets, triplets and multiplets. In very strong fields, however, the complex resolution of say the yellow doublet of sodium (D_1 giving four lines and D_2 six in weak fields), is transformed into a normal Zeeman triplet. This is known as the Paschen-Back effect and will be considered presently. Lithium is a better example, as it is impossible to produce a field of sufficient strength to show the complete effect with sodium.

To account for doublets or triplets on Bohr's theory the energy levels between which transitions occur (giving rise to radiation) must be subdivided. To accomplish this Sommerfeld introduced a third quantum number, j, called the "Internal quantum number," which if empirically allotted to certain terms gave results in agreement with experiment, if the changes in j in possible transitions followed a simple selection principle. The present view is to associate the doubling, etc., of the lines with the spin of the electron, the theory of which will be considered presently.

Landé's Formula. — The first real step towards explaining the anomalous Zeeman effect was made by Landé, who derived half empirically a formula known by his name, from which the magnetic splitting of any of the spectral terms could be predicted. It is based on the three quantum vectors, j the total momentum vector, l the azimuthal quantum vector and s the moment of the electron spin.

The formula is $$g=1+\frac{j(j+1)+s(s+1)-l(l+1)}{2j(j+1)}.$$

Landé called "g" the "Aufspaltungsfaktor," or factor of splitting, and it may be defined as the fraction by which we must multiply the normal distance of separation (corresponding to one Lorentz unit) of the magnetic levels of any term to obtain the levels involved in the anomalous effect.

Before taking up the subject of the spinning electron we must consider the subject of the ratio of the mechanical to the magnetic moment.

Magnetic Moment of Atom: Bohr's Magneton. — The mechanical moment of the atom, $L=l(h/2\pi)$, in which l, taking integral values gives the value of L in quantum units $h/2\pi$ corresponding to orbits of increasing diameter and energy.

We now require the corresponding magnetic moment. If w' is the angular velocity of the electron, its frequency is $w'/2\pi$ and the equivalent current $ew'/2\pi$ (e=electron's charge).

The magnetic moment of the circular current equivalent to the electron rotating in an orbit of radius r is

$$\mu = \frac{ew'}{2\pi}\pi r^2 = \frac{e}{2}w'r^2.$$

The mechanical moment is

$$mw'r^2 = l\frac{h}{2\pi} (m = \text{electron's mass})$$

or

$$\mu = \frac{e}{m}\frac{lh}{4\pi}.$$

The elementary quantum of magnetic moment corresponding to the elementary mechanical moment $h/2\pi$ is thus

$$\mu = \frac{e}{m}\frac{h}{4\pi} \quad \text{or} \quad \frac{eh}{4\pi mc} \text{ if expressed in electromagnetic units.}$$

This quantity is known as the Bohr magneton and may be defined as the magnetic moment of an atom with a mechanical moment $h/2\pi$.

The ratio of the magnetic to the mechanical moment is $-e/2mc$ the negative sign resulting from the electron charge $-e$, indicating that the north pole of the atomic magnet is opposed to the vector of the mechanical moment. This ratio of the two momenta multiplied by H gives the angular velocity of the Larmor precession which as we have seen is

$$w = \frac{e}{2mc}H$$

and this in turn measures the separation of the Zeeman components in the normal effect. In the case of the anomalous effect the ratio has values different from the one given above, as we shall see, the proportionality factor by which we must multiply $e/2mc$ being termed "g." As we shall see g has different values for the different spectral terms and may be defined as representing, when multiplied by $e/2mc$, the ratio of the magnetic to the mechanical moment of the atom in a particular state of excitation. For the normal Zeeman effect $g = 1$.

The Spinning Electron. — Inasmuch as the abnormal Zeeman effect is exhibited by the components of doublets, triplets and multiplets, we must first consider the type of atom concerned in the emission of these radiations and the physical processes involved. This is a necessary introduction to the study of the action of the magnetic field on the radiator.

One of the first attempts towards a solution of the problem has

already been outlined, namely, the hypothesis that the core of the atom had a magnetic moment of its own. This was found to be unsatisfactory, and in 1925 Goudsmit and Uhlenbeck introduced the idea of the spinning electron, endowing it with *rotation* in addition to mass and charge. Each electron is regarded as rotating about its own axis, the rotation giving rise to a magnetic field directed along the axis, just as the revolution of the electron in its orbit generates a field perpendicular to the orbital plane. The electron can thus be regarded as a small magnet which precesses in the magnetic field generated by its orbital motion, and we have as a result energy changes analogous to those discussed in the normal Zeeman effect. On this hypothesis the various types of spectra and the abnormal Zeeman effect have been very completely accounted for. On the assumption that the moment of momentum of the spinning electron $s=\frac{1}{2}$, *i.e.* one-half of the quantum unit which we employed in the case of the orbital motion, while its magnetic moment is that of one Bohr magneton. The ratio of the magnetic to the mechanical moment is thus twice as great in the case of the spinning electron as in the case of the orbital motion. Our model of the atom now takes the form indicated in Fig. 389. To the vector l (moment of momentum or orbital motion), is to be added the vector s, the momentum of spin, the resultant j, the total moment of momentum of the atom, being a quantum vector which can take only definite values depending upon the quantum values of l and s and their orientation. In the case of an atom with a single orbital electron l and s are either parallel or antiparallel as we shall see, and the parallelogram collapses to a line. The precession of the spinning electron causes the vectors l and s, which are rigidly coupled, to rotate about the resultant j, as indicated by the arrow, *i.e.* the parallelogram rotates on j as an axis, and we have an increase or decrease of the energy, according to the direction of the precession as in the case of the normal Zeeman effect, the magnetic field due to the orbital motion representing the applied field in the case of the Zeeman effect, and the electron representing the atom. The direction of precession will reverse when the vector s in Fig. 389 is turned down below the horizontal position. The rotation of the vector l about j as an axis means simply that the orbit "precesses" just as it does when acted upon by an external field. In producing the Zeeman effect we employ a heavy magnet, the field direction remaining fixed. In the present case we are dealing with a very light magnet



FIG. 389

(the orbit of the electron) and it "precesses" as well as the spinning electron. Like the vector l of our old model, j is measured in units $h/2\pi$, and can take integral or half integral values only, including 0. We cannot give a physical interpretation of the zero value of the orbital moment of momentum except in the case of an atom having more than one electron and consider them as rotating in opposite directions. Spectroscopic data, however, and the new wave-mechanics require that we must admit also the zero value for l even when only the rotation of a single electron is involved in the radiation. We will begin with a specific case, that of the sodium atom which consists of a core, the structure of which is equivalent to that of an atom of neon, with a single electron in rotation about it.

Assigning to l the values 0, 1, 2, and 3 and corresponding to spectral terms S, P, D and F (see page 686) we can construct the values of j as shown in the following table:

Spectral Term	S	P	D	F
$l=$	0	1	2	3
$j=l\pm s=$	$\frac{1}{2}$	$\frac{3}{2}$, $\frac{1}{2}$	$\frac{5}{2}$, $\frac{3}{2}$	$\frac{7}{2}$, $\frac{5}{2}$

s has the value $\frac{1}{2}$ in all cases, consequently the vectors l and s must be either parallel or antiparallel, otherwise the quantum conditions (j values differing by integral numbers) would be violated. This will be made clearer presently. This table shows us that all of the energy levels of an atom with a single valence electron are double except the S levels which are single. Transitions from the double P levels to the lowest S level give the principal series of doublets, of which the D lines form the first pair. With a single orbital electron, and a single value of $s=\frac{1}{2}$, for a value of $l=3$ we can have only two values of j viz., $3\frac{1}{2}$ and $2\frac{1}{2}$ as shown by Fig. 390 which illustrates the formation of the F term in the table. The transition involved in the formation of the first members of the series of doublets of sodium are shown in the diagram in the Chapter on Origin of Spectra.

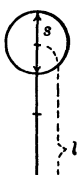


Fig. 390

In the case of the diffuse series the levels of both spectral terms (P and D) are double. In reality the lines are threefold, as one of the four transitions is forbidden. In the case of atoms with two or more orbital electrons, we can generally sum the l and s vectors of the several electrons and form the resultant j as before. We can now have the rotating parallelogram of Fig. 389, as the total

or resultant s may now have a value greater than unity. The relation between the N number of orbital electrons, the resultant s and the type of spectral series is shown in the following table:

N	Res. s	Type of Series
1	$\frac{1}{2}$	Doublets
2	0, 1	Singlets and triplets
3	$\frac{1}{2}$, $\frac{3}{2}$	Doublets, quartets
4	0, 1, 2	Singlets, triplets, quintets
5	$\frac{1}{2}$, $\frac{3}{2}$, $\frac{5}{2}$	Doublets, quartets, sextets

The compounding of the resultant l vectors and s vectors for the case $l=3$, $s=2$ is shown in Fig. 391a the resulting values of j being 5, 4, 3, 2, 1. The value 5 being for the parallel and the value 1 for antiparallels. For $l=3$ and $s=3/2$ the resultant j's are shown in Fig. 391b for which $j=4\frac{1}{2}$ (parallel) $3\frac{1}{2}$, $2\frac{1}{2}$, $1\frac{1}{2}$ (antiparallel). With an even number of orbital electrons we have integral values of j — with an odd number, half integral. This means that with given values of l and s, the axes of the spinning electrons orient themselves so as to give resultant j values differing by integers.

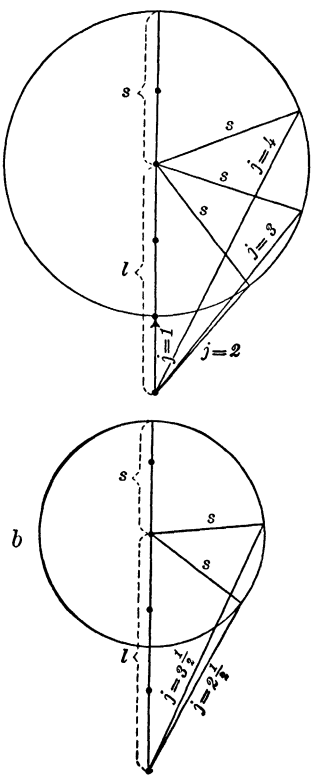


Fig. 391

Magnetic Levels in the Anomalous Zeeman Effect. — We will now consider what happens when we apply an external magnetic field to an atom of this type.

We now have the Larmor precession (about the direction of the magnetic field) of the resultant vector j of the last section, the angular velocity, however, being in general much less than that of the inner precession which gives rise to the doublets, triplets, etc. The direction of j is "space quantized" in the magnetic field in the same manner as was the plane of the orbit in the case first considered and the energy given to, or abstracted from, the system is measured by the projection "m" of j on the direction of H, the magnetic field.

As we have seen in the section treating of the ratio of the magnetic to the mechanical momentum vectors (page 688) it is this quantity which defines the angular velocity of the Larmor precession which in turn defines the energy changes and separation of the Zeeman energy levels.

We require then the ratio of the total resultant magnetic moment M (expressed in Bohr magnetons) to the total resultant mechanical moment j (expressed in multiples of $h/2\pi$). This ratio is $g=M/j$. In Fig. 392 l the mechanical moment of the orbital motion and s the mechanical moment of the spinning electron give the resultant j. We require now the magnetic moment which will be the sum of the components along the direction of j of the magnetic momenta of the orbital motion and spinning electron. The components perpendicular to j are without effect since their direction changes rapidly owing to the inner precession of l and s about j.

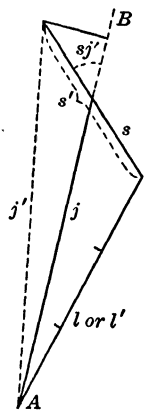


FIG. 392

The magnetic moment of the orbital motion is l Bohr magnetons, consequently l of Fig. 392 represents also the magnetic moment l'. The ratio of the magnetic moment of the spinning electron to its mechanical is, however, double the ratio for the orbital motion as we have seen, consequently the mechanical vector s must be doubled as shown in Fig. 392 to give the magnetic vector s'. The resultant magnetic moment is j' and its projection AB on the direction j is the sum of the components along j which we require.

$$AB=j+s\cos sj.$$

The ratio
$$\frac{M}{j}=\frac{j+s\cos sj}{j}.$$

Now
$$l^2=s^2+j^2-2js\cos sj$$

$$2js\cos sj=s^2+j^2-l^2$$

$$s\cos sj=\frac{s^2+j^2-l^2}{2j}$$

$$\therefore\frac{M}{j}=\frac{j+\frac{s^2+j^2+l^2}{2j}}{j}=1+\frac{s^2+j^2-l^2}{2j^2}.$$

This quantity which has been termed g determines the separation of the magnetic levels in the spectral terms.

It has been found however that a quantum refinement of this formula must be applied, writing for s^2, j^2 and l^2

$$s(s+1),\ j(j+1) \text{ and } l(l+1)$$

so that we have finally

$$g=1+\frac{s(s+1)+j(j+1)-l(l+1)}{2j(j+1)}.$$

This is identical with Landé's formula which was deduced empirically.

We may regard the coupling of the vectors l and s as analogous to the space-quantization of the electron orbit in a magnetic field. In the present case the electron, itself a spinning magnet, by its orbital rotation may be regarded as forming another magnet of almost infinitesimal mass. The axis of the orbit and the spin axis may therefore adjust themselves at quantized angles, which remain fixed, except when a transition from one energy level to another occurs. Precession occurs between the spinning electron and the field which it creates by its rotation, but the field precesses also, in other words, both fields precess about j as an axis, the direction of j in space remaining fixed. The only difference between this case and that of a Larmor precession in the field of a material magnet where the field direction remains fixed owing to the great mass of the magnet, is that in the present instance both magnets have the same mass consequently both precess about a common axis j. Since the frequency of the precession of l and s about j is enormously smaller than the orbital frequency, this means that the electron travels along a path having a wave form.

We resolve the magnetic momenta of the orbital motion and electron spin into components parallel and perpendicular to j. The integrated effect of the latter is zero, since both point towards, and rotate rapidly around, j.

The total magnetic moment of the atom is thus seen to be the sum of the projections of l and $2s$ on the j direction.

This gives us the g formula.

The quantum refinement above referred to is justified by the modern wave-mechanics, which substitutes a wave-system for the orbital electron, and the necessity for applying it means simply that the mechanical model which we have employed (spinning electron in orbital motion) cannot be made to give an exact solution of the problem *without* this refinement.

We will now apply this formula to a specific case, the anomalous effect shown by the D lines of sodium. As we have seen these lines result from transitions from the double P level ($^2P_{\frac{3}{2}}$, $^2P_{\frac{1}{2}}$) to the

single S level, as shown at left of Fig. 393. For the upper level $l=1$ and $j=l\pm s$ or $\frac{3}{2}$ and $\frac{1}{2}$, since $s=\frac{1}{2}$.

In the magnetic field the S and P levels are split up into magnetic levels, the separation of which in terms of the normal Zeeman distance can be found by the g formula.

For the $S_{\frac{1}{2}}$ level we have $l=0$, $s=\frac{1}{2}$ and $j=\frac{1}{2}$, substitution of which in the g formula gives us $g=2.$, *i.e.* the magnetic separation of the S levels is two Lorentz units.

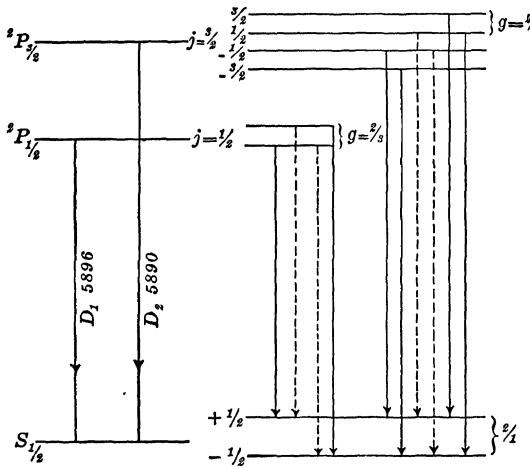


FIG. 393

For the $^2P_{\frac{1}{2}}$ level we have $l=1$, $s=\frac{1}{2}$, $j=\frac{1}{2}$ the g formula giving $g=\frac{2}{3}$ or a separation of $\frac{2}{3}$ that of the normal Zeeman separation (Lorentz unit) while for the $^2P_{\frac{3}{2}}$ level $j=\frac{3}{2}$, $s=\frac{1}{2}$ and $l=1$, the g formula giving $\frac{4}{3}$ as the separation between any pair of the four magnetic levels into which the $^2P_{\frac{3}{2}}$ level splits. The transitions are given in Fig. 393. (Distance between levels incorrectly drawn.)

Solid and dotted lines representing components polarized perpendicular and parallel to the magnetic field respectively. Only transitions for which m changes by 0 or ± 1 are allowable.

It is obvious that the lengths of the lines for the six possible transitions in the case of D_2 are all different and we have therefore a separation of the line into six components. It will be remembered that in the diagram for the normal effect, though we had a number of magnetic levels as in the present case, there were numerous coincidences in the length of the lines representing allowable transitions, owing to the equality in the separation of the components of the upper and lower level.

The Paschen-Back Effect. — It was observed by Paschen and Back in 1912 that as the strength of the magnetic field was increased in the case of doublets and triplets, the components of one line, instead of encroaching upon the territory occupied by the components of the other line, acted as though repulsive forces existed between the components of the two lines. Passage of one

line through another never occurred, or only very imperfectly, as unsymmetrical diffusion or widening of the components occurred before they reached the points occupied by their neighbors. With further increase of field the complicated group is transformed finally into a simple Zeeman triplet, symmetrical about the centre of gravity of the original doublet or triplet. The field strength required for this transition depends upon the separation of the components of the doublet or triplet. It was first studied in the case of the lithium doublet, the components of which are much closer than the D lines of sodium and also in the case of the close oxygen triplet 3947.44, 3947.62 and 3947.73.

Similar effects were also observed by Wood and Kimura [1] in the case of the close multiplets in the iodine spectrum which they discovered with an echelon in 1915. They found that many lines, previously regarded as single, consisted of groups of from three to six lines, giving in weak fields very complicated patterns, which, as the field strength was increased, fused to normal triplets. The interpretation is, however, somewhat different in this case.

Theory of the Paschen-Back Effect. — In the treatment of the anomalous Zeeman effect the components of the magnetic moment of the orbital motion and the electron spin perpendicular to the direction of j were considered as neutralizing when averaged for a complete period, in so far as any external action is concerned. This is true only when the frequency of precession of l and s around j is large in comparison to the Larmor precession.

Larmor Precession of the Orbit in Strong Fields. — In weak fields the coupling between l and s is stronger than the coupling between the external field and the magnetic vectors of the orbit and the electron spin, and the vector j precesses in the external field, l and s rotating about it at the same time at a fixed angle: the projection m of the vector j upon the field direction remains constant during this rotation.

As the field increases in strength this is no longer the case, for it begins to act upon the magnetic vectors of orbit and spin separately, and the parallelogram of Fig. 389 becomes distorted, finally ceasing to exist altogether, the two vectors l and s precessing in the field quite independently of each other.

The spinning electron now orients itself so that its magnetic axis is parallel or antiparallel to the field and the precession of its axis takes place independently of, and with twice the frequency of the precession of the orbit, l and s no longer uniting to a resultant j, the projection of which on H, the field direction, thus becoming equal to its actual value. This means that the Larmor precession

[1] *Astrophysical Journal*, xlvi, 181, 1917.

of the orbit will take place in the magnetic field with space-quantization, as in the case treated for the normal Zeeman effect.

The increment of magnetic energy now consists of the sum of two parts, that of the orbital motion and that of the spinning electron.

$$E_m = (m_1 + 2m_s)\omega \frac{h}{2\pi}.$$

This very brief statement of the theory of the Paschen-Back effect is intended merely to give the reader a general idea of the line of attack which has been made upon this extremely complicated phenomenon.

The complete treatment, covering the transition stage between weak and very strong fields, in which the complex pattern is gradually, by the fusion or disappearance of lines, ironing itself out into the simple Lorentz triplet, is beyond the scope of this book.

The Faraday Effect: Magnetic Rotation of the Plane of Polarization. — The discovery was made by Faraday that a transparent isotropic medium, when placed in a powerful magnetic field, acquires the property of rotating the plane of polarization, when the light traverses the medium in the direction of the lines of magnetic force. The phenomenon differs, however, from natural rotation, in that the direction in which the plane of the vibration turns depends upon whether the light rays are passing through the medium from the north pole of the magnet towards the south, or in the reverse direction. The rotation is therefore doubled if the light is reflected back through the medium, instead of being annihilated as in the case of quartz and other active substances. The effect is most pronounced with media having a high refractive index, such as bisulphide of carbon or dense flint glass. With a powerful Ruhmkorff magnet, the poles and cores of which are bored out to allow of the passage of light rays along the lines of force, the rotation can be easily observed with a thick piece of ordinary plate glass. Sun or arc light is passed through a Nicol prism, the hollow magnet cores, and the glass block between the poles. A second Nicol is placed in such a position as to extinguish the emergent beam. On throwing the current into the magnet, the field immediately becomes brilliantly illuminated, and by turning the analyzing Nicol until darkness is again produced the amount and direction of the rotation can be determined.

Owing to the rotatory dispersion this position will vary with the color, and the field will appear blue, purple and red in succession, as will be readily understood from Fig. 394, in which the

dotted arrow represents the original direction of the vibration, and the arrows the rotated red, green and blue vibrations. The analyzing Nicol in its original position is indicated by N. The rotation in this case is clockwise, and all of the colors are transmitted with more or less freedom, consequently the field appears nearly white. On turning the Nicol in the same direction it will extinguish the red first, leaving an outstanding color of a bluish green; the green goes next, leaving a purple field made up of the transmitted red and blue, and finally the blue disappears, leaving the field orange-red. If we examine the light through a spectroscope, as we turn the Nicol, we shall see a dark band enter the spectrum on the red side and leave it on the violet side.

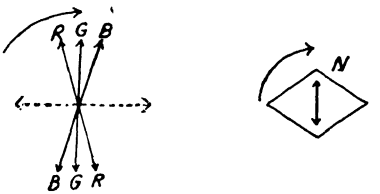


FIG. 394

Explanation of the Magnetic Rotation. — The explanation of the natural rotation in active substances which Fresnel gave, was that the plane vibration was decomposed into two oppositely polarized circular vibrations, which were propagated with different velocities. The same explanation will do for the magnetic rotation, provided that it can be shown that the refractive index of a medium in a magnetic field for circularly polarized light depends upon the direction of revolution. The matter was attacked experimentally by Righi and Becquerel independently, and both investigators found that the interference fringes, formed by two streams of circularly polarized light, one of which had traversed a block of glass placed between the poles of a magnet, were displaced when the magnetic field was formed. The direction of the displacement depended on whether right- or left-handed circular light was used, which showed that the effect of the field was to increase the refractive index for one type of vibration and diminish it for the other. It remained only to show that the actual decomposition of the plane vibrations into circular ones actually occurred by some experiment analogous to the one which Fresnel made with his battery of quartz prisms built of right- and left-handed crystals in alternation.

Resolution into Circular Components. — The experimental resolution of the light into its two circular components in the Faraday effect is a much more difficult problem than the one which confronted Fresnel, since we do not have at our disposal two liquids of the same index of refraction and of opposite magnetic rotation, with which hollow prisms might be filled, in the manner adopted

by Fleischl in the case of natural rotation. The problem has, however, been attacked and solved in a very beautiful manner by Brace.[1]

It will be remembered that in Fresnel's arrangement of right- and left-handed quartz prisms, the clockwise circular component which travelled at the higher velocity in one prism, travelled at the slower velocity in the following prism. The very ingenious idea occurred to Brace to reverse the direction of revolution of the circular vibrations at the boundary surface between the two prisms, which can be done with a half-wave plate of mica. By this artifice the same thing is accomplished as by employing prisms of dextro- and laevo-rotatory media in succession. A double prism of extra dense flint glass was employed with a half-wave plate cemented between the two components. The experimental difficulties were, however, found to be too great and no conclusive results were obtained. It subsequently occurred to Brace to make use of reflection instead of refraction, and look for evidences of a division of the ray into two circularly polarized rays. This at first sight seems to antagonize the law of reflection, but it must be remembered that the law of equality between the angles of incidence and reflection is based upon the fact that the velocity is the same before and after reflection. If we apply the Huygens construction for reflection to a case in which the velocity is less after reflection than before, we shall find that the angle of reflection *is less than* the angle of incidence. Suppose now that our two circular components in the magnetized medium encounter a reflecting surface which reverses the direction of revolution of each. The fast component now becomes the slow, and *vice versa*, and we have a division of the ray.

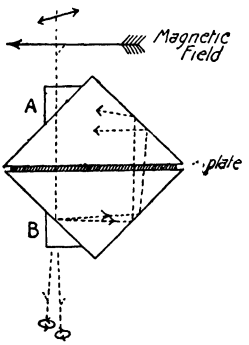


FIG. 395

The experiment in its final form made use of multiple reflections. A rectangular block of glass was made by cementing two right-angled prisms ($n=1.903$) together, with a half-wave plate of mica between them to reverse the direction of the circular vibrations (Fig. 395). The incident light enters the prism normally through a small auxiliary prism A, traverses the $\lambda/2$ plate, which turns its plane of polarization through 90°. The light is travelling perpendicular to the lines of force, consequently the magnetic field does not affect it. As soon, however, as it suffers total reflection, it travels along the lines of force, and is consequently broken up into circular

[1] *Wied. Ann.*, xxvi, 576, 1885; *Phil. Mag.* (6), i, 464, 1901.

components, one of which travels faster, the other slower, than the original disturbance before reflection. A division therefore results, and we have two reflected rays. The light travelled around the prism five times, undergoing twenty internal reflections, and emerged through a second auxiliary prism B at the top of the block. The source was a vertical slit powerfully illuminated with an oxy-hydrogen flame fed with sodium. On turning on the magnetic field the image in the telescope was seen distinctly doubled, and on examination with a Nicol prism the two lines were found to be nearly plane-polarized, due to polarization complications resulting from the total reflection and the $\lambda/2$ plate, which are fully explained in the original paper.

The velocity of right- and left-handed circularly polarized light in a magnetized medium was measured by Mills [1] by means of a Michelson interferometer. By means of a Brevais double plate, one-half of the field was illuminated with right- the other half with left-handed circularly polarized light, the fringes crossing the field in a direction perpendicular to the dividing line between the two halves of the plate. On exciting the magnet the fringes on one side moved up, while those on the other side moved down. The accelerated ray was found to be the one in which the direction of the circular vibration was the same as that of the current in the magnet.

Relation between the Field Strength and Rotation. — The angular rotation increases in general in proportion to the strength of the field, but this rule is not strictly followed, the most marked exceptions being shown by iron, nickel and cobalt. In the case of iron the relation between the field strength H and the rotation δ is shown in the following table. If the rotation were proportional to H, the figures of the last column would be approximately the same:

H	δ	$\frac{10^5\delta}{H}$
4420	1.72°	39
8060	3.47°	43
14100	4.41°	31
18500	4.45°	24
30100	4.36°	14

If instead of H we take the *magnetization*, we shall find that the rotation is proportional to this quantity. In the case of iron the magnetization increases with the field strength, but not at the same rate, finally becoming "saturated," beyond which point a further increase of field produces no increase in the magnetization. The rotation also attains a maximum value at the same point.

Direction of the Magnetic Rotation. — Shortly after the discovery of the magnetic rotation of the plane of polarization by

[1] *Phys. Rev.*, February, 1904.

Faraday, E. Becquerel observed that the addition of salts of iron to water diminished the rotatory power which it acquired in a magnetic field, an experiment which Verdet interpreted as proving that the iron salts possessed a rotatory power of opposite sense from that exhibited by the diamagnetic substances investigated by Faraday. The latter he called "positive" rotation (*i.e.* in the same direction as that of the current producing the field), as distinguished from the "negative" rotation shown by paramagnetic substances, but it was presently found that titanium tetrachloride, though diamagnetic, exhibited a "negative" rotation, a circumstance which caused Verdet to come to the conclusion that the magnetic properties of the material had nothing to do with its rotatory power. The negative rotation is now called "the paramagnetic rotation," and it will be more fully discussed towards the end of the chapter.

H. Becquerel showed however that there was a relation between the two properties, the rotatory power depending on the refractive index and another function which varied with the specific magnetism of the material. He attributed the negative rotation to the production of an internal magnetic field of opposite direction to that of the external field. At this period (1876) little or nothing was known about the origin of magnetic phenomena. In 1884 Kundt discovered that very thin films of iron gave enormous rotations when placed in a strong magnetic field, a film the thickness of which was only one-third of the wave-length of light giving a rotation of the polarization plane of over 4°. An iron plate of 1 mm. thickness, if it were possible to get any light through it would give, under similar conditions a rotation of 66 complete revolutions while a quartz plate of equal thickness gives a rotation of only 20°. The direction of rotation was found to be positive, *i.e.* metallic iron behaves like the diamagnetic substances.

Time Required for the Faraday Effect. — Bichat and Blondlot found that if a Leyden jar was discharged through a helix of insulated wire surrounding a tube filled with carbon bisulphide, the plane of polarization was found to turn with each impulsive rush of the oscillatory discharge, being waved to-and-fro at the rate of some 70,000 times per second.

To find out whether any time was required for the development of the effect, they illuminated the upper portion of a slit with the light of the spark and the lower portion with the light restored by the CS_2 tube, and examined it in a revolving mirror. The illuminated slit was seen spread out into a serrated band, but no discontinuity was found between the two illuminated halves, showing that the effect is practically instantaneous.

Abraham and Lemoine [1] making use of the method which they employed in the study of the Kerr effect (described in the Chapter on Electro-Optics) in which damped oscillations of very high frequency from a condenser are used both for the formation of the source of light (spark) and for the excitation of the magnetizing solenoid, found that the time which elapsed between the creation of the field and the rotary effect was less than 10^{-8} sec.

More recent investigations by Beams and Allison [2] have shown a slight time retardation which varies with the nature of the fluid used (in comparison with CS_2) varying from 1 to 10×10^{-9} sec., an interval of the order of time required for one complete revolution of the Larmor precession. Their results indicate that at least one Larmor revolution must occur before the magnetic rotation is observable.

The Kerr Magneto-Optic Effect. — The important discovery was made by Kerr [3] that plane-polarized light becomes elliptically polarized when reflected from the polished pole of an electromagnet. The incident light must be polarized either in, or perpendicular to, the plane of incidence, otherwise elliptical polarization results from the metallic reflection. On setting a Nicol prism in such a position as to completely extinguish the reflected light, and exciting the magnet, the light instantly reappeared, and could not be extinguished by further rotation of the Nicol, except by the introduction of a quarter-wave plate. The ellipticity is not very great, and we can regard the effect as a rotation of the plane of polarization. Employing normal incidence he found however, that in this case also elliptical polarization was produced and complete extinction by the analyzing Nicol could be secured only by the introduction of a quarter-wave plate. This means that the magnetic field introduces a component perpendicular to the original electric vector of the incident light, the so-called Kerr component.

Intimately connected with the Kerr effect is the rotation of the plane of polarization by thin films of iron in a magnetic field, which has been alluded to previously, a small trace of elliptical polarization being present in this case also.

This rotation by the iron films is to be referred to the unequal velocity of the two circular components into which the incident plane vibration is resolved. One of these is more powerfully absorbed than the other (compare with the unequal absorption by

[1] Abraham and Lemoine, *Comptes Rendus, 130,* 499, 1900.
[2] *Phys. Rev., 29,* 161; *30,* 66, 1927.
[3] *Phil. Mag* , May, 1877 and March, 1888.

tourmaline of the plane-polarized components) and this gives rise to the ellipticity. The theoretical treatment is very involved, and there is no very satisfactory interpretation of the phenomenon based on electron theory up to the present time. For a very complete treatment see Müller-Pouillet's *Lehrbuch der Physik*, Vol. II, Part 2 (1929).

It is worthy of comment that the sign of the rotation is negative in the case of the Kerr effect at normal incidence, and positive in the case of the transmission of light through an iron film. This shows that the Kerr phenomenon is not a simple rotation due to a slight penetration of the light into the steel in the act of reflection.

Theories of Magnetic Rotation. — Drude, in his *Lehrbuch der Optik* developed two theories of magnetic rotation, the first (his own) based on the hypothesis of molecular currents as conceived by Ampere and Weber to explain magnetism and diamagnetism. The creation of the magnetic field was supposed to set up a current (electron whirl) within the molecule, and the periodic electric forces of the light-waves caused periodic displacements of the centre of rotation, or in other words, waved the magnetic field of the molecular current to and fro. The introduction of expressions for this motion into the equation representing light propagation accounted for magnetic rotation of the plane of polarization, but called for a rotation of opposite direction on opposite sides of an absorption band; in other words an anomalous rotatory dispersion, represented by a curve similar in form to the curve of anomalous dispersion shown by absorbing media. Very few cases of such a rotation have been found, and these have been explained on a different hypothesis. The hypothesis of molecular currents was on this account abandoned. The second theory, due to Voigt was based on the Hall effect.

An electric current or a moving electron is subjected in a magnetic field to a deflecting force which is at right angles to the direction of the current and the lines of force. In a magnetic field, then, an electron which is thrown into vibration by light-waves will experience a force which will be proportional to the velocity with which it is moving, and by introducing an expression representing this force into the equations for wave-propagation a formula was derived which called for rotations of very large value *and* similar sign on opposite edges of an absorption band.

We will now consider some experimental observations in agreement with the predictions of the second theory.

Magnetic Rotation of Absorbing Gases and Vapors. — The experiments of Macaluso and Corbino showed that the vapor of

sodium in a magnetic field produced large rotations of similar sign for wave-lengths immediately adjacent to the *D* lines.

A small sodium flame, placed between the poles of an electromagnet, was traversed by a beam of polarized white light, in the direction of the lines of magnetic force. A Nicol prism was so oriented as to extinguish completely the light when the current was not traversing the coils. On forming the magnetic field a brilliant yellow light was found to be transmitted by the Nicol, which the spectroscope showed to consist of narrow bands symmetrically placed on each side of the *D* lines. By turning the Nicol first in one direction and then in the other it was easy to see that the rotation was of the same sign on opposite sides of the band.

With a view of testing the magneto-rotation dispersion formula, developed from theoretical considerations by Voigt (to be discussed presently) over a wider range of wave-lengths than was possible with a sodium flame, the experiment of Macaluso and Corbino was carried out on a larger scale by the author with the vapor of metallic sodium in a steel tube.[1] As the phenomena exhibited by the vapor are extremely beautiful, and very easily shown, the apparatus in its final form will be described in detail.

A piece of thin, seamless steel tubing (bicycle tubing) of such diameter as to permit of its being slipped easily through the hollow

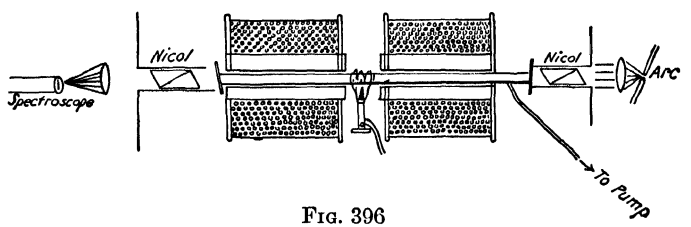


FIG. 396

cores of the electromagnet is procured. A short piece of small brass tubing is brazed into one end, through which the tube is exhausted. The ends are closed with small pieces of plate glass cemented with sealing wax. The arrangement of the apparatus is shown in Fig. 396. A piece of sodium the size of a walnut is rolled out into a cylinder between two boards and inserted into the tube just before the second end-plate is cemented on. The tube is at once placed in position in the magnet and exhausted. If a piston pump is used for the exhaustion, a glass topcock should be put between the pump and the tube to prevent back leakage of air. Care must be taken to have the lump of sodium midway between the poles of the magnet. The steel tube is now heated by means of

[1] *Phil. Mag.*, Oct., 1905; July, 1907.

a Bunsen burner, and the pump worked to remove the hydrogen liberated from the sodium, after which the burner is removed and the tube allowed to cool.

Light from a heliostat, or an arc lamp, is now passed in succession through a Nicol prism, the steel tube, a second Nicol, and then concentrated on the slit of a spectroscope. If the instrument has a large dispersion (a 14-foot concave grating was used in the present case) all of the phenomena now to be described can be seen.

The Nicols are set for complete extinction and a small flame placed beneath the tube. As soon as the vapor begins to form, two

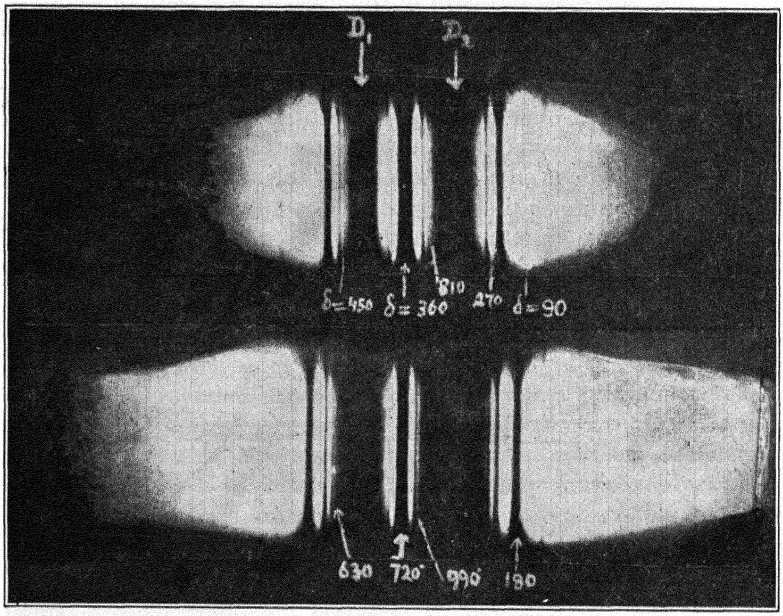


FIG. 397

very bright lines will appear in the position of the D lines the moment the magnet is excited. These lines represent the constituents of the white light, which are rotated through 90° by the vapor and passed by the analyzing Nicol. The lines are in reality double, though their duplicity cannot be made out when they first appear. As the density of the vapor increases the components separate, four lines being distinctly visible. The lines continue to separate, and presently a second pair appears between them for which the rotation is 270°, the dark regions between representing rotations of 180°.

In the former the two inner 90° lines are beginning to fuse to-

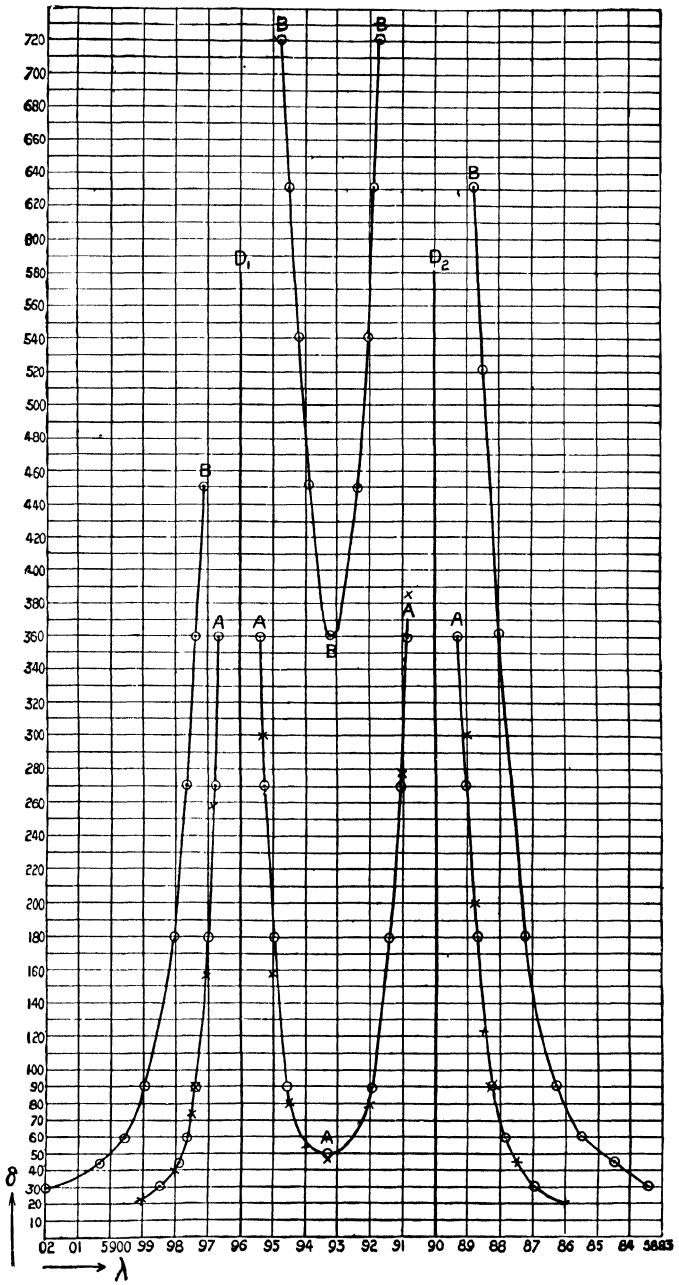

B
B
B
B
D1
D2
A
A
A
A
A
x
δ
λ
720
700
680
660
640
620
600
580
560
540
520
500
480
460
440
420
400
380
360
340
320
300
280
260
240
220
200
180
160
140
120
100
90
80
70
60
50
40
30
20
10
02 01 5900 99 98 97 96 95 94 93 92 91 90 89 88 87 86 85 84 5883

FIG. 398

gether, the centre being partially dark, however; in the latter the fusion is complete and the centre of the system of lines is bright. With a further increase in the vapor density the outer lines (90°) separate still further, and widen out into broad flares of light, other lines appearing between them corresponding to larger rotations, the system resembling a set of diffraction fringes, as shown in Fig. 397, which is from a photograph made with a large plane grating and a lens of 3 metres focal length. These bright lines represent rotations of 270°, 450°, 630°, etc., and by measuring their positions with an eye-piece micrometer, the wave-lengths corresponding to these rotations were determined. The centre of the system, as we may designate a point midway between D_1 and D_2, becomes bright and dark in succession, as many as eight complete alterations having been observed in some instances. This corresponds to a rotation of 1440°. If the burner is removed the changes take place very

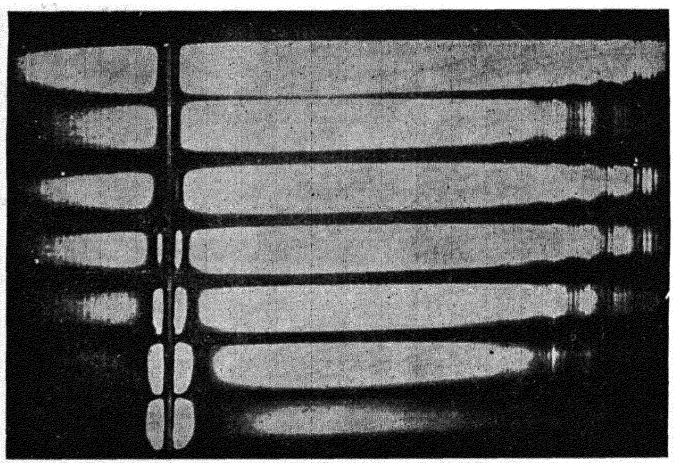


FIG. 399

rapidly, the centre "winking" bright and dark almost as rapidly as one can count.

The results obtained from measurements of the photographs are shown in the form of a curve in Fig. 398.

With vapor of considerable density the rotation is measured in a different manner. The Nicols are set in a parallel position, and the entire spectrum appears with the exception of the broad absorption band at the D lines. On each side of this absorption band a dark 90° rotation band appears. As we turn the Nicol these dark bands move, the one up, the other down the spectrum. By noting their positions we determine the values of λ corresponding to the rota-

tion of the Nicol in degrees measured from the position of extinction.

A series of photographic records obtained in this manner is shown in Fig. 399, which, taken collectively, exhibit the general form of the rotatory dispersion curve.

Owing to the great density of the vapor it was found that D_1 and D_2 could be considered as forming a single absorption band, and λ_m in Voigt's formula was given an intermediate value 5893. For a particular density and length of vapor column, the constant B was found from a single observation of δ. The observed and calculated values are given in the following table:

VALUE OF CONSTANT $B = 135{,}600$

λ	δ (OBS.)	δ (CAL.)
5980	5°	4°.47
5950	10°	10°.4
5933	20°	23°.2
5923	40°	38°
5917	66°	58°.9
5912.5	90°	89°.2
5874	90°	93°.1
5869	43°	43°
5864	40°	40°.6
5852	20°	20°.2
5833	10°	9°.2
5814	5°	5°.2

This table shows that with very dense vapor the rotatory dispersion is well represented by a single term formula, the observations being limited to a region not very near D_1 or D_2.

There is also rotation at many of the absorption lines of the band spectrum. This phenomenon will be discussed presently. The complete rotation spectrum is shown in the colored frontispiece, Fig. 5.

Ladenburg's Method of Determining Magnetic Rotation. — The rotation can be very accurately determined for *two* definite wave-lengths, one to the right, the other to the left of the absorption line, by employing a Savart plate (see Chapter on Polarization) between the polarizing Nicol and the gas-filled tube which shows interference fringes when the source is viewed through the analyzing Nicol. This method was originated by Ladenburg.[1] The fringes disappear when the polarizer is turned to a certain setting, and reappear with a shift of half of a fringe width, with a further slight rotation of the polarizer. If the polarizer is turned

[1] *Ann. der Phys.*, *38*, 249, 1912.

a degree or two from the zero position, so that fringes appear, and the magnet is then excited, the fringes disappear for two values of λ to the right and left of the absorption line, where the magnetic rotation has an opposite direction, but the same magnitude as the initial rotation of the polarizer from the zero position. In the region between these points and the absorption line, where the rotation is still greater, the fringes appear with the shift of a half fringe width alluded to above. This method, like the "Hacken-Methode" described in the Chapter on Dispersion Theory, enables us to obtain only two values for each observation, or photograph, but is more accurate than measuring the curved fringes obtained with the double prism of R and L quartz, described below.

Magnetic Rotation within an Absorption Band: Experiments of Zeeman. — It was established theoretically by Voigt [1] that, in the case of an absorption line separated into a magnetic doublet, by the Zeeman effect the rotation of the plane of polarization was positive for all periods lying outside of the components of the doublet and negative for all periods between the components, the light traversing the medium in the direction of the lines of force.

This was verified by Zeeman,[2] who made use of a method similar to the one which had been previously employed by Voigt in demonstrating magnetic double refraction.

The light of an arc lamp, after passage through a Nicol, was focussed upon the slit of a grating spectrometer, in front of the slit of which was placed a Fresnel bi-quartz prism, similar to the one employed by Macalus and Cobrino for the study of the rotation at the *D* lines.

This prism consists of two opposed wedges of right- and left-handed quartz through which plane-polarized light is passed in a direction parallel to the optic axes of the two wedges. The central ray A, traversing equal thicknesses emerges with the direction of polarization unchanged, for the rotation effected by one prism is unwound by the second. Rays to the right and left of A_1 such as B_1B are rotated more by one prism than the other and consequently emerge with their polarization direction rotated with respect to the middle ray by amounts proportional to their distances from A.

The Fresnel double-prism is placed as close as possible to the slit of the spectroscope, or an image of it is projected on the slit in the case of short focus instruments, and polarized light from an arc lamp passed through it. A second Nicol is mounted between the slit and the grating, which cuts off the vibrations which have

[1] *Wied. Ann.*, *67*, 359, 1899.
[2] *Proc. Amsterdam Acad.*, June, 1902.

been rotated into its plane of extinction by the quartz wedges, and the spectrum is traversed by a number of horizontal dark bands, corresponding to points on the slit at which the light is polarized in such a plane as to be stopped by the second Nicol. Between the first Nicol and the spectrometer the absorbing flame of sodium was mounted in a magnetic field, any rotation produced

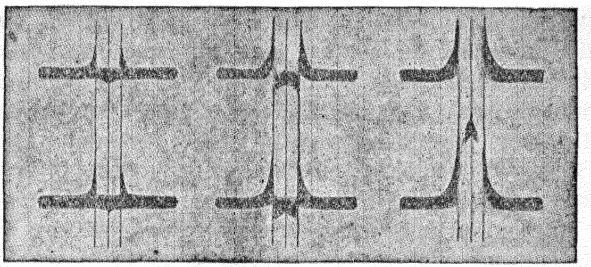


FIG. 400

by it adding itself to that produced by the quartz wedges. A rotation impressed upon any wave-length by the flame thus caused a vertical deviation of the dark band at the corresponding point of the spectrum, a shift equal to the width of a complete fringe corresponding to a rotation of 180°.

With a field of 15,000 c.g.s. units the dark absorption line was distinctly resolved into a doublet, and on increasing the amount of sodium in the flame the dark bands outside of the components

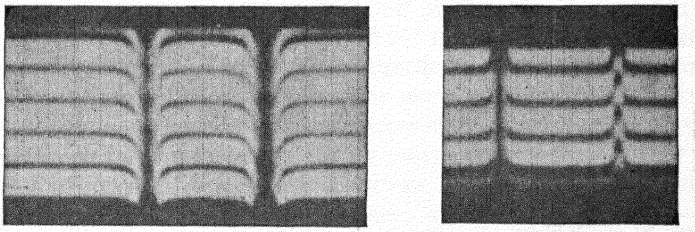


FIG. 401

curved upwards, while the portion of the band between them slid down in the opposite direction, as shown in Fig. 400, in which the appearances of the bands for different densities of the absorbing flame are shown. Photographs of the phenomenon are reproduced in Fig. 401.

Increasing the strength of the field caused the portion of the band between the components of the Zeeman doublet to move back towards its original position, which was in agreement with Voigt's

prediction that the negative rotation within the band decreased with increasing field strength. This is, of course, true only with fairly strong fields; in other words, for a given density of vapor, the negative rotation between the components reaches a maximum for a given strength of the magnetic field, after which it decreases with a further increase of the magnetization. No such maximum is found for the positive rotation outside of the components. The explanation of these effects will now be given.

Relation between the Faraday and Zeeman Effect. — The relation between the Faraday and Zeeman effect was pointed out by Larmor in his book *Aether and Matter*. We can express his conclusions in very simple language as follows: Assume that right-handed circularly polarized vibrations of frequency W are passing through an unmagnetized medium, and are reduced in velocity by the reaction of the atoms on the light-waves by a certain amount corresponding to the refractive index of the medium. The medium is now magnetized and the electron configuration in each atom rotates with a frequency ω (Larmor precession), say in the same direction as the circular light vibrations. If now we increase the frequency of the light to $W+\omega$ the refractive index will be the same as before, for we have increased the angular velocity of rotation of the light (*i.e.* its frequency) by an amount sufficient to compensate for the Larmor rotation of the electron configuration, and the circular vibrations will have practically the same phase relation with respect to the electron configuration as before. If the light is rotating in a direction opposite to that of the Larmor precession, we must reduce its frequency to $W-\omega$ to secure the same result. This means that the dispersion curves for right- and left-handed circular light for a magnetized medium are displaced laterally with respect to the curve for the unmagnetized medium, but are identical to it in form, the two frequencies $W+\omega$ and $W-\omega$ having the same refractive index as the original frequency W in the unmagnetized medium.

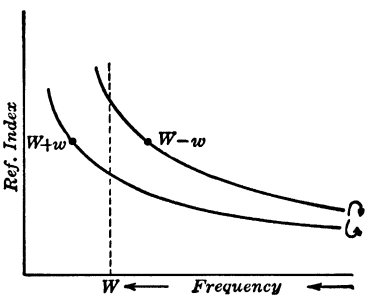


FIG. 402

The index for circular light of frequency W will be different according as it is right- or left-handed as shown by the two points at which the dotted line cuts the two dispersion curves (Fig. 402). This is all that is necessary to account for the rotation of the plane of polarization, for the incident plane-polarized vibra-

tion can be regarded as the resultant of two oppositely polarized circular vibrations, and these are propagated with different velocities, which as we have seen corresponds to a rotation of their resultant. Larmor gives the following deduction of Becquerel's formula for the rotation. The velocity of propagation of circular light of period $2\pi/\omega$ in a magnetic field H will be $V \pm (dV/dW)\omega$, where $\omega = eH/2mc$ (see Larmor precession in section on Zeeman effect) and $V =$ the velocity in absence of the field, the sign varying according as it is right- or left-handed.

If V_1 and V_2 are the velocities of the right- and left-handed components of an incident plane-polarized train of period $2\pi/\omega$ the rotation of the plane of polarization in length l of the medium will be $W(l/V_1 - l/V_2)$ the latter factor being the difference of the times of transit. In the present case it is thus $(lW/V^2)(dV/dW)$ so that the rotatory power of the medium is

$$\frac{lW}{V^2}\,\frac{dV}{dW}\,\omega.$$

If λ is the wave-length in vacuum and n its refractive index in the medium, $V = c/n$ and $W = 2\pi c/\lambda$. Thus the rotatory power is

$$\frac{le}{2mc^2} H\lambda \frac{dn}{d\lambda} \qquad \text{(electrostatic units)}.$$

The above agreement which applies to the normal precession velocity (normal Zeeman effect) can be applied also in cases where the effect is abnormal by substituting the modified value of the Zeeman displacement.

The following graphical method of explaining magneto-rotation is given by Zeeman. Magnetic rotatory dispersion differs from ordinary dispersion in that it has the same direction on opposite sides of the absorption line, and increases rapidly as the wave-length of the light is made more and more nearly equal to that of the line. That this is to be expected can be shown by combining the idea of the inverse Zeeman effect, with that of selective or anomalous dispersion. An absorption line of a vapor in a magnetic field is split into a doublet, one component of which absorbs completely right-handed circularly polarized light, the other left-handed.

There will accordingly be two dispersion curves for magnetized sodium vapor, one for right-handed, the other for left-handed circularly polarized light, separated on the wave-length scale by a distance equal to that between the components of the Zeeman doublet. These are shown in Fig. 403*a*. Remembering that plane-polarized light can be regarded as the resultant of two

oppositely polarized circular vibrations of the same period, it is clear from the figure that, outside of the region between the components of the absorption doublet a counterclockwise vibration will travel slower than a clockwise (since the refractive index is higher), consequently the plane of polarization (the resultant) will be rotated clockwise. Between the components of the doublet the reverse is true and for all wave-lengths within this region the rotation is in the opposite direction, and of greater magnitude (as was found by Zeeman), since the angular rotation is represented, for any wave-length, as the difference between the ordinates of the

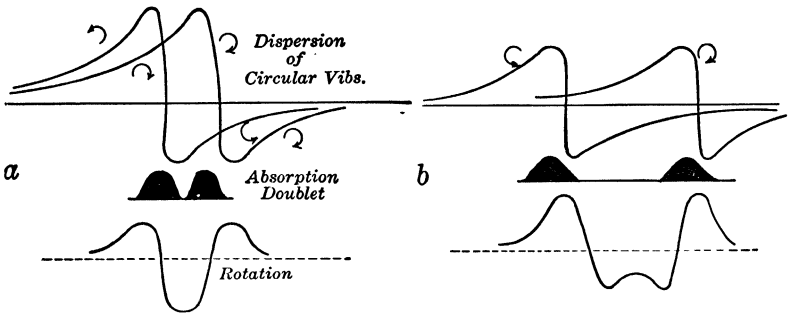


FIG. 403

two curves. The rotatory dispersion curve is shown below in each figure. As the field strength increases the components of the absorption doublet separate and with them the two dispersion curves, until we have the condition shown in Fig. 403*b*. In this case the difference between the ordinates of the curves between the components of the absorption line is less than in case *a* which explains why the rotation in this region decreases with increasing magnetic field, as was shown in the preceding section.

Ladenburg and Minkowski [1] have shown that, by measuring the magnetic rotation in the vicinity of absorption lines, in cases where the density of the vapor is known, we can calculate the number of resonating (or dispersion) electrons per atom, or the so-called "oscillator strength" of the atom for the resonance frequency in question, as has been done also by measuring the dispersion in the vicinity of the absorption lines of a saturated vapor. The p values for D lines of sodium calculated from magnetic rotation being for D_2 0.66 and for D_1 0.33, the sum being practically unity or one electron per atom (the valence electron).

Paramagnetic Rotation. — As was mentioned earlier in the chapter certain substances, notably the salts of the rare earths,

[1] *Zeit. für Phys.*, *6*, 153, 1921.

compounds of chromium and iron and chloride of titanium exhibit a negative rotation, *i.e.* a rotation in the direction opposite to that of the current producing the magnetic field which was explained by Jean Becquerel and more completely by R. Ladenburg as due to the Langevin orientation of magnetic atoms in the field. As a result of this the circularly polarized components of the inverse Zeeman effect for the absorption bands of these substances, will be of different intensity, one even disappearing entirely if the temperature is so low that absence of thermic agitation permits of a complete orientation.

By a graphical method similar to that given for explaining the normal or diamagnetic rotation by combining the idea of dispersion with the Zeeman effect, Dorfmann [1] accounts for paramagnetic rotation by considering the absorption for one circular component much smaller than that for the other. This gives the selective (or anomalous) dispersion curve on a smaller scale for the former component than for the latter.

If now the two dispersion curves corresponding to right- and left-handed circularly polarized light are plotted with a lateral displacement equal to the separation of the components of the Zeeman doublet (the fainter component corresponding to left-handed rotation) and the magnetic rotation computed as before from the difference between the values of n for left-handed and right-handed vibrations, we obtain the curve shown in Fig. 404 in which the black rectangles represent the Zeeman doublet into which the absorption line splits in the magnetic field. The refractive index curves are dotted and the curve of magnetic rotation is solid, the direction of rotation reversing as the absorption doublet is crossed.

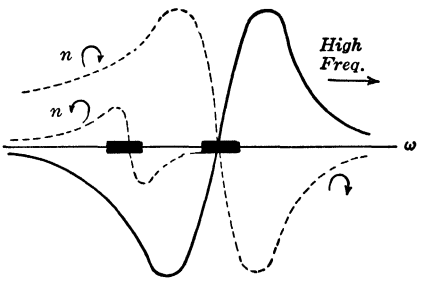


FIG. 404

If one absorbing circular component is absent the rotation is the resultant of the ordinary dispersion curve, determined largely by ultra-violet absorption, with the magnetic dispersion curve of the single circular absorption component. This method and Fig. 404 illustrating it are given by Ladenburg in the last edition of Müller-Pouillet's *Physics* where a more exhaustive treatment of the whole subject will be found.

As we have seen the quantum theory permits of two orientations

[1] J. Dorfmann. *Zeit. für Phys.*, *34*, 898, 1925.

of atoms which have a magnetic moment, either parallel or antiparallel to the field (shown by Stern-Gerlach experiment). If now an electron configuration capable of responding to a certain frequency, causing absorption of that frequency, is placed in a magnetic field, the two possible orientations will occur, and the atoms in one orientation will respond to and absorb circularly polarized light travelling in the direction of the field, of a slightly different frequency (one component of inverse Zeeman doublet) while those in the other orientation will absorb the other circular component. Now if the number of atoms in one orientation is less than the number in the other, the absorption of one circular component will be less than that of the other and we shall have a dissymmetry in the intensities of the two components of the Zeeman doublet seen in absorption, which, as we shall see presently, accounts for paramagnetic rotation and anomalous rotatory dispersion.

At high temperatures thermic agitation destroys more or less the orientation of the atomic magnets, and the effects described above become conspicuous only at very low temperatures. Moreover, at very low temperatures, one circular component in absorption may disappear entirely, as was shown by Jean Becquerel, in the case of the very narrow absorption bands of mineral crystals containing the elements of the rare earths group.

This has been explained as due to an increase in the number of atoms with electrons in the lower magnetic level. If the upper level is single and the lower double, then at very low temperatures there will be a tendency for atoms in the lowest energy state to preponderate, and as the high frequency (short wave-length) component of the inverse Zeeman doublet, results in a transfer of electrons from the lowest of the two lower levels to the upper single level, this component becomes enhanced at the expense of the other. There are of course other less simple cases for which the above argument does not hold.

It follows from these considerations that the paramagnetic rotation is negative on one side of the absorption band produced by the paramagnetic part of the molecule: the effect increases inversely proportional to the absolute temperature, and in some part of the spectrum increases with decreasing wave-length. This rotation is always combined with the common diamagnetic rotation, a general property of all substances produced by the Larmor precession, and it depends on the relative magnitude of the two effects which one preponderates.

Zeeman Effect and Magnetic Rotation of Crystals at Low Temperatures. — Some very remarkable experiments have been made

by Jean Becquerel, and by H. du Bois and J. Elias [1] with crystals at low temperatures, which throw a great deal of light upon the much disputed question of the nature of the magnetic rotatory dispersion in the vicinity of absorption bands.

The effects of low temperatures on the absorption bands of tysonite and xenotime have been mentioned in the Chapter on Absorption. Becquerel placed the crystals at the temperature of liquid air in a strong magnetic field and found that the position of the bands changed, the phenomenon being of the same nature as the Zeeman effect, only on a much greater scale, the separation of the components of some doublets amounting to a distance greater than that between the *D* lines. Still more remarkable was the discovery of the fact that, in the direction of the magnetic field, the bands corresponding to the absorption of the circular components of given sense are not all displaced in the same direction in the spectrum. The effect is very clearly indicated in Fig. 405 for

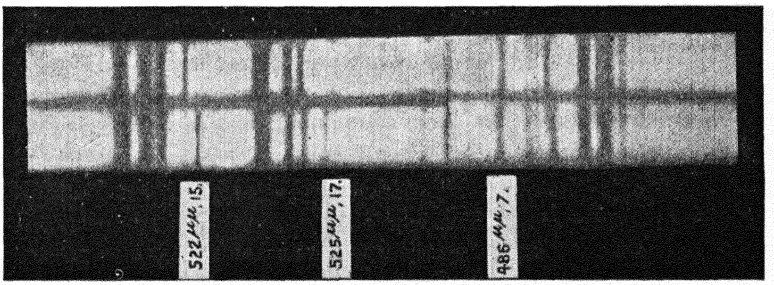


FIG. 405

the two absorption bands 5221 and 5252 of xenotime. This photograph was made by Becquerel with a Rowland concave grating in the second order, and the scale is 1 mm. = 1.6 Ångström units.

The upper spectrum represents the absorption when we employ circularly polarized light of a certain sense, passing it through a crystal plate 8 mms. thick, immersed in liquid air in a strong magnetic field. On reversing the direction of rotation of the circular light we obtain the lower spectrum. As will be seen, the 522 band moves to the *right*, the 525 one to the *left*. The latter shows dissymmetry clearly, the line having almost disappeared in the upper spectrum. This effect was found by Becquerel only at very low temperatures (−259°). The component which the magnetic field displaces towards the blue increases at the expense of the other component.

One of the most important points brought out was that, in

[1] *Ann. der Phys.*, *27*, 233, 1908.

every case, the magnetic rotation was in the same direction on opposite sides of the double band into which the magnetic field splits the original band, and in the opposite direction within the band, precisely as with absorbing vapors such as sodium.

Cases in which we appear to have opposite rotations, such as the neodymium band studied by the author, are shown to result from superpositions of bands, which can be separated by lowering the temperature. Moreover, if we have dissymmetry in the intensity of the circularly polarized bands to such extent that one practically disappears, we find what appears to be a band showing opposite rotations on its two edges. In reality one of these rotations is the rotation within the band analogous to the negative rotation observed between the Zeeman components of the sodium lines.

Becquerel found that the longitudinal inverse Zeeman effect shown by a majority of the crystals containing rare earths appeared as a simple circularly polarized doublet, but with incomplete polarization, *i.e.* when circularly polarized light was passed through the crystals, both components of the doublet appeared, one of them being very faint, however. This was subsequently found to be due to the fact that each component was in reality a close doublet, with components of very unequal intensity and oppositely rotating circular polarization.

The magnitude of the separation of the Zeeman components in the case of some of the absorption lines of these crystals was as much as nine times the normal separation, and in some cases reversed circular polarization occurred in the longitudinal effect.

No very satisfactory theory of these effects has been proposed up to the present time, though Ladenburg considers that many of them can be explained as paramagnetic rotations.

Becquerel, in comparing the magnetic rotation in the vicinity of an absorption line of a crystal, to the rotation of a gas, such as sodium vapor at low pressure, calculated that the ratio of the number of "dispersion electrons" to the number of crystal atoms was $1 : 10^8$ and that cooling the crystal to $-188°$ caused a threefold increase in the number, as shown by the increase in the magnetic rotation, and anomalous dispersion.

More recently Becquerel, Kammerlingh Onnes and W. de Haas have found that the negative rotation for wave-lengths not in the immediate vicinity of absorption bands increased proportionally to $1/T$, an effect attributable to paramagnetic action. The rotations of a plate of tysonite 1 mm. thick at 20°.4 absolute (the boiling point of hydrogen) and 4°.2 (liquid helium) for various wave-lengths are as follows:

	6391	4850	3800
20°.4	59°	113°	230°
4°.2	228°	458°	941°.

Voigt Effect: Magnetic Double Refraction. — Voigt's theory predicted magnetic double refraction for light traversing an absorbing medium in a magnetic field in a direction perpendicular to the lines of force. Zeeman gives a graphical method of showing this, similar to the one previously given for explaining the positive and negative rotation for wave-lengths outside of and between the circularly polarized doublet of the inverse effect.

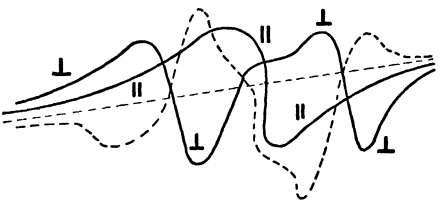


FIG. 406

The double refraction results from the circumstance that plane-polarized light with its electric vector E parallel to the field is propagated with a velocity different from that of light with E perpendicular to H. For a medium giving a Zeeman triplet by absorption the two dispersion curves for parallel and perpendicular polarized light are plotted as in Fig. 406 and the resultant (difference between the ordinates) is shown by the dotted line.

Voigt [1] placed a sodium flame between the poles of a magnet and passed through it the light of an arc, polarized at 45° with respect to the field direction, analyzing the emergent beam with a Babinet compensator placed in front of the slit of a spectroscope provided with a Nicol prism mounted immediately behind the slit. The spectrum was thus traversed with horizontal dark bands, which curved sharply in opposite directions on adjacent sides of each sodium absorption line, as called for by his theory, and as indicated by the dotted line of Fig. 406 for the regions outside of the triplet.

Zeeman [2] and Geest first observed the complicated patterns which occur when the experiment is made with a resolving-power sufficient to enable observation of the deformation of the bands between the components of the quartet and sextet into which D_1 and D_2 are separated. They calculated the effects to be expected by drawing dispersion curves similar to those of Fig. 406 and verified them by visual observations, illustrating their paper with drawings.

In 1912 Voigt and Hansen photographed all of these effects with a large plane grating in the third-order spectrum.

[1] Voigt, *Göttinger Nach.*, 5, 1902.
[2] Zeeman and Geest, *Proc. Amsterdam Acad.*, Dec., 1904.

Cotton-Mouton Effect. — There is another type of magnetic double refraction which was discovered by Cotton and Mouton which occurs when light passes through pure liquids in a direction perpendicular to the magnetic field. This is analogous to the phenomenon of electric double refraction discovered by Kerr, which will be considered in the next chapter.

Magneto-Optics of Resonance Radiation. — The very surprising discovery was made by Wood and Ellett [1] in 1923 that the almost complete plane-polarization of the resonance radiation of mercury vapor, excited by plane-polarized light of wave-length 2537, was completely destroyed by the magnetic field of the earth, if properly oriented with respect to the apparatus. An extended study of the action of weak magnetic fields on the resonance radiation of both mercury and sodium vapor, showed that the field could destroy polarization when originally present, and produce polarization in cases in which it was initially absent.

In the case of mercury the source of light used was a small Cooper-Hewitt mercury arc in quartz, partly immersed in a large vessel of water to keep down the vapor pressure within the arc and to prevent reversal of the λ2536.7 line. It is important to neutralize the earth's field with a Helmholtz coil and to keep the air of the room free from mercury vapor by proper ventilation and by avoidance of spilled mercury.

The radiation from a wide slit, placed close to the arc, was passed through two large quartz lenses, arranged to give a slightly convergent beam, and then through a quartz prism of about 60°, the base of which was perpendicular to the optic axis. With lenses of crystalline quartz this disposition is necessary, since if the prism is placed in the usual position between the lenses the rotatory power of the second lens introduces polarization in all azimuths in both images, though the two beams which emerge from the prism are plane-polarized. Later on in the work a lens of fused quartz was employed for converging the polarized beam. This arrangement is preferable as smaller and more intense images of the slit are obtained when the prism is mounted between the lenses. With this arrangement it is possible to illuminate the vapor in the bulb with a polarized monochromatic radiation of wave-length 2536.7 with the electric vector either vertical or horizontal. The two polarized spectra formed by the prism are received on a screen of barium platino-cyanide which shows the highest phosphorescence at the two images of the slit formed by the 2536 radiation. A screen of black paper, perforated by a hole 1 or 2 mms. in diameter coated on one side with the phosphorescent barium

[1] R. W. Wood and A. Ellett, *Proc. Roy. Soc.*, *103*. 396, 1923; *Phys. Rev.*, *24*, 243.

compound, is now mounted in contact with the quartz bulb and one of the polarized images brought into coincidence with the small hole. We now know that the bulb is traversed by the concentrated ultra-violet radiation which excites the resonance radiation.

The polarization of this light has to be studied by photography, of course. When the polarization is very feeble its presence is best shown by mounting in front of the bulb a quartz wedge of small angle cut parallel to the optic axis, and photographing the fringe system formed by the wedge, through a double image prism of quartz mounted in front of the quartz objective of the camera. This method can be used only with monochromatic light. The behavior of the wedge is first studied by mounting it between two Nicols pointed at a soda flame. In this way we find the position of the wedge with respect to the electric vector of the light, which gives the fringes at maximum visibility. Next we substitute for the first Nicol one or two inclined glass plates which produce feeble polarization, shown by the low visibility of the fringes. This preliminary study is very important as we must be thoroughly familiar with the action of the wedge under all conditions, as it is to be employed in the study of invisible light.

The arrangement of the apparatus is shown in Fig. 407. When measurements of the degree of polarization were to be made the quartz wedge A was removed and the times of exposure were varied so that the weaker of the two images formed by the double image prism B had in one exposure the same density as the stronger image in the other. The resonance lamp D consists of an evacuated quartz bulb 2 cms. in diameter, blown as thin as possible, to which is attached a long side tube containing a drop of mercury. This side tube was immersed in a cooling bath in order to reduce the pressure of the mercury vapor within the bulb and thus eliminate unpolarized secondary resonance set up in the vapor outside the path of the primary beam. This secondary resonance is largely unpolarized and at room temperature practically obscures the fringes in the main beam. Its intensity, however, may be reduced to less than 2% of that due to the main beam by cooling the side tube in a mixture of ice and salt.

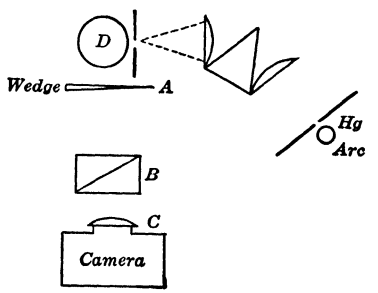


FIG. 407

The effect of impressed magnetic fields in various orientations

with respect to the incident light beam and its plane of polarization was now investigated. The field used for this purpose was that near the end of a short solenoid about 40 cms. in diameter. Investigation of the field of this solenoid with a flip coil showed that a field uniform to within less than 1% could be produced throughout the volume of the resonance bulb. We will consider three typical cases:

Case I. When the direction of the impressed magnetic field H is perpendicular to the exciting beam and to its electric vector E (Fig. 408, Case I) the polarization of the light emitted in the

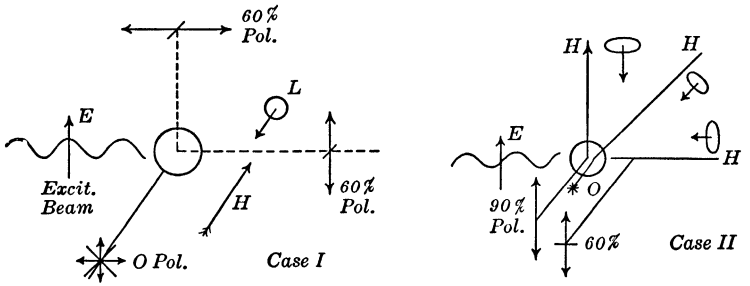


Fig. 408

direction of the field is rapidly decreased with increasing strength of the field. The arrowed circle L in the figure will be explained presently.

That the failure to obtain 100% polarization in the absence of the field (the expected value for an isotropic oscillator) was not due to collisions occurring between absorption and reëmission, was proved by reducing the density of the mercury vapor by cooling the side tube to $-50°$, the vapor over *solid* mercury showing the same percentage of polarization as before. An exposure of 15 minutes with the mercury at $-50°$ gave an image of about the same density as a 20-second exposure at $-18°$.

With a magnetic field of sufficient intensity to destroy the polarization in the rays parallel to H a threefold increase in intensity and 60% of polarization was found for the ray leaving the bulb parallel to the electric vector of the exciting beam (normally of intensity $\frac{1}{10}$ that of the ray perpendicular to E), while the 90% polarization of the ray nearly parallel to the exciting beam was reduced to 60%.

Case II. Case II with the electric vector of the exciting beam vertical as before, the magnetic field is turned gradually from the vertical direction to the horizontal (Fig. 408, II). The state of polarization for the emergent beam (perpendicular to paper) for

the three positions of the field is shown by the arrows on the three parallel rays, each ray joined to the H line associated with it. For the vertical field we find 90% polarization, for the oblique field zero polarization and for the horizontal 60%. Careful observation showed that the plane of polarization rotated with H as the percentage decreased. Light emitted horizontally (to the right) parallel to H is unpolarized. With H parallel to E 100% polarization is to be expected, instead of 90% as found.

The cause of this discrepancy was found by Keussler [1] who, in repeating the experiments, with a more accurate method of measuring the polarization, found only 80%, with weak fields, *but* 100% in a strong one, of 7000 Gauss.

He explained this as due to the hyper-fine structure of the 2536 exciting line which was first correctly determined by the author.[2]

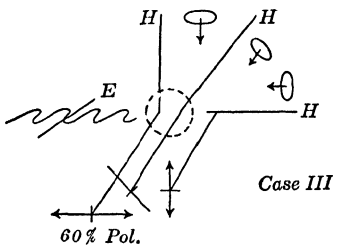


FIG. 409

CASE III. Electric vector horizontal and the magnetic field rotated as in previous case. The emitted light is 60% polarized in each case, with its electric vector perpendicular to the field, the plane of polarization turning with the field, without going through the zero stage as in Case I.

Theory of the Effects. — Most of these effects were explained by Wood and Ellett by the rather artificial and formal method, suggested by C. Darwin, and based on Lorentz's first method of explaining the Zeeman effect.

CASE I. The motions of the electrons are considered as decomposed into three simple motions, a linear vibration parallel to the field, and two opposed circular motions perpendicular to the field (L in the figures). In Case I the two circular motions are excited by the electric vector E of the incident light (Fig. 408) while the linear vibration is not (as it is perpendicular to E). The resultant of the two opposed circular motions, seen in the direction of the field H, is unpolarized light.

CASE II. Here we have, with H vertical only, the linear oscillation parallel to E excited, and with H horizontal the two circular excitations and not the linear which is now perpendicular to E. In both cases the light emitted as shown in Fig. 408 is plane-polarized.

CASE III. In this case the linear component is always perpendicular to E and therefore never excited while the planes of

[1] Keussler, *Phys. Zeit.*, *27*, 313; *Ann. der Phys.*, *82*, 793.
[2] *Phil. Mag.*, *50*, 764, 1925.

rotation of the circular components are always parallel to E and therefore emit plane polarization which turns with H, as the plane of rotation is always perpendicular to H.

Explanations of all of the effects from the viewpoint of the quantum theory were given in numerous papers cited below.[1]

On the quantum theory we can explain Case I as follows:

Since the light is entering the vapor perpendicular to the magnetic field, the perpendicular components of the inverse Zeeman triplet are absorbed, electrons being raised from the lower (single) level $m=0$ to the upper magnetic levels $m=\pm 1$. Bohr's polarization rule is that transitions in which m changes by ± 1 give the perpendicular (or circular) Zeeman components, while transitions involving no change of m give the parallel components. On their return to the lower level they emit circularly polarized light of opposite rotations parallel to field, which we see as unpolarized light, since each atom contributes only one circular vibration, and all rays are incoherent. Perpendicular to H we have plane-polarized light. The question immediately comes up as to why we have polarized light in the absence of the field, with no orientation of the atoms. Bohr's explanation of this is that with no orientation, an atom behaves as a quasi-elastic isotropic oscillator.

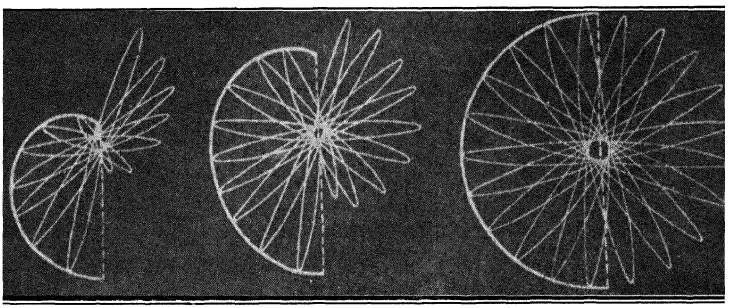
Case II. On the quantum theory with H vertical and observation direction horizontal as in Fig. 408, the absorption as transition will be from the lower level $m=0$ to the upper level $m=0$, since E is parallel to H and the absorption is for the middle or parallel component of the inverse Zeeman triplet. The light emitted as shown in Fig. 408 will be plane-polarized parallel to E in each case. The calculation of the angle of 54° for zero polarization will be found in a paper by Eldridge [2] based on the Larmor precession, and in *Atoms, Molecules and Quanta* by Ruark and Urey based on the quantum theory.

Hanle's Observation of the Larmor Precession. — Shortly after the publication of the results obtained by Wood and Ellett, Hanle drew attention to the fact that rotation of the plane of polarization was to be expected in fields too weak to completely destroy it. On the classical theory of the Zeeman effect, we have seen that an electron, bound by a quasi-elastic force, changes its linear path to that of rosette form, when placed in a magnetic field. In a

[1] Hanle, *Naturwissenschaften, 11*, 690, 1923; *Zeit. für Phys., 30*, 93, 1924; Breit, *Phil. Mag., 47*, 832, 1924; *J. O. S. A., 10*, 439; *Phys. Rev., 25*, 242; Rasetti, *Lincei Rend., 33*, 38, 1924; Gaviola and Pringsheim, *Zeit. für Phys., 25*, 667; *34*, 1; Foote, Ruark, and Mohler, *J. O. S. A., 7*, 415, 1923; Joos, *Phys. Zeit., 25*, 130, 298, 400, 1924; Pringsheim, *Zeit. für Phys., 23*. 324; Van Vleck, *Proc. Nat. Acad., 11*. 612, 1925.

[2] J. A. Eldridge, *Phys. Rev., 24*, 234. 1924.

very weak field, only a small portion of the complete rosette may form before the radiation ceases. This state of things is represented in Fig. 410*a*, in which the radiation has died out before a complete rotation of the plane of polarization has occurred. The initial plane of polarization in zero field is shown by the dotted line and the emitted light will be partially depolarized and rotated in clockwise sense. Figures 410*b* and *c*, show the effect of stronger fields, the polarization being completely destroyed in the latter. This rotation is in the direction of the magnetizing current and is in reality



a *b* *c*

FIG. 410

the Larmor precession. This rotation was observed by Hanle in the case of mercury resonance radiation, and subsequently by Ellett for sodium and by Soleillet for cadmium vapor. Ellett [1] operated with sodium vapor at a pressure corresponding to only 90° C., obtaining rotations which increased from a few degrees at 3 Gauss to 21° at 13 Gauss.

By calculating the angular velocity of the Larmor precession for the weak magnetic field giving a measured rotation Hanle determined the time that the electron remained on the upper level to which it had been raised by absorption. This time-interval, 10^{-7} sec., was in good agreement with Wien's measurement with canal rays.

Effects of Rapidly Alternating Fields. — Calculations showed that the alternating magnetic field of concentrated sunlight was of the order of magnitude of 3 Gauss, and Wood and Ellett showed by experiment that no depolarizing action on mercury resonance radiation was produced by such radiation after passage through a dark red filter.

Rasetti and Fermi [2] and Breit and Ellett [3] experimented with

[1] Ellett, *Nature*, 1925; *Jour. Optical Soc. Am.*, *10*, 427.
[2] Rasetti and Fermi, *Zeit. für Phys.*, *33*, 246.
[3] Breit and Ellett, *Phys. Rev.*, *25*, 888.

high frequency magnetic fields obtained with electron tubes oscillating at frequencies of from 1 to 5 million; at the highest frequencies there was no depolarizing action at all, for the field reversed before the Larmor precession had turned the plane of polarization through more than a small angle. A small rotation of the plane in the opposite direction occurred with the reversal of the field. This to-and-fro oscillation of the plane of polarization (which showed of course as a slight depolarization) became less and less as the frequency increased. This in fact was an experimental measurement of the rotation velocity of the Larmor precession, which turned out to be about the same as that calculated from the Zeeman effect.

Magneto-Optics of Sodium Resonance Radiation. — Similar experiments were made with sodium vapor. The lamp for exciting the resonance consisted of a long hydrogen tube 5 mms. in diameter provided with a short lateral tube containing sodium. The central portion with the sodium tube passed through an oven, the temperature of which determined the rate at which sodium vapor diffused into the main tube, through which a slow stream of hydrogen passed. The electrode bulbs were outside of the oven, and the hydrogen was admitted at one end through a long and very fine capillary and pumped out at the other end, at such a rate as to keep the pressure in the discharge tube at about half a millimetre. This lamp proved superior and more durable than any form of sodium arc or discharge tube, and excited a very intense resonance. The new hot cathode sodium arc would be still better.

The light from a slit placed close to such a discharge tube, and rendered parallel by a condensing lens of short focus, was passed through a large Nicol prism and then focussed at the centre of the resonance lamp which consisted of a pyrex glass bulb about 3.5 inches in diameter to which were attached two side tubes, one for exhausting and the other containing a mixture of shavings of metallic calcium and common salt. By heating the latter, sodium is set free and may be distilled over into the resonance lamp as required.

The latter was kept in communication with the pump and held at a temperature between 140° and 180° in a current of hot air rising from a large asbestos chimney with a ring gas burner at the bottom.

The radiation is not as completely polarized as that of mercury, nor is its polarization as easily affected by a magnetic field. When excited by plane-polarized light in the absence of a magnetic field the radiation emitted perpendicular to the exciting beam and to its

electric vector is about 6% polarized. This low initial value of the polarization may be due to the presence of hydrogen, which it seems impossible to eliminate entirely. It is well known that the presence of very small traces of hydrogen greatly reduces the polarization of the 2536 resonance radiation of mercury.

The effect of an impressed magnetic field is in general the same with sodium as with mercury, though much more intense fields are required and the values of the polarization are never as great.

In the case of mercury the effect of the field reaches practically its maximum value with intensities of 2 or 3 Gauss, while with sodium fields of 80 to 100 Gauss are required.

The chief point of interest in the case of sodium resonance is the fact, pointed out by Gaviola and Pringsheim, that in the magnetic field only the light of one of the two yellow lines (D_2) is polarized. An image of the polarized fringe system obtained with a Savart plate was projected on the slit of a spectroscope and the fringes appeared in the D_2 line only. This is explained by the difference between the magnetic levels of the two lines, shown in Fig. 411.

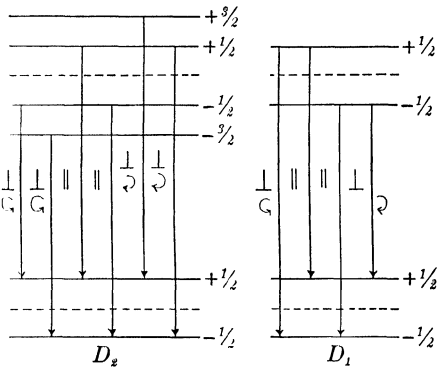


FIG. 411

As we have seen transitions in which the magnetic quantum number m does not change, give rise to an emission of light perpendicular to the field, plane polarized parallel to the field, and no radiation in the field direction, while transitions in which m changes by ± 1 cause plane polarization perpendicular to the field and circular polarization along the field.

Referring now to Fig. 411 for the case of exciting light parallel, and emitted light perpendicular to the field, for D_1, we see that the absorption is *for the two circular components only* by which electrons are raised to the two upper levels from each one of which they can return to either of the two lower levels, the emission consisting of a mixture of perpendicular and parallel polarization in equal amounts, or unpolarized light.

In the case of D_2, however, it is only by absorption of the two outer circular components that levels $+\frac{1}{2}$ and $-\frac{1}{2}$ are reached, these being the only levels from which returns giving parallel polarization are possible. Absorption of the inner circular com-

ponents of the sextette raise electrons to levels $\pm 3/2$ from each of which only a single return transition is possible, these giving plane polarization perpendicular to the field.

The mixture of parallel with a larger amount of perpendicular polarization resulted in a showing of about 30% of polarization in the reëmitted radiation. By applying a magnetic field of 1200 Gauss, Ellett succeeded in obtaining 100% polarization, the two outermost components of the inverse Zeeman sextette being separated by a distance greater than the width of the exciting line. In consequence of this no absorption transitions to upper levels $\pm 1/2$ occurred, and no parallel polarization was present in the reëmitted light.

Magneto-Polarization Effects in Vanishing Fields. — We must now examine a little more critically some of the effects which have been treated in the previous sections. Going back to the first case of mercury vapor in which 90% of polarization was found with the electric vector of the exciting beam vertical (Fig. 409) in the absence of a magnetic field, and zero polarization with a field of 1.5 Gauss parallel to the direction of observation, we will consider what happens as the field diminishes in strength. We are accustomed to consider the oppositely circularly polarized components of the longitudinal Zeeman effect as coming from different atoms, and therefore incoherent. In a strong field we can separate their radiations with a spectroscope, while with a weak field we recognize them (superposed) by the absence of polarization. If this assumption of incoherence is assumed to hold for the weakest fields we come to the conclusion that, no matter how weak the field, there must be complete depolarization, which is contrary to experimental observation, for as we have seen full polarization occurs in zero field and a rotation of the plane of polarization as the field rises from zero to the value giving depolarization. This means that the magnetic energy levels are no longer sharply defined in very weak fields, and in zero field there is no magnetic quantization at all; a condition known as degenerate. The energy change involved in transition from one level to another is independent of any change of orientation of the atom occurring during the transition. Bohr considered that such a degenerate system behaved like an isotropic oscillator and in this way explained the 90% polarization of mercury vapor in the absence of any field.

By applying a magnetic field parallel to E, the electric vector of the exciting beam, we still have 90% of polarization with mercury, for the oscillation of the electron (of the virtual oscillator) is parallel to the lines of force, and therefore uninfluenced by them. Bohr's conception was applied only in the case of atoms showing

the normal Zeeman effect, but Heisenberg extended it to cases involving an abnormal effect, obtaining the expected polarization for *zero* field, by considering a small field applied parallel to E (which is without influence on the virtual oscillator), and calculating the polarization in this case from the magnetic energy levels and the known intensities of the Zeeman components, as explained in the previous section. In this way a calculated value of 60% of polarization for excitation by the D_2 line alone, and zero polarization for D_1 excitation is obtained.

This was tested by Datta [1] who excited the resonance with the polarized light of the D_2 line, employing the author's quartz block method of separating D_1 from D_2 light, which was described in the Chapter on Polarization, page 349.

Polarization in Step-Up Optical Excitation. — Hanle has given us a very striking proof of the reliability of the method given in the last section of calculating the polarization of resonance radiation in the absence of any magnetic field, by assuming a field applied in such a direction as not to affect the electron motion of a classical oscillator, *i.e.* parallel to electric vector, and then making use of the magnetic energy levels.

Hanle and Richter [2] excited the mercury tube with polarized light and found the lines of the emission spectrum plane-polarized in different percentages *and in different directions*. Richter had observed that if the light from the quartz mercury arc was focussed at the centre of the resonance bulb through a double-image prism of calcite one cone of radiation was green and the other blue. In the first case the electric vector of the exciting light was perpendicular, in the second parallel to the direction of observation.

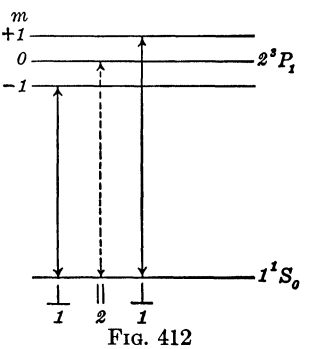


FIG. 412

Applying the method of Bohr and Heisenberg to the "step-up" excitation, we consider a magnetic field applied parallel to the electric vector of the exciting light. In the Zeeman effect for mercury the lower S level remains single and the 2^3P_1 level splits into three magnetic levels as shown in Fig. 412. The exciting beam is polarized with its electric vector E parallel to the field H, consequently only the parallel component is absorbed (dotted arrow pointing up), and emitted (dotted arrow pointing down) the emitted light is polarized par-

[1] G. Lal Datta, *Zeit. für Phys.*, *37*, 625, 1928.
[2] Hanle and Richter, *Zeit. für Phys.*, *54*, 811.

allel to E. In the second stage of the excitation (Fig. 413), the absorption occurs from the 2^3P_0 metastable level, to which electrons have been brought from 2^3P_1 by collisions with nitrogen molecules, and as this level remains single in the magnetic field, it matters not how the electrons reach it, and the first stage of the excitation need not be considered. By the absorption of 4046 transition b occurs (the parallel component) and from the 0 level reached, emission transitions occur as follows: For the 4046 line transition b, giving complete polarization *parallel* to H. For 4358 — transition f and g giving complete polarization *perpendicular* to H. For 5461 — transitions m, n, o, giving 14% polarization parallel to H, the percentage being calculated from a parallel component of intensity 8 and two perpendicular components of intensity 3.

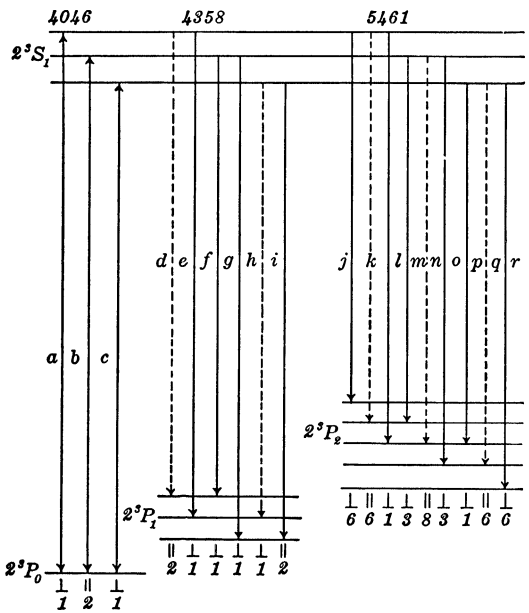


FIG. 413

Under certain conditions Hanle found that the percentage polarization in a magnetic field was increased or decreased by the presence of a foreign gas. If one of the inert gases, helium for example, at fairly high pressure is mixed with the mercury vapor, it has very little effect upon the intensity of the resonance radiation, while hydrogen at very low pressure destroys it, as was first shown by the author, and Hanle found that argon at 2 mms. decreased the polarization without changing the intensity. The collisions with argon atoms change the orientation of the mercury atom without damping the vibration.

Hydrogen on the other hand, which damps the vibration, reduced the intensity but diminished the rotation of the plane of polarization and increased the percentage of polarization. In other words the emission was extinguished by collisions with hydrogen atoms before the precession had time to manifest itself.

Magneto-Optics of Band Spectra. — It was thought for many years that the fine lines which make up band spectra were unresponsive to a magnetic field. The first observations showing an influence were made by the author during the investigation of the rotatory power of magnetized sodium vapor in 1895 described previously. With dense vapor it was observed that, with crossed Nicols, a spectrum of numerous bright lines in the red-orange and green-blue region appeared as soon as the magnetic field was turned on, showing that rotation of the plane of polarization occurred in the vicinity of the lines which made up the complicated band spectrum, which we now know results from absorption by diatomic sodium molecules.

Similar results were later obtained with iodine and other vapors, and the spectra were named magnetic rotation spectra. The general appearance of the spectrum obtained with sodium is shown in Fig. 5 of the colored frontispiece. Over two hundred lines can be

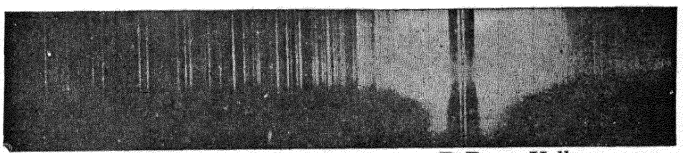


FIG. 414

counted in the red, and about one hundred and thirty in the blue-green, region. A photograph of a portion of the spectrum is shown in Fig. 414, the broad bands at the centre being due to the *D* lines.

An extended study of the magnetic rotation spectrum was made by F. E. Hackett [1] in collaboration with the author. A photograph of the spectrum in the blue-green region, in coincidence with the iron comparison spectrum, is shown on Plate 10, Fig. 2. Just above (Fig. 1) is a photograph of the fluorescent spectrum already discussed. The similarity between the two is obvious. Figures 3 and 4 show small portions of the magnetic rotation and fluorescent spectra in the green region on a larger scale. These photographs were made with a Rowland concave grating. All trace of the rotation disappears if hydrogen or nitrogen at a pressure of more than a few centimetres is present in the tube; this is also true for the fluorescence. Iodine vapor also gives a very beautiful bright-line spectrum. A few crystals are introduced into a small glass bulb which is highly exhausted and sealed off from the pump. This bulb, when placed between the perforated conical pole pieces of a magnet (Nicols crossed) and gently warmed, restores light of a

[1] *Astrophysical Journal*, xxx, No. 5, December, 1909.

most beautiful emerald-green color, which the spectroscope shows to be discontinuous.

Zeeman Effect for Band Spectra. — As has been pointed out already, the abnormally high rotation of the plane of polarization in the vicinity of some of the lines forming the band spectrum of iodine, first described by the author, and more completely investigated in collaboration with Ribaud, is evidence of a small inverse Zeeman effect for these lines. Repeated observation by various investigators failed to show any direct effect of a magnetic field on the emission lines of band spectra until 1908 when Dufour [1] recorded a direct, though very abnormal, effect in the case of the band of the sub-fluorides and chlorides of calcium (CaF), barium and strontium. He made his observations parallel to the field, employing a quarter-wave plate and Nicol, as originally used by Zeeman for converting the circular into plane vibrations.

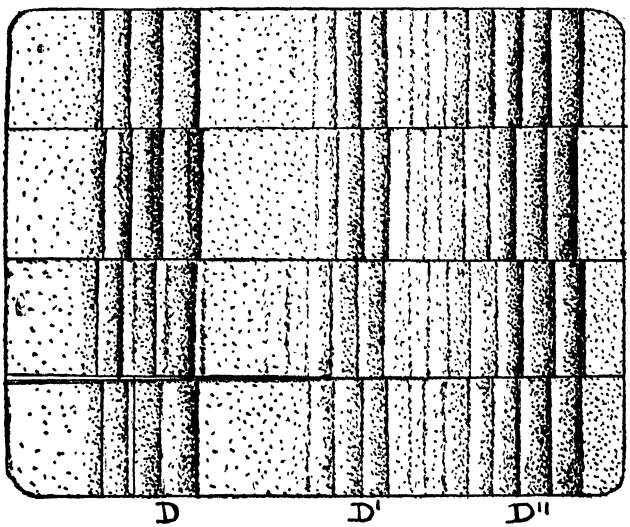


FIG. 415

The displacement of the lines indicated that in one band the circular components were of the normal type while in another band they were reversed. Figure 415 is a drawing made from Dufour's photograph, the upper and lower spectra without field and the middle spectra with field and the quarter-wave plate first in one position and then in the other. In the band D'' the shift indicates normal circular polarization, while in bands D and D' the shift is reversed, indicating reversed circular polarization.

[1] A. Dufour, *Le Radium*, *5*, 291; *Phys. Zeit.*, *10*, 124.

In the case of the transverse effect the parallel and perpendicular components form two almost coincident pairs, the middle component, seen in the normal effect, being absent.

Dufour also recorded that, in the case of the secondary (band) spectrum of hydrogen, now recognized as due to the molecule H_2, he found some lines that showed no effect and still others the reversed effect.

Fortrat [1] observed a similar behavior in the case of the nitrogen bands at 2370, 2479 and 2596. In 1913 he found also an effect similar to the Paschen-Back effect in the case of the cyanogen bands at 3000, the fine doublets, splitting up in a complicated manner in weak fields, fused together into a single line in strong ones.

Positive and Negative Rotation of Band Spectra Lines. — The rotation of the plane of polarization in the vicinity of the *D* lines is positive for wave-lengths lying outside of the components of the Zeeman doublet of the absorption line and negative for wave-lengths within it, as we have seen.

The lines of the band spectra are due to absorption by the molecule Na_2 and it is a matter of great theoretical interest to determine whether it is positive or negative.

During the course of the experiments on sodium vapor this question was very carefully investigated by the author.

The rotation is so slight, however, that the question could not be settled by merely rotating the analyzing Nicol, as this caused the whole continuous spectrum to appear, obliterating the bright lines.

The nature of the rotation in the band spectrum was finally ascertained by employing a Fresnel double prism of quartz (page 708). Wedges of about 12° angle were employed as it is desirable to have the bands rather broad, masking off, at the slit, everything except a very narrow strip of light along the top and bottom of a dark band. This reduces the amount of light scattered or diffused by the grating and darkens the background against which the lines are to appear.

With this arrangement of the apparatus the magnetically rotated line should penetrate the dark band from above or below, according to whether the rotation is positive or negative. If we excite the magnet and gradually heat the sodium tube, we see sharp needles of light shoot down from the continuous spectrum into the dark region immediately to the right and left of the *D* lines, as has been described by Macaluso and Corbino, Zeeman and others. The appearance of this spectrum is shown in colored frontispiece Fig. 6. If we reverse the magnetic field the needles of light shoot

[1] Fortrat. *Comptes Rendus*, *157*, 991. 1913.

up from below. The direction in which the plane of polarization is rotated by the *D* lines indicates that they are caused by negative electrons. The important question to be answered is whether the absorption lines of the band spectra rotate the plane of polarization in the same or in the opposite direction, and whether they all behave alike.

The magnetic rotation spectrum being much brighter in the red and orange than in the green and blue region, the first observations were made in this part of the spectrum. The spectroscope was a medium-sized instrument, consisting of a telescope and collimator of about 180 cms. focus, furnished with a plane grating.

The sodium tube was heated until the fine black absorption lines in the red appeared distinctly in the continuous spectrum above and below the horizontal dark band due to the Fresnel prism. The current was then thrown into the magnet, the self-induction of which is so great that the field did not rise to its full intensity for several seconds, so that there was plenty of time to see exactly what happened. As soon as the switch was closed numerous needles of light commenced to penetrate the dark region, *some of them shooting down from above, others shooting up from below,* as shown in Fig. 6 of the colored frontispiece. Of these, some only extended halfway or less across the dark band, while others crossed it completely. On opening the switch the luminous needles slowly withdrew from the dark background into the bright region from which they came. The phenomenon is most beautiful and shows us at once that some of the absorption lines rotate the plane of polarization in the positive direction, while others rotate it negatively.

A photograph of the phenomenon is reproduced in Fig. 416.

An attempt was made to study the direction of rotation by

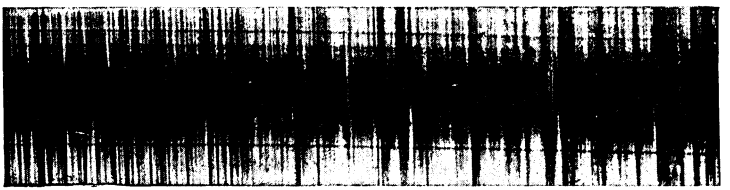


FIG. 416

Hackett and Wood in the blue-green rotation spectrum, but there was little or no trace of the bright needles of light shooting up or down into the dark interference band. Their failure to appear in the green is due to the smaller angular rotation of the plane of polarization.

The methods employed at the time for determining the directions of these small rotations were not wholly satisfactory and it was not until the subject was attacked again in collaboration with G. Ribaud,[1] using a very large Weiss magnet and iodine vapor contained in a small glass bulb, that a complete solution of the problem was obtained. Iodine vapor gives a rotation spectrum very similar to that of sodium, and is much less troublesome to work with.

Small glass bulbs about 2.5 cms. in diameter, highly exhausted and containing a crystal of iodine, were mounted between the poles of a large Weiss electromagnet. The bulbs were supported in a brass tube of 3 cms. internal diameter, furnished with two lateral holes for the passage of the light, and heated electrically by a spiral of nickel wire placed below the bulb (Fig. 417). A cover of mica forced the heated air rising around the bulb to escape through the side holes, and prevented iodine crystals from depositing on the walls in the path of the beam of light, as was invariably the case if the brass tube was open at the top. The source of light was a quartz-mercury arc arranged "end-on," the observations being restricted to the seven or more absorption lines of iodine which are covered by the broadened green mercury line.

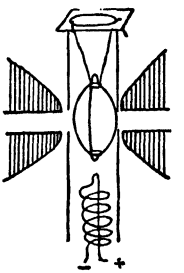


FIG. 417

Observations were made with a 20-plate Hilger echelon, each plate of 15 mms. thickness. The polarizing Nicol is rotated until sufficient light is restored to render the absorption lines visible (10 to 15 degrees). The magnet is then excited and the spectrum brightens at the points where the rotation is in the opposite direction to that in which the Nicol has been rotated, and darkens where the rotation is in the same direction.

It was immediately obvious, with this instrument, that certain lines became broader when the magnet was excited, and that a reversal of the direction of the field caused them to become so fine as to be almost invisible, as a result of the brightening of the regions bordering them. This brightening was in some cases much stronger on one side of the line than on the other.

Suppose that the rotation is positive to right and left of one absorption line, and negative to the right and left of another. The former will appear narrower when the field is excited as a result of the brightening of the edges of the dark line. The latter will, however, appear broader than in the absence of the magnetic field. If the rotation is anomalous, the centre of the absorption

[1] *Phil. Mag.* June. 1914.

line will appear slightly shifted as a result of its becoming brighter on one side and darker on the other. This method of observation gives results quite as conclusive as those obtained with the Fresnel prisms, and is well adapted to cases where the maximum rotation is less than forty or fifty degrees.

To obviate the necessity of reversing the field to observe these changes, a half-wave plate of mica was placed over one-half of the slit of the collimator, the principal directions of the plate coinciding with those of the polarizing prism. The two conditions seen with magnetic fields of opposite direction are thus visible simultaneously, one above the other and in exact coincidence. Any shift due to anomalous rotation would thus be doubled, but no shift was seen. The theory and method of using the half-wave plate will be found in the original paper.

The lines observed are numbered as in the section on resonance spectra.

The curves of rotation for these lines are shown by Fig. 418: they are only roughly quantitative.

In this group of lines the direction of the rotation changes as we pass from line to line, which explains perfectly why the angular magnitude of the rotation is so small in comparison with that exhibited by the *D* lines. The + rotation due to a given line is nearly neutralized by the − rotations of its two neighbors. We also see at once why the rotation on one side of a line may be much greater than on the other. For example, we have very strong rotation to the right of line 4 since the opposed rotation due to the faint line 4′ is very small, and the effect contributed by the next line 5 is of the same sign. The rotation to the left of 4 is, however, very small since line 3, with its opposite rotation, lies very close to 4.

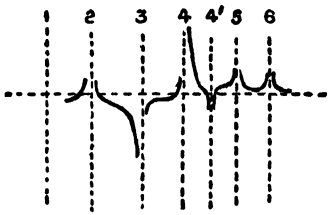


FIG. 418

It is very probable that the same condition holds for the vapor of sodium, at least in the green region. In the red and orange portion of the spectrum it is probable that the + and − rotations observed with the Fresnel double prism were due to close groups of lines with rotations of the same sign. It was observed that the luminous needles which penetrated the dark bands were almost invariably found on one side or the other of broad absorption lines, which were undoubtedly unresolved groups of fine lines. Suppose we have a group of a dozen lines, the first four rotating the plane of polarization in the same direction, while for the remainder the

sign changes in passing from line to line. It is clear that if the spectroscope does not resolve the lines the Fresnel prism will show a strong rotation in the vicinity of the first lines, that is on one side of the group, and no rotation at all on the other; in other words, we apparently have a broad line which shows rotatory power on one side only, which was exactly what was found in the earlier work with sodium.

Examination for the Zeeman Effect. — Since the selective rotatory power of the vapor in the vicinity of absorption lines can be explained by a longitudinal Zeeman effect, it was of interest to see whether any evidence of such an effect could be observed.

A double circular analyzer (two $\lambda/4$ plates of mica, one rotated through 90° with respect to the other) was mounted between the iodine bulb and the analyzing Nicol (azimuth 45° with respect to the neutral lines of the plates). One obtains in this way two fields of view separated by a fine line, one corresponding to right-handed, the other to left-handed vibrations.

If a longitudinal Zeeman effect exists, the absorption line, which runs across both fields of view as a continuous line in the absence of the magnetic field, should be displaced in opposite directions in the two fields as soon as the magnet is excited. No trace of such a shift was observed.

Since the absorption lines 3 and 4 are separated by a distance of about 1/20 A.U. we should have certainly been able to detect a shift of 0.01 A.U.

From this we must suppose, that if the Zeeman effect exists, it is less than 0.01 A.U. for a field of 20,000 Gauss.

CHAPTER XXII

ELECTRO-OPTICS

Faraday, immediately after his discovery of the magnetic rotation of the plane of polarization, sought in vain for some analogous phenomenon resulting from an electric field.

The first positive results were obtained in 1875 by Kerr, who imbedded the wire terminals of an induction coil in a block of glass placed between crossed Nicols and found that there was a restoration of light when the coil was in operation. The restoration was not instantaneous, however, first appearing about two seconds after the coil was set in operation, and not reaching its maximum value for thirty seconds. The restored light could not be extinguished by rotation of the analyzing Nicol, which showed that the phenomenon was not a simple rotation of the plane of polarization, but an elliptical polarization resulting from the fact that the medium had become doubly refracting. The phenomenon was observed at its best when the plane of polarization made an angle of 45° with the direction of the wire electrodes, and vanished when the plane was either parallel or perpendicular to this direction. In the case of the resin colophonium a similar effect but of opposite sense was observed, the two substances acting like positive and negative crystals. The slowness with which the phenomenon appeared and disappeared indicated that, in all probability, it was to be referred to a strain induced in the glass resulting either from the continued action of the field, or possibly from temperature changes due to traces of conductivity. Kerr himself was of the opinion that an orientation of the molecules occurred as a result of the action of the electric field, and that it was hindered in a measure by frictional forces.

The Kerr Electro-Optic Effect in Liquids. — If the phenomena discovered by Kerr were merely of mechanical or thermal origin they would have very little theoretical interest. His discovery that similar effects were produced by liquids, in which any considerable strain could not be set up, indicated that the effect was, in all probability, a true electrical one. He examined a large number of fluids and determined the nature of the double refraction, *i.e.* whether positive or negative, by means of a strip of glass, which when compressed or expanded, as the case might be, compensated

the effect produced by the electric field. It was found that positive double refraction, similar to that exhibited by quartz in the absence of an electric field, was exhibited by bisulphide of carbon, bromine, molten sulphur and phosphorus. Negative refraction analogous to that of Iceland spar was found in the case of many oils.

The phenomenon results from the circumstance that light polarized parallel to the field traverses the medium with a velocity different from that of light polarized in a perpendicular direction.

For a quantitative study of the effect it is customary to pass the light between two parallel metal plates immersed in the liquid, the plane of polarization being inclined at an angle of 45° to the direction of the electric field, which is maintained between the plates by means of a static electric machine or other source of high potential. The emergent light is elliptically polarized and the magnitude of the effect can be determined by means of a doubly refracting plate as compensator, placed between the Kerr cell and the analyzing Nicol. If the observed effect is large, this may be a double-image prism, the Nicol being set for equality of intensity of the two images. Or we may use the Babinet compensator as described in the Section on Circular and Elliptical Polarization. A more sensitive method is that of the half shadow polarimeter.

The Kerr Constant. — The magnitude of the effect for a given substance is measured by its Kerr constant. In measuring this the non-homogeneous nature of the field at the ends of the cell must be allowed for, since the field extends beyond the ends of the plates. Chaumont [1] has given the following formula for this

$$l' = l + \frac{a}{\pi}\left[1 + \frac{d}{a}\log_{\text{nat.}}\left(1 + \frac{a}{d}\right)\right]$$

in which a is the distance between plates of length l and thickness d. The applied potential is to be considered as operating over the distance l' instead of l. The difference of path of the two components of the vibration (π parallel and σ perpendicular) is determined by the methods employed in the study of elliptical polarization.

The path-difference

$$d = \frac{l}{(n_\pi - n_\sigma)}$$

in which n_π and n_σ are the refractive indices of the medium in the electric field for the parallel and perpendicular components.

The relative retardation measured in wave-lengths is

$$\Delta = \frac{d}{\lambda} = \frac{l}{\lambda(n_\pi - n_\sigma)}.$$

[1] *Ann. der Phys.*, *5*, 31, 1916.

The Kerr constant $$j=\frac{\Delta}{lF^2}=\frac{n_\pi-n_\sigma}{\lambda F^2}$$

the effect being proportional to the square of the field. F is measured in electrostatic units (1 c.g.s. unit = 300 volts/cm.).

For sodium light the Kerr constants of various substances are given in the following table:

Benzene	0.6	10^{-7}
Carbon bisulphide	3.21	
Water	4.7	
Monobromnaphthalene	9.1	
Chlorbenzene	10.	
Nitro-toluene	123.	
Nitro-benzene	220.	

If the magnitude of the effect is sufficiently great, the cell may stand a potential sufficient to introduce a path-difference of one whole wave-length in which case the field will appear dark as before the application of the electric force. If this is possible, without a breakdown of the liquid we can determine the Kerr constant without going to the trouble of determining the elliptical polarization by using the following formula

$$j=\frac{a^2}{lE^2}$$

in which E is the applied potential necessary to give the dark field, and a the distance between the plates. A favorable condition for this method is a large value of l, the length of the plates.

Ilberg [1] attempted to use this method for determining the absolute value of the Kerr constant for nitro-benzene but found that the potential gradient between the plates was not uniform, which resulted in a variable illumination of the field. This effect is evidently closely related to the circumstance that the slight conductivity shown by most fluids when the field is first put on, disappears after a few moments. The conductivity is probably due to dissociated impurities which are swept to the electrodes; as the conductivity appears again if the polarity of the field is reversed. Ilberg accordingly used an alternating electric field but modified the apparatus so that the method just given for a steady field could be employed. In brief he used two Kerr cells in series, the first transmitting flashes of light of very brief duration and only at the peak voltage of the cell. These flashes passed through the second cell and Nicol, and were visible or invisible as the voltage on the second cell was altered. In this way the potential

[1] Ilberg, *Phys. Zeit.*, *29*, 670, 1928.

necessary for the retardation of one wave-length was determined. The original paper must be consulted for the details of this very ingenious arrangement. Values of the Kerr constant for nitro-benzene for six different regions of the spectrum were determined

λ	Kerr Constant
467	31.1×10^{-6}
491	28.8
527	26.
569	23.
603	21.6
639	20.2

This enormous value of the constant for nitro-benzene makes it especially adapted to methods now to be described.

The Method of Abraham and Lemoine. — The method originated by Abraham and Lemoine in their investigations of the time required for the development of the electro-optic effect is one of great interest and one that can be applied in modified form in many fields of investigation in which the measurement of very minute intervals of time is required.

By this method it has been found possible to measure a time interval of one *two hundred millionth* of a second. A rifle bullet, moving with a velocity of 500 metres per second, would travel

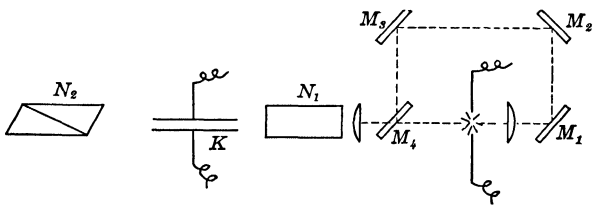


Fig. 419

about $\frac{1}{400}$ of a millimetre, and a light-wave about 1.5 metres, during this interval of time. They illuminated the field of their instrument with the light of a spark produced by the discharge of the fluid condenser, the double refraction of which was being studied. If the light from this spark was made to travel over a distance of 400 cms. before passing between the plates of the condenser all trace of the double refraction disappeared, while a retardation due to the introduction of a path of only 80 cms. reduced the double refraction to one-half of its maximum value. The analyzing Nicol was set for extinction by substituting a lamp for the spark.

The arrangement of the apparatus is shown in Fig. 419. The metal plates of the Kerr cell K are connected to a source of high

potential (static machine or transformer) wires leading from them to the spark gap. The light of the spark can be made to pass either around the rectangle formed by the four mirrors M_1-M_4 and then through the Nicol N_1 set at 45° to the field, the Kerr cell and the Nicol N_2 (set for extinction), or directly through the Nicol N_1 the intervening mirror M_4 being removed in which case the spark can be seen through the crossed Nicols owing to the double refraction of the Kerr cell. The time interval between the flash of the spark and the arrival of the light at the cell can be increased by moving the mirrors M_1 and M_2 to the right, or a single mirror placed at a variable distance can be employed, the spark being mounted a little to one side. As the length of the path is increased the spark, as viewed through the instrument, decreases in brilliancy and finally disappears.

Determination of Velocity of Light with Kerr Cell. — A. Karolus, in 1925, substituted a pair of Kerr cells operated with high frequency potential for the toothed wheel in the Fizeau method for determining the velocity of light. He was able to interrupt the beam 10 million times per second, and obtained the first eclipse of the returning beam reflected from a mirror at a distance of only 15 metres. More recently, working with various frequencies between three and seven million per second and light paths between 250 and 330 metres obtained by multiple reflections in the laboratory, Karolus and O. Mittelstaedt [1] have made 750 measurements of the velocity of light, obtaining as a final result

$$C_0 = \frac{299{,}778 \text{ km.}}{\text{sec } \pm 20}$$

in good agreement with Michelson's latest value of $299{,}796 \pm 4$.

Determination of Duration of Fluorescence with the Kerr Cell. — The first application of the method of Abraham and Lemoine to the study of fluorescence and phosphorescence was made by the author in 1921.[2] Small crystals of barium platinocyanide were placed as close as possible to the spark, and viewed through a nitro-benzene Kerr cell. The greenish image of the crystal could be seen without difficulty, the eye being shielded from the direct light of the spark. If the crystal had remained dark for a ten millionth of a second before fluorescing it would have been invisible.

In this case a search was being made for a "time-lag" in the fluorescence, such as had been observed in the case of mercury

[1] Karolus and Mittelstaedt, *Phys. Zeit.*, *29*, 698, 1928.

[2] R. W. Wood, "Time Interval between Abs. and Emis. in Fluorescence," *Proc Roy. Soc.*, *99*, 362, 1921.

vapor. The method was improved by Gaviola [1] whose fluorometer measures the duration of the fluorescence. This apparatus consists of a pair of Kerr cells simultaneously excited by a high frequency field, corresponding to radiation of $\lambda = 15$ metres (Fig. 420). One cell transmitted light flashes of this frequency, which illuminated the fluorescent substance, the fluorescent light then passing through the second Kerr cell, also between crossed Nicols. The two cells thus performed the function of the rotating discs in the Becquerel phosphoroscope (Chapter on Fluorescence and Phosphorescence). If the time of light transit from first cell to second by way of F is as large as the time of a half oscillation of the electric field, there will be no light transmitted by the Nicol N_2. The Nicol N_4 is set perpendicular to N_3 which is parallel to N_2; K is a double image prism which gives in general two images of unequal intensity derived from the two components of the light which have received a displacement of phase in the Kerr cell.

This is measured by turning N_4 until the two images are of equal intensity. The phase displacement is, within limits, proportional to the time required by the fluorescent substance to reëmit its absorbed light. By substituting the mirror M for the fluorescent substance, we obtain another value for the phase displacement, and the difference between the two is the measure of the duration of the fluorescence. In carrying out an investigation, the "mirror curve" is first determined. The mirror S' is placed close to the inclined glass plate M, and the phase displacement measured (Fig. 420*b*). The mirror M is then substituted for M' and a second determination made. With S' close to M' these two observations will agree, but as S' is moved, say 50 cms., to the right, delaying the time of arrival of the light flashes at cell K_2, one finds a smaller value for the displacement of phase, since

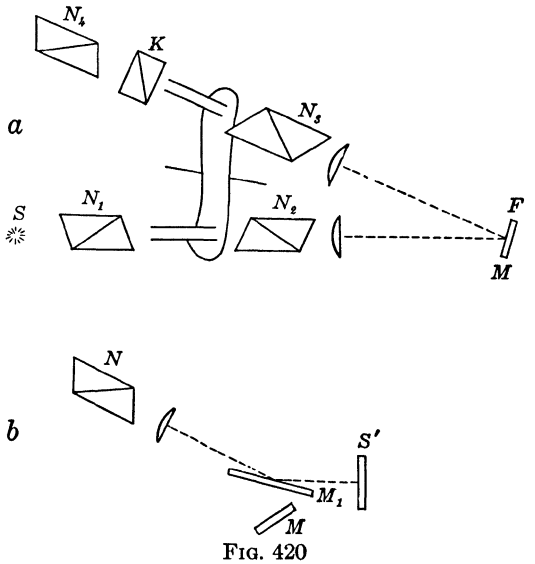


FIG. 420

[1] *Zeit. für Phys.*, *42*, 853, 1927.

the electric field has partially subsided. By further removal of S' continually decreasing values are obtained, until a minimum is reached after which the values increase to a maximum, the flash being delayed until the potential on the cell has come to its next maximum value. This distance, for the frequency employed in Gaviola's experiment ($\lambda = 18.4$ m.) was 9 metres. In measuring the duration of fluorescence a color filter was employed which corresponded in hue to that of the fluorescence. This was important on account of the chromatic dispersion of the Kerr effect for nitrobenzol. The Nicol N_4 was turned from its zero position for equality of the two images formed by the double image prism, first for the fluorescence, then for the mirror S. This gave the phase displacement of the two components in each case. By comparing the readings taken with fluorescent light with those with the mirror moved to a distance, the duration of the fluorescence could be determined. For solutions of organic dyes it was of the order of 2 to 4, 10^{-9}.

Modulation of Light with Kerr Cell: Light-Beats. — As has been pointed out, the periodic interruption of a train of monochromatic light-waves should give rise to two new frequencies $N+n$ and $N-n$, in which n is the frequency of the light and N that of the interruption. The moving fringes of an interferometer, as one mirror is advanced, was shown to be a case of light-beats due to the change of frequency of the light reflected from the moving mirror by the Döppler effect. A much more convincing experiment, however, would be to show that there was an actual change of wave, detectable with a spectroscope, due to a mechanical interruption with a high speed periodic shutter; such a shutter is available in the Kerr cell operated by a very high frequency electrical potential, and a very interesting experiment was described by Rupp [1] in which the green light of a thallium resonance lamp was found to be quite freely transmitted by a cell containing thallium vapor, if the light passed first through a Kerr cell excited at a frequency of $12 \cdot 10^8$ sec.$^{-1}$, while it was absorbed if the frequency was reduced to $5 \cdot 10^8$. This is a case in which an absorption cell serves as a spectroscope for detecting a change of wavelength. The electrical analogy of the experiment can be shown with a Framm frequency metre. If a 60-cycle current is passed through the metre and interrupted two times per second the metre will record frequencies of 58 and 62.

The Stark Effect. — Zeeman's discovery of the effect of a magnetic field on the frequency of emitted radiation caused an intensive search for a similar effect in electric fields. For fifteen years the search was fruitless, and Voigt proved mathematically that no

[1] *Zeit. für Phys.*, *47*, 72, 1928.

effect was to be expected in the case of vibrations of electrons bound by quasi-elastic forces. We now know that the failure of all of these earlier experiments was because none recognized that a strong electric field could not be maintained in a luminous gas, whether excited electrically or by flame, on account of its conductivity.

The importance of lowering the conductivity was recognized first by J. Stark [1] in 1913, who discovered that the hydrogen lines were split up into polarized components, when the luminous canal rays behind a perforated cathode were subjected to a powerful electric field. The canal rays, discovered by Goldstein were mentioned in the first chapter in connection with Stark's discovery of the Döppler effect due to the high velocity with which the luminous positively charged ions shoot through the perforations of the cathode. Stark mounted an auxiliary electrode close to, and immediately behind the perforated cathode and found that it was possible to maintain a field of several hundred thousand volts per

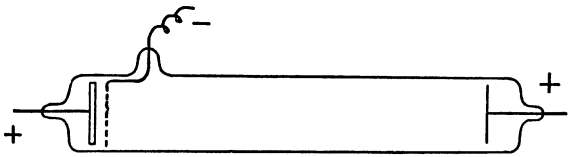


FIG. 421

centimetre between them, in spite of the presence of the luminous ions (Fig. 421). The spectroscope, when pointed from the side at the region between the two electrodes, showed the so-called transverse effect analogous to the transverse Zeeman effect, the line splitting into a number of symmetrically polarized components disposed to the right and left of the original line, the separation increasing directly with the field strength. This is known as the linear Stark effect. In very strong fields we have an additional effect proportional to the square of the field strength, known as the quadratic effect.

With a different disposition of the auxiliary electrode Stark observed also the longitudinal effect parallel to the lines of electric force. His observation on hydrogen may be summarized as follows: Each Balmer line is divided into a number of components, the number increasing with the serial number of the line, the red line H_α showing the smallest number (nine). Observed perpendicular to the field some of the components are polarized parallel, others perpendicular to the field. In a direction parallel to the field, the above-mentioned perpendicular components appear unpolarized, while the parallel components are absent.

[1] Preuss. Akad. Wiss Berlin: *Ber.*, *40*, 932, 1913.

The separation of components is on a much larger scale than in the case of the Zeeman effect, the outer components of the violet line, for example, being 33 A.U. apart for a field of 74,000 volts per cm. while the separation of the outer component of the Zeeman triplet for the same line is only 0.8 A.U., in a field of 45,000 Gauss.

Stark's method, while well adapted to quantitative measurements, can be used only in special cases.

The effect was also discovered independently, and at almost the same time, by Lo Surdo [1] though he apparently did not realize its exact significance until Stark's paper had appeared.

He employed a much simpler apparatus, a simple vacuum tube in which one electrode was an aluminum wire a millimetre or two

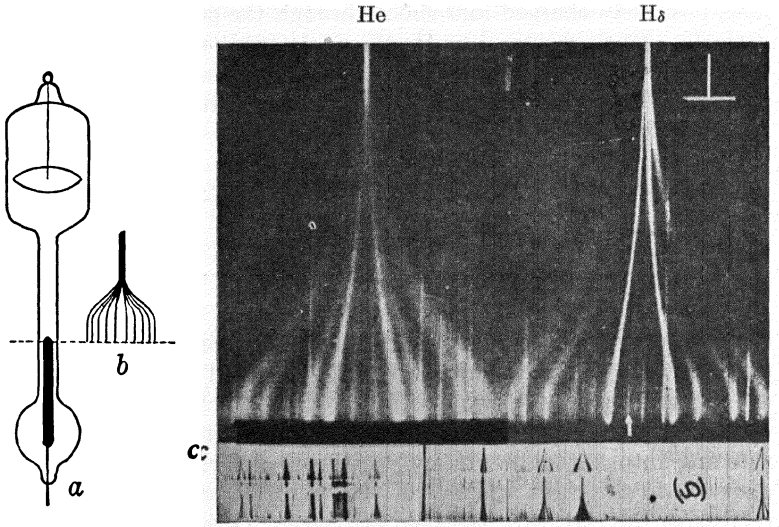


FIG. 422

in diameter fitting snugly into a capillary glass tube (Fig. 422*a*), the effect occurring in the immediate vicinity of the end of the cathode wire, where the potential gradient is very high. A sharply focussed image of the discharge at this point must be formed on the slit of the spectrograph, when the lines will be found to have the appearance shown in Fig. 422*b*. The field falls to practically zero, a short distance above the cathode surface and from this point on the line is single.

This method is applicable also to the study of the effect in the case of the spectra of metals, the metal forming the cathode and vaporizing under the impact of the canal rays.

[1] *Rend. Acad. d. Lincei, 22*, 665. 1913; *Phys. Zeit., 15*, 122, 1914.

Photographs taken by J. S. Foster [1] with an improved form of Lo Surdo tube are reproduced in Fig. 422*c*; contact prints below for || and ⊥ polarized components, and above, enlargements of the Hδ line of hydrogen and the 4144 line of parhelium.

Theory of the Stark Effect. — In one of his earlier papers Stark stated that, to explain the complicated splitting of the Hβ hydrogen line in an electric field on the classical theory, required the combined action of at least fourteen electrons, while the quantum theory accounts for all of the components, their relative intensities and their states of polarization, as a result of the action of an electric field on a single electron only.

In the case of the Zeeman effect the change in the energy of the atom produced by the magnetic field was referred to the angular momentum of the Larmor precession, that is the change involved the kinetic energy chiefly.

In the Stark effect for hydrogen-like atoms it is only the potential energy that is affected, at least in fields of moderate strength. It will be well to illustrate the manner in which the electric field alters the potential energy of the atom by the following analogy. Assume a mass of matter at a certain height above the earth's surface. It has potential energy which will be transformed into kinetic by the fall of the mass. Now assume that before the fall commenced a strong uniform gravitational field was superposed on the earth's field, and assume further that the earth, but not the mass of matter, is prevented from moving under the influence of this field. The potential energy of the mass is increased or decreased by this superimposed field (according to its direction) and more (or less) kinetic energy will appear when the mass falls, the magnitude of the change being proportional to the height of the mass above the earth's surface. If we consider the potential energy as referred to the earth's surface, there will be no change, if the mass is initially at the surface.

Take now the case of the hydrogen atom, consisting of a + nucleus (the earth) with a − electron moving in an elliptical orbit, having potential energy with respect to the nucleus. Instead of considering the electron's charge as distributed around the orbit, we can consider it as concentrated at a point so chosen that its potential energy with respect to the nucleus is the same as its average potential energy for one complete revolution. This point is termed the electric centre which will be referred to as the E.C. In the case of an elliptical orbit, owing to the high velocity at perihelion, it is obvious that the electron spends more time in the

[1] *Proc. Roy Soc.*, *114*, 47, 1927.

aphelion part of its path, and the E.C. will be somewhere between the centre of the ellipse and the aphelion point.

In the case of a circular orbit the E.C. will coincide with the nucleus, while with elliptical orbits it can be proved to be at a point situated halfway between the centre of the ellipse and the unoccupied focus. Its distance from the nucleus is thus $\frac{3}{2}A_n\epsilon$ in which A_n is the semi-major axis of the ellipse and ϵ its eccentricity: $A_n\epsilon$, the so-called linear eccentricity, is the distance of the centre from each focus, which explains the $\frac{3}{2}$. The E.C. corresponds to the centre of gravity of the mass of matter in our gravitational analogy.

Suppose now that the orbit is inclined at an angle β to the direction in which the electric field is to be applied. The E.C. will, in this case, lie at a distance $\frac{3}{2}\ A_n\epsilon \cos \beta$ above or below a plane taken through the nucleus and perpendicular to the direction in which the electric field is to be applied. This will be referred to as its distance "up" or "down field," and corresponds to the height of the mass above the earth's surface. When the electric field is applied the potential energy is changed by an amount proportional to this distance and to the strength of the field. This represents the main part of the energy change except for very strong fields in which, as we shall see, we have a quadratic term.

The change of average potential energy of the atom in its first or excited state, due to the field, will thus result in an alteration of the wave-length of the light emitted when the electron drops to an inner orbit, provided the energy change for the atom in this second state differs from that in its first. To account for the emission of the groups of sharp lines in the Stark effect, we must quantize the orbit, as in the case of the Zeeman effect, the quantization in this case being such as to give the electric centre a fixed value up or down field for each state. The inclination of the orbit and its eccentricity may change, but the relation between them must be such as to keep the E.C. at a fixed distance from the yz plane through the nucleus. This point will be made clearer further along.

For outer orbits (higher energy levels) of greater total quantum number we have a larger number of different orientations and eccentricities available, consequently a larger number of Stark components. We will now follow out what happens when an electric field parallel to the x axis is applied to an atom having a single electron moving about the nucleus in an elliptical orbit. We will suppose, at first, that the field is only of sufficient strength to quantize the orbit, without producing an observable Stark effect. The orbit orients itself in certain definite ways, as with the

Zeeman effect in very weak fields. The electric centre for any given quantum relation is displaced a definite amount "up" or "down" field with respect to the nucleus, *i.e.* it remains at a fixed distance from the *yz* plane. This distance does not alter as the field is increased, but the potential energy does, since the change of potential energy is represented by $\Delta E = eF\bar{x}$ in which e is the electronic charge, F the strength of the electric field and $\bar{x}$ the distance of the electric centre up or down field. If $\bar{x}=0$, as with a circular orbit perpendicular to the field, or in any position for

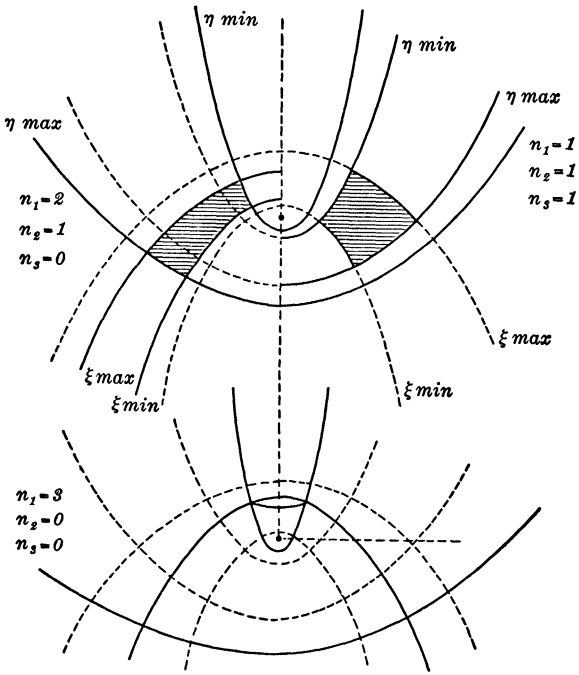


FIG. 423

that matter, or an elliptical orbit with its major axis in the *yz* plane, no change of energy will be produced by the field. We will now proceed with the development of the theory which is not as simple as in the case of the Zeeman effect, for the electron's path in space is much more complicated.

The problem is to determine the orbital electron's motion when subjected to a steady outside force pulling always in the same direction (and not directed perpendicular to the path as in the Zeeman effect).

The solution was worked out independently by Schwartzchild [1]

[1] *Ber. Ber.* 548, 1916; Epstein, *Ann. d. Phys.*, *50*, 489.

and Epstein in 1916, who followed the method employed by Jacobi for the problem of the motion of a particle subjected to the gravitational force of two fixed centres of attraction. Jacobi showed that in this case the most suitable coördinate system to which the motion is to be referred, was a system of confocal ellipsoids and hyperboloids described around the two centres of attraction as foci. By moving one of the attracting centres to infinity and giving it a correspondingly increased attracting force, we have a uniform gravitational field superposed on the central force exerted by the other body on the particle. This case is analogous to that of the electron moving in an orbit under the central force of the nucleus when placed in a uniform electric field, and the system of ellipses and hyperbolas goes over into a system of confocal parabolas described about the nucleus as focus, with their common axis parallel to the field direction as shown in Fig. 423.

The parameters of the dual system of parabolas are ξ^2 and η^2, which, taken together with the angle Ψ (around the axis) form the coördinate system with respect to which the electron's motion is referred. By rotation around the axis (measured by Ψ) the parabolas trace out rotation paraboloids, and it can be shown that the electron path is confined to a space somewhat resembling an anchor-ring traced out by the rotation around x of the curved quadrangle bounded by four intersecting parabolas (shaded in Fig. 423). The shaded quadrangle to the left is for a different quantization of the orbit, later explained.

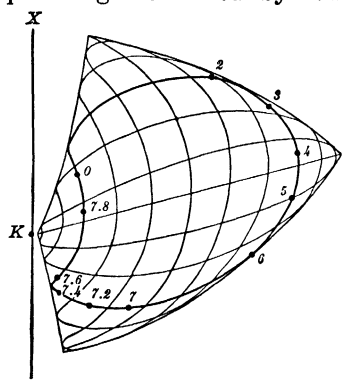


FIG. 424

The mathematical analysis shows that the electron circulates in one of these quadrangles in a path somewhat as indicated in Fig. 424,[1] eventually covering the entire surface, the quadrangle turning about the x axis at the same time, making approximately one revolution for each circuit of the electron in the quadrangle. The electron thus moves around inside the anchor-ring formed by the rotation of the quadrangle with parabolic sides about the x axis. The reader may find it helpful, in the study of what is to follow, to glance over the next section devoted to a study of the electron's motion by stereoscopic models.

Referred to a system of rectangular coördinates, the parabolas $\xi = \text{const.}$; $\eta = \text{const.}$ have equations

[1] Buchwald, *Ann. d. Phys.*, *81*, 965, 1926 (also kinematograph diagrams).

$$\frac{y^2}{\xi^2}+2x=\xi^2; \quad \frac{y^2}{\eta^2}-2x=\eta^2.$$

The position of the electron in the xy plane can be expressed as the intersection of two parabolas of parameters ξ^2 and η^2, and its position in space by a third coördinate, the angle Ψ (defined above) with respect to a fixed meridian plane.

As ξ and η increase, the parabolas become more open, while for $\xi=0$, $\eta=0$ the corresponding parabolas collapse to a line coincident with the x axis. The mathematical investigation has shown that the path of the electron, for a given quantization, is confined to the space bounded by the anchor-ring formed by rotation around x of the quadrangle formed by the paraboloids having maximum and minimum values of ξ and η as shown in Fig. 423, these maximum and minimum values depending upon the way in which the orbit is quantized.

Three quantum numbers are necessary for the specification of the orbit, and the calculation of the energy levels. The impulse coördinates corresponding to the position coördinates ξ, η and Ψ are p_ξ, p_η, and p_Ψ and the quantum numbers n_1, n_2 and n_3 are given by

$$\int p_\xi\, d\xi = n_1 h; \int p_\eta\, d\eta = n_2 h; \int^{2\pi} p_\Psi\, d\Psi = n_3 h.$$

Since p_Ψ is constant we have $2\pi p_\Psi = n_3 h$; $p_\Psi = n_3 \frac{h}{2\pi}$;

n_1 and n_2 can have positive integral values including zero, while n_3 can take only positive integral values, since for zero value of p_Ψ the equations show that $\xi_{\min}=\eta_{\min}=0$, *i.e.* these limiting parabolas shrink to the x axis, and there remain only the intersecting parabolas $\xi_{\max}$ $\eta_{\max}$ and since there is no rotation about x in this case, the electron, traversing eventually every point in the plane included between the parabolas would have to pass through the nucleus (which lies in this plane). On this account this case is excluded in the earlier treatments, but treated by wave-mechanics this case also is allowed. The sum of these three quantum numbers is equal to the total quantum number $n-1$ [1] and we have therefore

$$n_1+n_2+n_3=n-1.$$

If $n_1=n_2=0$, the parabolic quadrangle shrinks to a point or we have two intersecting paraboloids and the electron traverses their (circular) intersection as shown in Fig. 425 (diagram 1). This is the condition for the ground level for hydrogen.

[1] In the earlier treatments the relation was given as $n_1+n_2+n_3=n$ instead of $n-1$ which is the one derived from wave-mechanics.

If $n_1=0$ and $n_2=1$ the path of integration in the first of the above integrals shrinks to zero and $\xi_{min}=\xi_{max}$, *i.e.* we have a single ξ paraboloid only, which intersects, say, the η_{min}, η_{max} paraboloids (for which $n_2=1$) along two circular arcs, and the electron traverses the surface of the ξ paraboloid enclosed between these arcs. Fig. 425 (diagram 2).

In general as we increase n_1 or n_2 one of the corresponding parabolas becomes more open (longer **focus**) while the other shrinks (shorter **focus**) as shown in Fig. 425 (diagrams 4 and 6).

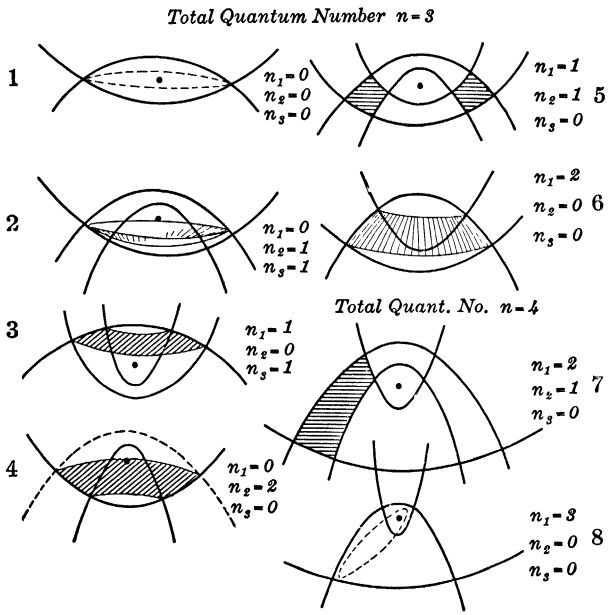


FIG. 425

Diagrams 7 and 8 are for total quantum number $n=4$ and show the further change in the form of the η parabolas as we pass from $n_1=2$ to $n_1=3$. In the latter case $n_2=0$ and the two ξ parabolas have fused, the greatly inclined orbit (dotted ellipse) lying on the surface of the ξ paraboloid, between the circular sections.

We see from this that for $n_1=n_2=0$ the orbit is a circle perpendicular to the field, and that with increasing values of n_1 (or n_2) the orbit tips up from this plane, becoming more nearly parallel to the field, as shown by dotted line in diagram 8 of Fig. 425 and also by the stereoscopic photographs, to be described presently.

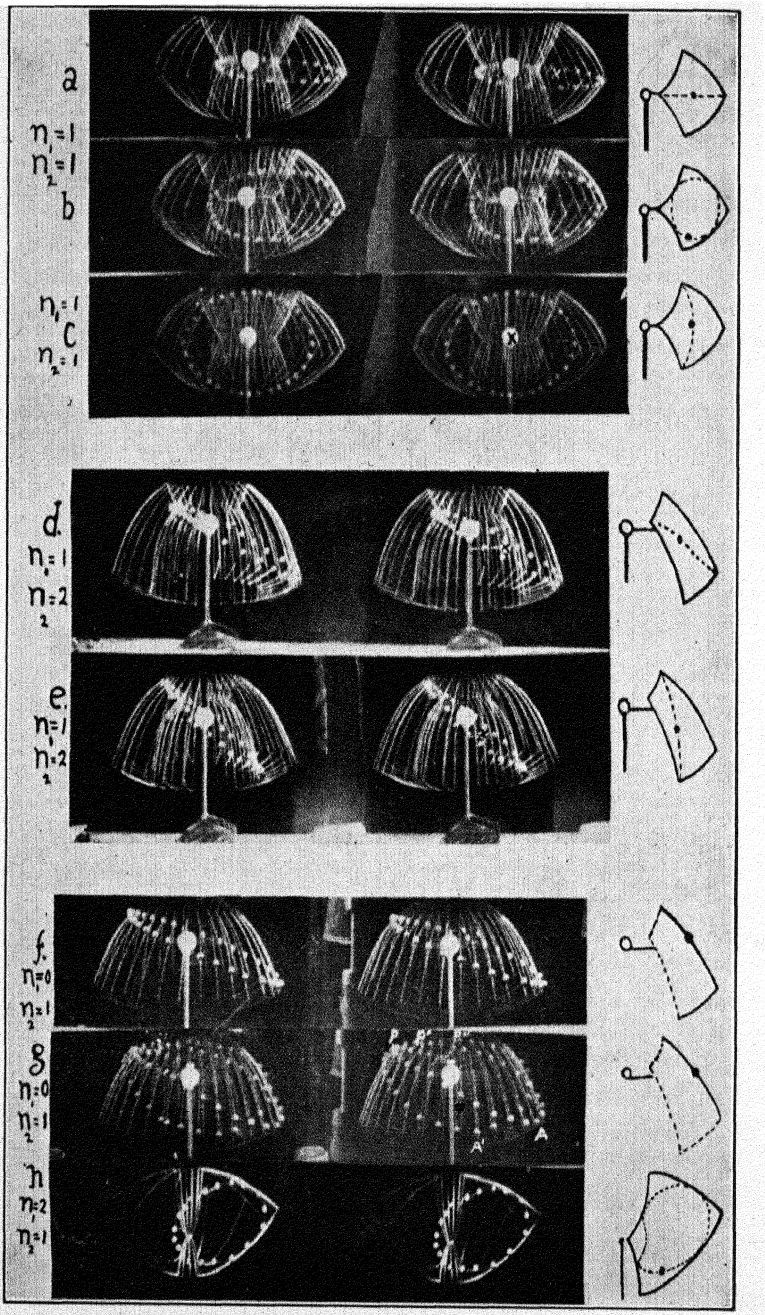


PLATE 15. MOTION OF ELECTRON IN STARK EFFECT.

As the total quantum number increases (higher energy levels) we have more possible values for n_1 and n_2 and a greater number of possible orbits with varying obliquity. It has been shown that the change of energy produced by the field is proportional to $(n_2-n_1)n$, in which n is the total quantum number.

The Electron's Motion in Space in the Stark Effect: Stereoscopic Models.[1] — The movement of the electron within the anchor-ring described in the last section is almost impossible to visualize or illustrate by any two-dimensional diagram, but a method of making stereoscopic three-dimensional models, devised by the author, enables us to get very clear pictures of what takes place when the quantum numbers are assigned in the various possible ways. As has been stated the electron covers the entire surface of the parabolic quadrangle as the latter rotates about a line passing through the nucleus and parallel to the field. It must therefore reach the apices of the four angles of the quadrangle, its motion (referred to the quadrangle) in this case being along a line (straight or curved) joining opposite angles. This line then opens out into a curve which is tangent always to the four sides of the quadrangle, the curve gradually changing its form until it collapses to a line joining the other two angles, as indicated in Fig. 424. During these changes the quadrangle is rotating about the x axis, which makes the actual path of the electron in space extremely complicated for it must eventually make contact with every point on the surface of the anchor-ring. The photographs reproduced on Plate 13 show this path very clearly if viewed with a stereoscope. They were prepared by making a wire frame of the form of one of the parabolic quadrangles arranged to rotate about a vertical axis (Fig. 426). Within this quadrangle an elliptical loop of thinner wire was mounted, on which a small ball of white wax (the electron) could be moved. A wire index and a circle graduated into 32 equal parts at the base made it possible to rotate the frame by equal jumps, the electron being advanced a suitable distance along its wire for each advance of the quadrangle in its circuit around the nucleus.

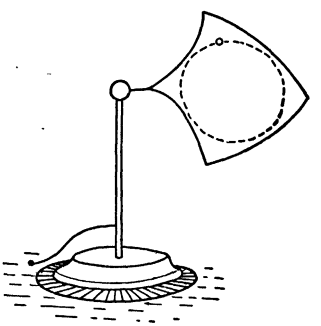


FIG. 426

The apparatus was mounted in a strong light, in front of a black background (the open door of a large dark room) and successive exposures made with a stereoscopic camera, one for each

[1] *Phys. Rev.*, *38*, 346.

advance of the rectangle and electron. The "anchor-ring" is thus built up step by step, the path of the electron within it appearing as a white dotted line within a wire cage when the picture is viewed with a stereoscope, or by the method described on page 407.

Two parabolic wire quadrangles were constructed, representing the two different quantum relations of Fig. 423. We will take first the condition $n_2=n_1=1$ which gives us $n_2-n_1=0$ or no energy change, the electric centre being neither up or down field. The three photographs *a*, *b*, *c*, illustrate this condition, the path along which the electron was moved as the quadrangle was rotated being indicated by the diagram to the right. The two periods were made equal in this case, *i.e.* the number of divisions along the electron's path were equal to the number of steps into which the 360° rotation of the quadrangle was divided. This gives the electron a closed orbit, and corresponds to zero field. In the upper figure, in which the electron was moved to-and-fro once along the horizontal wire during one complete rotation, the orbit is an ellipse lying in the *yz* plane with the nucleus at one focus and the electric centre at *x*, also in the *yz* plane and therefore neither up or down the field, as required for $n_2-n_1=0$.

If we join the other two corners with a straight wire, and repeat the experiment, the electron's path will be an ellipse (as a little calculation shows) with the nucleus at the *centre* an impossible motion. In this case (*c*) the orbit must be a circle, and a little consideration shows us that the line joining the upper and lower angles must be the arc of a circle, whose centre is the nucleus. As the electron is moved along this wire and the frame rotated, the electron's distance from the nucleus is constant, and its orbit is a circle inclined at an angle to the electric field. In both of these limiting cases, the "electric centre," marked with a cross, lies in a horizontal plane passing through the nucleus, that is, it is neither "up field" nor "down field." In the transition stages (*b*) in which the horizontal line opens out into a loop, tangent so the parabolic quadrangle at four points, for each circuit around the quadrangle, the orbit in space passes gradually from the elliptical to the circular form, reaching its maximum inclination to the field when the loop has contracted to the circular arc joining the upper and lower angle. The electric centre coincides with the nucleus in this case. If we make the period of revolution of the quadrangle equal to the period of the motion within the quadrangle the change in the shape and orientation of the orbit will not be progressive, *i.e.* there will be no precession, for if we start by moving the electron along the horizontal wire, and pass gradually to the vertical arc, and then back to the horizontal wire, our orbit is at first an ellipse

with perihelion to the left of the nucleus, changing to the inclined circle, after which the motion retrogrades back to ellipse with perihelion placed as at the start. As a matter of fact the two periods are not equal (except for zero electric field) and the transition from the horizontal to the vertical motion in the quadrangle occurs only as a result of the field.

In taking our photographs we must therefore divide our circle of rotation into, say, 32 equal parts and our electron path on its wire into, say, 28 properly spaced divisions. The orbit, no longer truly elliptical, "precesses," and at the end of a complete cycle in the quadrangle the perihelion of the ellipse will have rotated one-eighth of a complete revolution. In reality the difference of period is much smaller.

For the quantum relation $n_1=2$, $n_2=1$ we have the parabolic quadrangle as shown by Fig. 423 (at left) and by photographs *d* and *e* (Plate 13).

The orbit is now very much inclined to the field, the electric centre being displaced "down field." These photographs also were taken with the two periods equal (condition for zero field) and illustrate the orbital positions for one trip of the electron around or across the quadrangle. Here we notice that the electric centre (marked with a cross) remains at a constant distance "down field" from the nucleus in spite of the change of inclination of the orbit, for as the orbit becomes more inclined, the ellipse becomes less eccentric (shown in the photographs by the position of the nucleus) and the electric centre approaches the nucleus, thus compensating the effect of the increasing inclination of the orbit, the major axis remaining unchanged.

To show the path in space for a complete cycle in the quadrangle, we must make the periods unequal and employ at least two complete revolutions of the quadrangle about the nucleus. The easiest way to accomplish this is to take the case $n_1=0$, $n_2=1$ which, as we have seen, causes the two ξ parabolas to fuse to a single one, the electron travelling about over the surface of this *paraboloid* between the two circular sections cut by the η parabolas. In this case we have only to make our electron travel to-and-fro along the outer parabolic arc of the unsymmetrical quadrangle just mentioned. This was divided into 28 properly spaced divisions, and two complete revolutions given to the quadrangle, with two to-and-fro trips of the electron along the arc. Photograph *g*, Plate 13 was made in this way, while *f* shows the elliptical orbit made under the same conditions, but with the two periods equal. The former gives a very clear idea of the complicated form of the orbit even under this simpler condition. In photograph *g*

the path starts at the back just above the nucleus and curving to the left around the paraboloid passes just below the nucleus at the front, aphelion being at A. In its second circuit, it crosses the first part of the path near P, the perihelion point of the first loop, and reaches aphelion at A'. The end of the path is at perihelion P''. Precession is thus shown in clockwise direction as viewed from above. Photograph h was taken with a wire quadrangle of the form shown in Fig. 424 on which the rate of progress of the electron along its orbit is shown.

In the cases in which the motion is within the anchor-ring, the eccentricity and orientation change continuously and the electric centre moves with simple harmonic motion in an elliptical orbit parallel to the zy plane, with the nucleus at its centre.

Energy Levels and Stark Components. — The change in the energy resulting from the field is shown by the equations of the

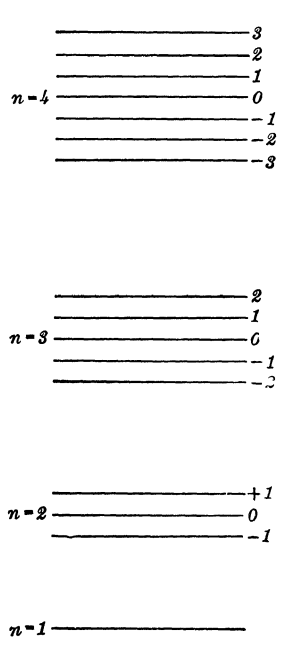


FIG. 427

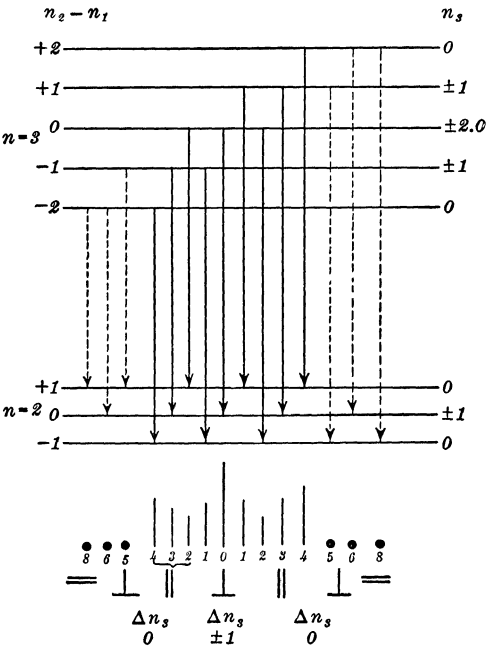


FIG. 428

electron's motion to be proportional to $(n_2 - n_1)n$, that is, to the total quantum number n multiplied by the difference between the new quantum numbers n_2 and n_1.

In the Stark effect the original levels for hydrogen are each broken up into a number of equidistant levels, the separation of

the sublevels being greater as the total quantum number increases as shown in Fig. 427.

We will now take the case of the effect for H_α, the red hydrogen line, for which we require the values of n_2, n_1 and n_3 for the states $n=3$ and $n=2$, the levels between which the transition giving H_α occur.

The possible values for $n=3$ are given in the following table:

$n_1=0$	$n_2=2$	$n_3=0$	two values of each corresponding to opposite directions of rotation (designated by $\pm$)
	$n_2=1$	$n_3=\pm1$	
	$n_2=0$	$n_3=\pm2$	
$n_1=1$	$n_2=1$	$n_3=0$	
	$n_2=0$	$n_3=\pm1$	
$n_1=2$	$n_2=0$	$n_3=0.$	

It is to be noted that the $\pm$ signs of n_3 refer only to the direction of rotation and do not invalidate the relation $n_3+n_2+n_1=(n-1)=2$. For the level $n=2$ we have:

$n_1=0$	$n_2=1$	$n_3=0$
	$n_2=0$	$n_3=\pm1$
$n_1=1$	$n_2=0$	$n_3=0.$

Assigning these values to the energy levels (Fig. 428) we have nine possible orbits for the upper level, and four for the lower. The only allowable transitions (shown by arrows) are those for which $\Delta n_3=0$ or ±1.

The state of polarization is as follows:

	Direction of Rays	
	$\perp$ to Field	$\parallel$ to Field
$\Delta n_3=0$	Components plane-polarized parallel to field	No radiation
$\Delta n_3=\pm1$	Components plane-polarized perpendicular to field	Components = Unpolarized Components (two opposite circular polarizations).

Components corresponding to the dotted line transitions have not been observed experimentally, owing probably to their faintness. The displacements of the components from the field-free line are obtained by subtracting the energy in the final state from that of the initial state, *i.e.* they are proportional to

$$\underset{\text{initial or upper level}}{n'(n_2'-n_1')}-\underset{\text{final or lower level}}{n''(n_2''-n_1'')}.$$

The numbers thus obtained are given below the components in Fig. 428.

This linear effect, with separation of the components proportional to the field-strength depends upon the fixed distance of the electric centre up or down field. The precession of the orbit resulting from the field is accompanied by a change of eccentricity and inclination so balanced as to keep the E.C. in a plane perpendicular to the field.

In the case of all atoms, save those of the hydrogen type, the elliptical orbit revolves about the nucleus in its own plane, this "precession" giving rise to the difference in energy of terms of equal quantum number as was explained in the Chapter on Origin of Spectra. The E.C. in this case is rotating rapidly about the nucleus in the absence of any exterior field and has no preferred position. When the field is applied the E.C. is up field and down field in alternation, and on the average no change of energy is produced. The linear effect consequently is not present. A quadratic effect occurs however in such cases, displacements being proportional to the square of the field-strength.

The Quadratic Stark Effect. — The linear Stark effect, proportional to E, the field-strength, occurs only in the case of hydrogen and hydrogen-like atoms such as ionized helium. It results, as we have seen, from the change in potential energy of the atom due to the action of the field on the dipole formed by the nucleus and electric centre, the magnitude of the change depending upon the distance up field or down field of the E.C. from a plane through the nucleus perpendicular to the field.

In very strong fields, however, the distance of the electric centre from this plane may be altered by the field, and this gives rise to a quadratic term, or to a Stark effect proportional to the square of the field strength, which is superposed on the linear effect in the case of hydrogen-like atoms, or exists alone by itself in the case of the heavier atoms which show no linear effect as we shall see presently. This change in the position of the electric centre with respect to the nucleus may come about in two ways.

Consider first a circular orbit perpendicular to the field. The E.C. normally coincides with the nucleus but the electric field applies a force perpendicular to the plane of the orbit which is thus forced "down field," retaining its shape, however, and the E.C. thus lies below the nucleus. The field now alters the potential energy of the atom with its displaced E.C. as in the linear effect, the change being proportional to the product of the force into the distance of the E.C. from the nucleus. If we double the field, we double this distance and multiply it by the doubled

field-strength, the energy change thus being as the square of the field. It is negative in this case, for the force exerted by the field is opposed to that due to the attraction of the nucleus for the displaced E.C., hence the Stark component lies on the long-wave-length side of the original line. If this effect is superposed on the linear effect the entire Stark pattern, due to the latter, will be shifted towards longer wave-lengths and hence will be unsymmetrical about the position of the unmodified line. This is in agreement with experiment. The field, however, may displace the electric centre in another way. Assume a circular orbit parallel to the field, as in Fig. 429. In this case also the E.C. coincides with the nucleus and there is no linear effect. Suppose the field so directed as to drive an electron down in the direction of the arrow. The effect will be to increase the electron's velocity in the lower part of its orbit and decrease it in the upper, the motion being analogous to that of a weight attached to a string and whirled in a vertical plane in the earth's gravitational field. Since the electron thus spends more time in the upper part of its orbit, the electric centre is elevated above the nucleus and the field now produces a change of the potential energy as in the linear effect. As before the change of energy is represented by the product of E and the distance up or down field of the E.C. and it is proportional to the square of E.

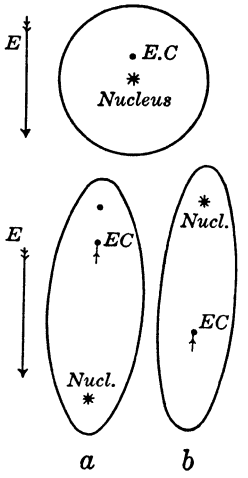


FIG. 429

Applying this action of the field to an elliptical orbit we may take the two orbits oriented as in Fig. 429, *a* and *b*. The linear effect for these atoms will be as follows. Orbit *a* will have its potential energy increased by the field and will yield the short-wave-length component. The reverse will be true for orbit *b*, which will give the long-wave-length component. Considering now the shift of the E.C. by the field, we find that it is elevated in both cases. This would mean that both of the linear effect components would be shifted in the direction of short wave-lengths, for the energy of orbit *a* is increased, while that of orbit *b* is decreased. Herzfeld has shown that the law of attraction between the nucleus and electron (considered as a linear oscillator however) determines which of the two effects will preponderate. They become equal only when the attraction is as the 6th power of the distance. For lower powers the first mentioned effect (which we

may term orbit displacement) preponderates and in the case of a coulomb field (inverse square) is alone of any appreciable influence.

Becker's theoretical treatment of the quadratic effect is concerned chiefly with the second effect (change of orbital velocity) and predicts a shift towards shorter wave-lengths, which is contrary to what is observed in the case of the alkali metals.

Quadratic Effect of the Alkali Metals. — The observation of the quadratic effect is attended with difficulties which do not obtain in the case of the linear effect. The gas must be at very low pressure, otherwise it is impossible to establish the powerful fields necessary, and the light then becomes too faint for satisfactory observation.

In some cases the inverse effect by absorption has been utilized, notably in the case of the alkali metals. This makes it possible to employ a powerful source of light, namely, the emission line corresponding to the absorption line under investigation, and to operate with vapors at very low pressure. Ladenburg was the first to utilize this method in his study of the Stark effect for the *D* lines of sodium vapor at a pressure of 10^{-5} mms., in fields up to 160,000 volts per cm. In this case the source of light was a sodium vacuum tube of special construction giving the *D* lines at great intensity and without trace of reversal.

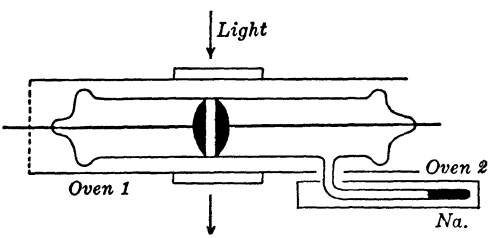


FIG. 430

The light of this lamp was passed between two circular flat electrodes placed very close together and in contact with the tube wall as shown in Fig. 430. The tube was heated in oven 1 and sodium distilled into it from oven 2, at a slightly lower temperature.

The absorption lines appeared as very narrow dark cores in the emission lines of the lamp, a powerful Lummer-Gehrcke interference spectroscope being used. The inverse effect was not observed directly, as the shift was too small to be seen even with the high dispersion employed, but it was recognized by the electrical double refraction which occurred in the immediate vicinity of the absorption lines.

Electrical Double Refraction in Vicinity of Absorption Lines. — The magnetic double refraction occurring near absorption lines and its relation to the Zeeman effect was discussed in the previous

chapter. A similar method is available for the study of the inverse Stark effect as was shown by Ladenburg and Kopfermann.[1] They calculated the expected shift and state of polarization of the lines D_1 and D_2 by employing the principle first stated by Kramers that the assignment of quantum numbers m to displaced energy levels in the electric field are the same as for a magnetic field since the effect in the electric field is quadratic, the $\pm$ levels coinciding.

In the magnetic field the (upper) $2P$ and (lower) $2S$ levels of D_1 split into two, $m = \pm\frac{1}{2}$ which coincide in the case of the electric field.

The $2p$ level of D_2 splits into four $m = \pm\frac{3}{2}$ and $m = \pm\frac{1}{2}$ in the magnetic and into two in the electric field. The energy levels are shown in Fig. 431. The parallel and perpendicular components of D_1 in the electric field have equal energy values and hence equal wave-lengths. Therefore they coincide, giving an unpolarized line shifted towards red. In the case of D_2, however, we have one p component and one s component of the same energy value giving an unpolarized line shifted towards red, and an s component of greater energy giving a polarized line with a lesser shift towards the red.

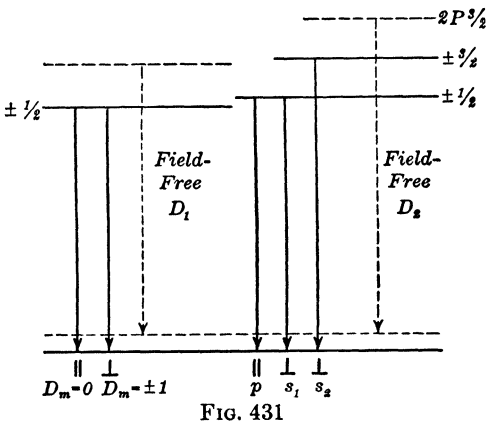


FIG. 431

This was confirmed experimentally by Ladenburg and Kopfermann, who observed anomalous electric double refraction in the vicinity of D_2, due to the difference in the position of the absorption line for parallel and perpendicular polarized light. This makes the velocity of light for a wave-length close to the line vary with the direction of polarization, which is the physical cause of the double refraction, the case being analogous to the one treated in the Chapter on Magneto-Optics (page 717). They also observed the Stark effect with the resonance radiation of sodium vapor, by observing the change in the absorption coefficient of normal unexcited vapor for the light of the resonating vapor in an electric field.

[1] Ladenburg, *Phys. Zeit.*, *22*, 549, 1921; *Zeit. für Phys.*, *28*, 51, 1924; Kopfermann and Ladenburg, *Ann. der Phys.* *73*, 659, 1926.

Grotrian and Ramsauer [1] investigated the inverse effect for potassium, employing two plates 80 cms. long separated by a distance of 1.5 mms., and mounted in a glass tube. Using a concave grating, they were able to photograph the dark absorption lines which appeared in the continuous spectrum of the carbon arc (the light traversing the tube between the metal plates) when potassium vapor at very low pressure was formed in the tube.

Two photographs were taken in coincidence, one with, the other without an electric field between the plates, and shifts towards the red were observed with the field on, for the doublets 4047 — 4044 and 3447.7 — 3446.7, the second and third members of the series. In both cases the shift of the short-wave-length member of the doublet was greater than that of the line of longer wave-length. The former was also more diffuse than the latter, as it was made up of two lines shifted by slightly different amounts (as in the case of D_2 of sodium). With a field of 96,800 volts/cm., they actually separated these two components in the case of the 4044 line. By inserting a Nicol prism in the path of the light, they found that the component of greater shift was alone observed with light polarized parallel to the field (as with D_2 of sodium). Both lines were equally sharp in this case, as the short-wave-length component was not present. With perpendicular polarization, 4047 was shifted the same amount as before, while 4044 showed a lesser shift, since we have here the short-wave-length component of the diffuse line observed with unpolarized light.

The Photo-Electric Effect. — When a metal surface is illuminated by light, X-rays or the γ-rays of radium, electrons are ejected, in number proportional to the intensity of the radiation, and with velocities proportional to its frequency of vibration. This is called the photo-electric effect, and its history begins with the discovery by Hertz [2] in 1887 of the influence of the ultra-violet light from one spark gap upon the discharge of a neighboring gap. If the distance between the electrodes was made just sufficient to prevent the discharge, illumination by ultra-violet light from another spark caused the passage of sparks. Hertz found that it was the action of the light upon the negative electrode only which occasioned the discharge. In 1888 Hallwachs [3] discovered that certain substances, when negatively charged, lost their charge when illuminated with ultra-violet light, and that if originally uncharged, they acquired a positive charge as a result of the illumination. Lenard and J. J. Thomson discovered independently

[1] Grotrian and Ramsauer, *Phys. Zeit.*, *28*, 846, 1927.
[2] *Wied. Ann.*, *31*, 983.
[3] *Wied. Ann.*, *33*, 301.

in 1898 that the action resulted from the expulsion of negative electrons brought about by the absorption of the light, in other words cathode rays leave the surface of the plate.

In the earlier investigations attention was focussed chiefly on the intensity of the emission and its dependence on the state of polarization of the light, and the condition of the metal's surface, but in 1902 P. Lenard directed attention to the importance of studying the velocity with which the electrons left the surface. This he found to be independent of the intensity of the light which illuminated the metal. He showed moreover that the maximum velocity acquired by the electrons depended only on the frequency, the shorter the wave-length the higher the electron velocity.

Two distinct methods of investigation have been used. In the first, the metal surface is illuminated in a vacuum, and the positive charge which it acquires when illuminated is measured with an electrometer. This method gives us the maximum velocity with which the electrons leave the surface, for the potential continues to rise until it becomes great enough to hold back the electrons which have the highest velocity.

In the second method the illuminated plate is placed in opposition to and at a short distance from a second plate, which is connected to earth through a galvanometer, which measures the current resulting from the discharge of the electrons. This method gives us the number of negative electrons which leave the surface of the metal for a given intensity of illumination. The number has been found to be proportional to the intensity of the light, while the velocity, which is between 10^7 and 10^8 cms. per second, is independent of it.

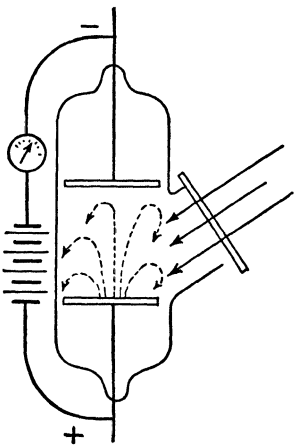


FIG. 432

Or the two plates may be placed in circuit with a galvanometer and a battery so connected as to apply a retarding potential, *i.e.* with its positive pole connected to the illuminated plate (Fig. 432). As the potential of the battery is increased the current decreases, reaching zero value when no electrons are able to reach the negatively charged plate. This gives us a measure of the maximum electron velocity under the conditions of illumination. This condition is shown in Fig. 432 in which the electron paths are indicated by dotted arrows. If V is the voltage required to

just arrest the current, the velocity of the swiftest electrons is given by the formula $\frac{1}{2}mv^2 = eV$ or $v = 595\sqrt{V}$ (in kilometres per sec.). It is usual, however, to express electron velocity in terms of volts, *i.e.* a 4-volt electron has a velocity of $v = 1190$ km./sec. An electron has a velocity of V volts when its velocity v, and hence its energy $\frac{1}{2}mv^2$, is reduced to zero by traversing a potential difference of V.

Electron Velocity and Frequency of Light. — Lenard's discovery that the velocity with which the expelled electron left the metal depended only on the frequency of the light and not at all upon its intensity, led Einstein in 1905 to make the very startling suggestion that the energy of light was not distributed uniformly over the wave-front as had always been assumed, but that it was localized in distinct pockets, or "quanta" which separated as the light spread out, retaining however their energy and individuality. The effect of a collision of one of these quanta with an atom would thus be independent of the distance of the light source. With strong illumination the quanta are close together, with feeble illumination they are widely separated. The experimental facts are represented by Fig. 433, where the ordinates represent currents, the abscissae, the voltage applied between the photosensitive plate and a collecting anode. With a sufficient positive voltage applied on the cathode no current will flow, however great the illumination. At a certain positive voltage, V_p ("stopping potential") currents begin; these increase as the positive voltage is reduced, reaching finally a saturation value at a voltage V_s, which for blue light, with an alkali metal such as potassium, is about one volt removed from the stopping potential, and occurs at zero applied field only if the anode and cathode are of the same material. If the illumination is varied in any ratio, the saturation currents C_1, C_2 will vary in the same ratio. Later work by Compton and Richardson, established that the saturation voltage V_s is a measure of the contact difference of potential between anode and cathode. The effective field V is the sum of the applied field and the contact electromotive force. For electrons emitted at the stopping potential the velocity v is given by the relation

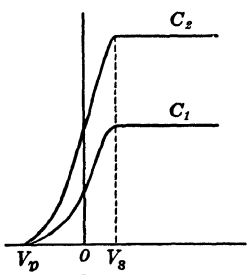


FIG. 433

$$eV = e(V_p + V_s) = \tfrac{1}{2}mv^2$$

where e is the charge on the electron and m its mass. It is this maximum velocity which Lenard found to depend only on the fre-

quency of the light. These characteristics of the photo-electrons became later the foundation of the light quantum hypothesis of Einstein.

Einstein identified the energy of an electron ejected by light of frequency ν with Planck's energy quantum $h\nu$, writing

$$E = h\nu.$$

The value for E found by experiment is, however, smaller than the above value as some of the energy of the light is spent in getting the electron out of the metal. Einstein gave the equation

$$\tfrac{1}{2}mv^2 = h\nu - p$$

in which p is the "work-function" (Abtrennungs arbeit) for an electron belonging to an atom on the surface of the metal.

The quantity p is characteristic of the metal and is closely related to Richardson's work-function in the case of the emission of electrons from hot bodies. It plays only a small part in the photo-electric effect except for visible light. In the case of high frequency, ultra-violet light and X-rays it is so small in comparison to $h\nu$ that it can be neglected. There is now an upper limit of wave-length beyond which no electrons are ejected. Light quanta of this frequency, designated by $h\nu_0$, are only just able to get the electrons out of the metal without giving them any velocity. For this condition $0 = h\nu_0 - p$ or $h\nu_0 = p$ and our equations become $\tfrac{1}{2}mv^2 = h\nu - h\nu_0$. The frequency ν_0 causes the emission of the electron with zero energy. Lower frequencies cause no emission. This shows us that the work function p is to be identified with the energy value of a light quantum at the long wave limit at which the photo-electric effect vanishes.

Further investigations have shown that p may be subdivided into ω and p' the work done respectively in getting the electron out of the atom and out of the metal as a whole, the latter quantity having a close relation with Richardson's work-function in the case of electron emission from hot bodies. It plays only a very small part in the photo-electric effect, except in the case of light of low frequency. Of greater interest is the work ω necessary to remove the electron from the atom, for it will depend upon the electron's location in the atomic structure, and a great deal has been learned about atomic structure by studying the photo-electric effect with X-rays and the γ-rays of radium, in which the energy values of the quanta are enormously greater than in the case of light.

Ladenburg's early measurements with ultra-violet light showed that the relation between electron energy and the wave-length of the light was linear, as was pointed out by Joffé in 1909.

Plotting maximum electron energy as ordinates and light frequency as abscissae, we obtain a straight line, the slope being the same for all metals, the point at which the line cuts the axis of abscissae, however, depending on the value of p for the metal. In these experiments we have electrons of all velocities below the maximum which originate in the layers below the surface of the metal.

The proof of Einstein's law connecting the energy of the photo-electron with that of the light quantum was given by the experiments of Richardson and K. T. Compton who showed that the electron energy increased proportionately with the frequency of the radiation, and that the constant of proportionality was equal to the value of Planck's constant h, which can be determined in this way from the slope of the line above referred to. Millikan made a very careful series of observations on the alkali metals, which respond to a wider range of frequencies than the metals previously employed, and hence permit of a more accurate determination of the slope of the line, which should be exactly equal to h/e. From the known value of e, the electron's charge and the measured slope of the line a value was found for

$$h = 6.56 \times 10^{-27}$$

in close agreement with the value found by Planck from the laws of radiation.

The relation between velocity and wave-length has been found by Wien and Stark to extend into the range of X-ray frequencies.

Relation between Electron Current and Wave-Length. — The number of electrons emitted by a metal surface increases, as the wave-length of the light decreases. In determining this, due account must be taken of the change of intensity as we pass down the spectrum. If E_1 is the saturation current measured by the galvanometer, which is the measure of the number of electrons emitted, and E_2 the energy of the radiation of wave-length λ, we must plot the quantity E_1/E_2 against λ in order to obtain a curve showing the relation between the number of electrons thrown off and the wave-length of the light. Ladenburg found that the E_1 curve had a maximum at $\lambda = 2120$, falling rapidly on each side of this point, E_2 increased steadily with increase of wave-length, while E_1/E_2 increased very rapidly with decrease of wave-length, the curve resembling a dispersion curve in the ultra-violet region.

This is not at all what we should expect if the total energy of the absorbed light is spent in the liberation of electrons. For light

of frequency ν the quantum of energy $h\nu$ is given to the electron. If all absorbed energy is thus spent in electron liberation we should expect the number of electrons ejected for absorbed energy Q to be

$$N=\frac{Q}{h\nu}$$

or the number ejected per absorbed calorie to increase with the wave-length of the radiation, since increase of λ is the equivalent of decrease of ν. It would be natural to expect, for a given value of Q, the ejection of few electrons of high velocity for the short λ region, as here we have light quanta of high energy value, but fewer in number, than in the long λ region, where the light quanta are more numerous but of smaller energy value. Here we should expect an emission of a large number of electrons of small velocity. But we know that by a further increase of λ we finally reach a point λ_g at which the emission ceases entirely, the energy of the quanta being insufficient to liberate the electrons. There must then be some point near λ_g at which the rate of emission begins to fall off, say λ_q in Fig. 434*a*.

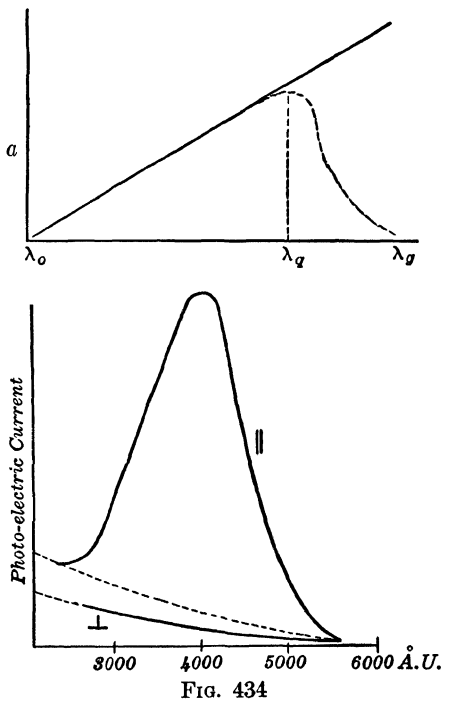


FIG. 434

This is contrary to experimental results, the number of ejected electrons being at most only a few per cent of the quantum equivalent.

Ladenburg's results, and all subsequent observations have shown that the number of electrons ejected per calorie of absorbed energy increases with decreasing wave-length. It is probable that we are observing the dotted portion of the curve in Fig. 434*b*, which represents the falling off in number as we approach the long-wave-length limit where emission ceases. Properly conducted experiments with X- and γ-rays might give us the straight portion between λ_0 and λ_q.

Photo-Electric Effects by Polarized Light. — Normal and Selective: The long series of investigations by Elster and Geitel, which resulted in the observation that some metals, when amalgamated with mercury, responded to visible light, and the invention of the photo-electric cell, culminated in the important discovery that a liquid alloy of sodium and potassium, when illuminated obliquely by polarized light, showed great changes in the photo-electric emission as the plane of polarization was rotated.

With light polarized so that the electric vector is in the plane of incidence the current may exceed by as much as fifty times that with the electric vector parallel to the surface. An extended study of this effect by Pohl and Pringsheim, who carried their measurements through a wide spectral region, showed that this "vectorial effect" is localized in a narrow region of the spectrum. While the "normal" effect, as they termed it, increased uniformly toward high frequencies, the curve for the "selective effect" showed as a hump, similar in appearance to the spectral maxima occurring in the curves obtained by Elster and Geitel with sodium films converted into the hydride. Their results with sodium-potassium alloys for light incident at an angle of 60° are shown graphically in Fig. 434. There is no such irregularity in the bulk absorptive power of the alkali metals or their alloys. If the photo-electric currents were simply proportional to the light absorbed, the currents would not vary by more than two or three times from one plane of polarization of the light to the other.

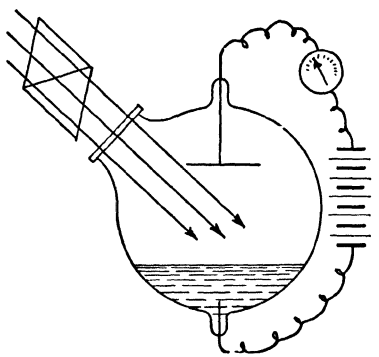


FIG. 435

A clew to the origin of this "polarization selectivity" has been found by Ives,[1] from a study of the optical conditions at the surface of a metal. Upon calculating the energy density in the standing wave or interference pattern above a metal mirror, using the ordinary optical constants, it is found that, at the plane of intersection with the mirror the energy densities are quite differently related to each other for the various conditions of incidence and polarization than they are for the bulk absorption. If, as is probable, the photo-electric action is localized in an extremely thin layer it is the energy at the surface which is in question. Thin films of alkali metals on such underlying metals

[1] *Phys. Rev.*, *38*, 1209, 1931.

as platinum and silver show strong polarization selectivity which can be correlated in a very satisfactory manner with values of the energy density computed from the optical constants of the underlying metal.

Pohl has described a very striking experiment for demonstrating the dependence of the electron emission on the plane of polarization of the light. A cell provided with two electrodes, one a fluid alloy of sodium and potassium is illuminated with polarized light of wave-length 3800–4200. The galvanometer shows a difference in the photo-current in a ratio of 1 : 100 as the polarizing prism is rotated (Fig. 435). The wave-lengths for the maximum values of the selective effect for a number of metals are given in the following table:

	λ
Lithium	2800
Sodium	3400
Potassium	4350
Rubidium	4800
Strontium	3600
Barium	3800

The position of the maximum appears to depend on surface conditions. Pohl and Pringsheim found that a potassium surface brought into the condition of a black colloidal modification by the action of oxygen had its selective maximum at $\lambda = 4050$.

Directional Distribution. — The question as to whether there is any tendency of the photo-electrons to leave the surface in a direction coinciding with that of the electric vector of the light (as has been observed in the case of X-rays) is of considerable importance in connection with the theory of the effect.

A very elaborate apparatus was employed by Ives [1] and his collaborators and a very thorough investigation, both theoretical and experimental, carried out. They used a small spherical cathode surrounded by a spherical anode, the resulting field being a central force of the inverse square type, giving elliptic or hyperbolic electron paths according as the outer sphere was negative or positive with respect to the inner one (retarding or accelerating field respectively).

They found a symmetrical angular distribution of the photo-electrons around the normal to the surface, with no tendency to follow the electric vector, which Fry explained as probably due to the fact that the electric vector rotates within the metal, that is, we have elliptical polarization as in metallic reflection. With

[1] T. C. Fry and H. E. Ives, *Phys. Rev.*, *32*, 44, 1924; Ives, Olpin and Johnsrud *Phys. Rev.*, *32*, 57.

polarized light at normal incidence there was also no tendency of the electrons to emerge in the direction of the electric vector. The exact mechanism by which the electrons are released from the metals is by no means clear, for the experiments of Ives have shown clearly that it is not simply a forced oscillation of increasing amplitude terminating in expulsion.

In the case of photo-ionization of gases the conditions are simpler, and Lawrence and Chaffee [1] have found, in the case of potassium vapor, that twice as many electrons are emitted in the direction of the electric vector, as at right angles to it, these latter being in all probability scattered electrons.

Photo-Electric Effects of X-Rays. — In the case of the release of electrons by X-rays, frequency of which is much greater than that of light, the process is much clearer. By forming a magnetic spectrum of the electrons emitted from metal films under irradiation by monochromatic X-rays, de Broglie [2] showed that a "line spectrum" is produced. The lines correspond to electrons released from the various K, L, etc., deep energy levels of the atom in striking agreement with the Bohr theory of atomic structure.

The subsequent and more precise experiments of Robinson and his collaborators [3] have shown that the lines are sharply defined on the high velocity side, and hazy on the opposite side due to energy losses by collisions which occur even in very thin films.

The striking difference between the optical and X-ray effect lies in the circumstance that in the former, where free electrons are concerned, we have electrons emitted with velocities depending on the energy of the light quanta (*i.e.* on the frequency) whereas in the latter we have a group of discreet velocities corresponding to the binding energies of the inner electrons of the atoms. $h\nu = E_1 + \frac{1}{2}mv^2$, E_1 having different values for levels K, L, etc.

Ionization of Gases and Vapors. — We have seen in the Chapter on the Origin of Spectra that there is a band of continuous absorption beginning at the head of the principal series of the alkali metals and extending down into the region of higher frequencies. This band was first observed by the author, and we now know that it results from the impact of light quanta of greater energy value than that necessary to give a line at the head of the series. We should therefore expect ionization, or the complete expulsion of the electron from the atom.

The first clear demonstration of this effect was given by Foote,

[1] *Phys. Rev.*, *37*, 1233, 1931.
[2] *Jour. d. Phys. et Rad.*, *2*, 265, 1921.
[3] *Proc. Roy. Soc.*, *104*, 455; *113*, 282; *128*, 92.

Mohler and Chenault,[1] using caesium vapor, detecting the photo-ionization by the variation of space charge by the positive ions produced. Their method depended on the discovery by Kingdon and by Hertz that the current produced by applying a potential between a hot filament and a surrounding conductor is limited by the space charge (electrons from the filament). If a positive ion is produced in this region, it may, owing to its low velocity, cause the liberation of 10^4 to 10^6 electrons from the atoms of a vapor in which the cylinder and filament are immersed. This causes an increase in the current, and they obtained the curve shown in Fig. 436, the small peaks being due to the absorption of the lines of the principal series, by which excited atoms are produced, which, by collision with neutral atoms (according to the theory of J. Franck), form an ionized molecule (*i.e.* Cs_2^+ + electron), the ionization potential of the molecule being lower than that of the atom.

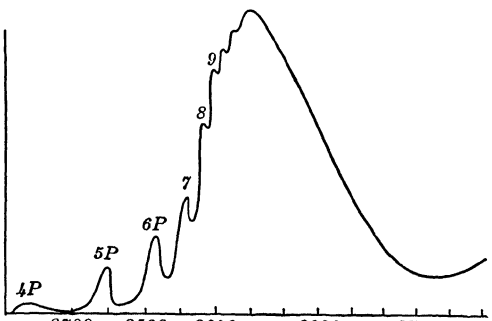


FIG. 436

Photo-Conductivity: The Internal Photo-Electric Effect. — The phenomenon of the production of electrons by light is exhibited not only by free metal surfaces and metallic vapors, where the effect is *external*, but in many cases where the electrons remain inside the substance, and produce variations of conductivity and electromotive force. The earliest and best known case of an *internal* photo-electric effect is the *selenium cell.* In this a bridge of the metallic form of selenium is formed between gold or other metallic wires, and the initially high resistance is greatly reduced by the incidence of light. The response is not directly proportional to the intensity of illumination, and is not instantaneous.

It is only very recently that the physical processes involved in this change of conductivity have been discovered by the brilliant investigations of the whole field of photo-conduction by Gudden and Pohl, who found that in addition to a primary photo-electric liberation of electrons within the medium, there occurred all sorts of secondary effects, which in the case of selenium were so great as to completely mask the primary effect.

In certain insulating crystals they were able to produce the

[1] *Phys. Rev.*, *27*, 37, 1926.

primary effect, unaccompanied by secondary effects, which could be made to appear by changing the conditions. For example, they found that a crystal of zinc-blende, placed between two electrodes which were in circuit with a battery and galvanometer, gave, when illuminated, a current proportional to the intensity of the illumination, if the illumination was not too intense or of too long duration, and that it started and stopped without lag, two characteristics of a pure photo-electric effect. These relations do not hold if the crystal has been under illumination for a certain time, due to changes in its conductivity resulting from the flow of the primary current. They showed that such a change occurred by determining the voltage-current curve for a crystal of zinc-blende in the dark, obtaining two quite different curves for a fresh specimen and for one in which the current had been flowing for some time. They inferred from this that the natural conductivity (*i.e.* in the dark) was electrolytic, involving a transfer of matter through the crystal.

The primary current, due to the photo-electrons, may have superposed on it a secondary current resulting from the motion of the positive ions, and Gudden and Pohl found that in cases in which these were immobilized in the medium, they could be set free by the action of infra-red radiation, the conductivity being nearly double when infra-red acted simultaneous with the ultra-violet, though producing no change when acting by itself. If, however, they gave a preliminary ultra-violet excitation a subsequent infra-red illumination gave increased conductivity. They obtained very interesting results by modifying the character of pure crystals, such as rock salt, by the introduction of small amounts of impurities, or by coloring them by X-rays. A very complete account of their experiments will be found in the recent treatise on *Photoelectric Phenomena* by Hughes and Du Bridge, in which the subject is most fully treated in all of its aspects.

Photo-Voltaic Effects. — Various voltaic cells, operating only when illuminated, have been constructed and studied since A. C. Becquerel's discovery in 1839 that a pair of platinum or silver electrodes coated with silver chloride and immersed in dilute sulphuric acid, produced a current when a strong light was thrown on the cell. The tendency at the present time is to regard the phenomenon as a photo-chemical one, and not a pure photo-electric effect.

The most interesting case, and one which is more nearly photo-electric in character is Lange's cuprous-oxide cell. Grondahl [1] discovered that the action of light had some effect on his rectifier,

[1] *Jour. Am. Inst. Elec. Eng.*, *46*, 215, 1927.

which consisted of a copper plate coated with a film of Cu_2O, and pointed out that it had possibilities as a photo-cell. Lange [1] developed the idea, and by coating a copper plate with a layer of Cu_2O on which a semi-transparent layer of metal was deposited, obtained a continuous current on joining the outer layers with a wire, if the plate was illuminated with sunlight, the flow of electrons across the boundary setting up an E.M.F. of sufficient magnitude to operate a relay, or even drive a small motor. Such cells are coming into rapid use for light detection and measurement.

[1] *Phys. Zeit.*, *31*, 139, 964, 1930.

CHAPTER XXIII

THERMAL RADIATION

In this chapter we shall discuss the type of radiation which practically all liquid and solid substances emit in virtue of their temperature. Most of our sources of light, natural and artificial operate in this way. The incandescent mass of the sun, the glowing particles of carbon in the candle and gas flame and the glowing wire of the electric light pour out energy of this type. Examination with a spectroscope shows us that the spectrum is continuous, the red appearing at a temperature of about 550° C, followed by the other colors in succession, the position of the point of maximum intensity drifting towards the region of shorter wave-lengths with rising temperature. The quality of this type of radiation, by which we mean the distribution of energy in the spectrum, is, under certain conditions which will be specified presently, independent of the nature of the substance, and dependent only on the absolute temperature of the radiating mass. On this account the laws governing it can be regarded as more fundamental than those which relate to the emission of light by gases, excited electrically or by other means, in which the character of the radiation is characteristic of the gas. Of the actual physical processes which are taking place in the temperature radiator we speak with more hesitation than in the case of radiating gases where we are dealing with isolated atoms and molecules.

Gases and vapors do not in general radiate as a result of high temperature and for many years it was supposed that there was no exception to this rule. We know now, however, that in certain cases radiation of restricted type, *i.e.* confined to certain definite regions of the spectrum, occurs in the case of some gases as a result of thermal excitation. These cases will be dealt with later on.

This difference in the behavior of liquids and gases is of fundamental importance and it appears probable that much information of value would be gained if we could study the radiation of some substance having a critical temperature above 700° C. Mercury would appear to be a suitable substance, and if some method could be devised of preventing the explosion of a quartz capillary tube containing the metal, it is probable that we could study the transition which occurred when the white-hot metal reached its

critical temperature, which is supposed to be in the neighborhood of 900°.

R. J. Strutt, now Lord Rayleigh, made experiments of this nature many years ago, without obtaining satisfactory results. It is of no use to increase the wall thickness of the tube, since the inner layers give way or tear before the strain reaches the outer layer. A possible solution might be to employ a nest of concentric tubes, each containing mercury in amount sufficient to generate pressure a little short of the exploding pressure. Pressure would thus be applied to the outside wall of each tube, the principle being that employed in wire-wound guns.

It seems probable that the powerful thermal radiation of conductors of electricity, metals, graphite, etc., results from the presence of free electrons. These, when subjected to acceleration by the thermic agitation of the atoms, may be regarded as the sources of the radiation. If free electrons are absent, as in the case of perfect insulators, fused quartz for example, there is almost no radiation, even at very high temperatures. Bound electrons may however be sources of radiation. They give rise to absorption bands, and fused quartz containing a trace of neodymium radiates strongly when heated, giving a spectrum of widely separated and brilliantly colored bands, corresponding in position to the absorption bands of the rare earth. Even in the case of metals we have to consider bound electrons as well as those that are free. They are responsible for the colors of certain metals, such as gold and copper, while silver has a region of great transparency in the ultra-violet at wave-length 3000.

The so-called permanent gases do not radiate even at the highest laboratory temperatures. This is due to the fact that sufficient energy to raise the electrons from the ground level to higher energy states cannot be communicated to the atoms. This is not the case, however, for certain vapors, iodine and selenium and many metallic vapors for example, as we shall see presently.

The Relation between Emission and Absorption. — We will now investigate the very intimate relation which exists between the emission of heat or light-waves by a substance, and its power of absorbing the same waves.

That some relation existed between the emissive and absorbing power in the case of radiant heat was indicated by the experiments of Leslie, Melloni, Provostaye and Desains, and others. Their measurements were, however, made for the most part with apparatus of insufficient sensitiveness, the spectrum regions being only roughly determined by means of absorbing screens. It was determined, however, without question, that bodies which pos-

sessed a strong emissivity acted also as powerful absorbers of the radiant heat which they emitted, and the approximate equality of the emitting and absorbing powers was recognized.

Ritchie's Experiment. — This relation was shown by a very simple and ingenious experiment devised by Ritchie.[1] Two air-tight metal chambers were connected by a glass tube containing a drop of fluid, the whole forming an air thermometer. Between them a third metal chamber of the same size was mounted, which could be heated by filling it with boiling water. One surface of this heat radiator was covered with lampblack, the other with the substance under investigation, for example powdered cinnabar. (See Fig. 437.) The surface of the air thermometer which faced the radiating lampblack surface was coated with cinnabar, while the surface which faced the cinnabar radiator was coated with lampblack. With the apparatus arranged in this way no movement of the fluid drop occurred when boiling water was poured into the radiator, which established the fact that the emitting and absorbing powers were equal. Let us assume that the lampblack radiates powerfully, the cinnabar feebly. The powerful radiations coming from the former are but slightly absorbed by the latter, while the feeble radiations from the latter are strongly absorbed by the former, the heating of the two chambers of the thermometer being equal. Calling E the amount of heat emitted by the cinnabar and A its absorbing power, e and 1 the corresponding expressions for the lampblack, the lampblack surface emits an amount of radiant heat e, of which the cinnabar surface absorbs the amount eA. The cinnabar surface emits an amount E, which the lampblack completely absorbs (since its absorbing power $=1$). The equality of temperature indicated by the thermometer shows us that $eA=E$ or $E/e=A$. Now E/e is the ratio of the amount of heat emitted by cinnabar to the amount emitted by lampblack at the same temperature. This ratio we will call the emissivity of the cinnabar, and our equation shows us that it is equal to the absorbing power.

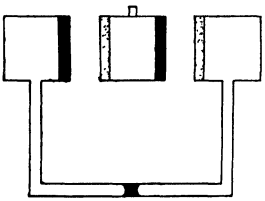


FIG. 437

Kirchhoff's Law. — This relation was reduced to a more definite form in 1859 by Kirchhoff (and independently by Balfour Stewart), who showed that it must be true for each wave-length in the emitted spectrum, and formulated the law which has since gone by his name. "*At a given temperature the ratio between the emissive and absorptive power for a given wave-length is the same for all bodies.*"

[1] *Pogg. Ann.*, *28*, 378, 1833.

The theoretical considerations from which Kirchhoff's law is deduced will be discussed later on in the chapter. It may be remarked that we frequently meet with the statement that the absorption of light by flames which contain the vapors of metals is a necessary consequence of this law. This is by no means the case, for Kirchhoff's law is to be applied only to radiation which results from temperature. In the case of the emission by flames, the phenomena are in part connected with chemical changes which are taking place.

Kirchhoff's law states that in the case of radiation which results solely from temperature, $E/A = \text{const.}$ for all bodies at the same temperature. The value of the constant is a function both of the temperature and wave-length, and is equal to the emissive power of a perfectly absorbing body, that is, a body which at the temperature in question completely absorbs, without reflection, all radiation falling on it, no matter what its wave-length. This amounts to saying that at a given temperature no substance can emit more light of a given wave-length than a perfectly black body. Paschen has compared the intensity of the light of the two D lines in the sodium flame with the total intensity of a region, completely enclosing the D lines, in the continuous spectrum of a black substance heated in the same flame. The total intensity of the D radiation was more than twice as great as that of the region of the continuous spectrum which enclosed them, from which the inference can be drawn that something other than temperature is concerned with the emission of light by the sodium flame. The same thing was found by Kayser and Paschen in the case of the ultra-violet bands of the arc, which were much brighter than a corresponding region of the spectrum of the positive crater, notwithstanding the fact that the temperature of the latter is higher than that of the arc proper.

Cotton has called attention to the fact that there are two distinct relations which are almost invariably confused: a qualitative rule, which connects the absorption and emission for a given substance, and a quantitative rule, which establishes relations between different bodies. From the former we can only draw the conclusion that if a body emits certain radiations it absorbs them when they come from without.

Absorption by Flames. — The absorption by flames and the reversal of spectral lines are special cases of the qualitative rule. These cases have been studied by Gouy, who sought to determine by experiment whether flames were transparent to the radiations which they emitted. His method consisted in comparing photometrically the light of different thicknesses of radiating gas. If

no absorption occurred, doubling the thickness should double the intensity of the illumination. He found, in the case of every line examined, that after the line had attained a certain brightness absorption manifested itself. If E and A be the emissive and absorptive powers of unit thickness of the flame and we increase the thickness, we shall find that the emissive power (*i.e.* the intensity sent out by the thick layer) approaches E/A as a limit.

This can be readily seen by calculating the amount of light sent out in the direction B by a flame of thickness 6 (Fig. 438). Assume unit thickness to emit light of intensity 100 ($E=100$) and to have the power of absorbing one-half of the light of the same wave-length, which traverses it. The intensity of the light from the element 1 is reduced to 50 by the absorption of element 2, and on emergence from element 6 has an intensity of only 3.37. The total amount of light emerging from element 6 is obviously the sum of the numbers in the different elements, which in this case is 197.62. If the number of elements were increased indefinitely this sum would be 200, the limiting intensity which is equal to E/A, as defined above.

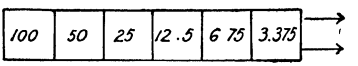
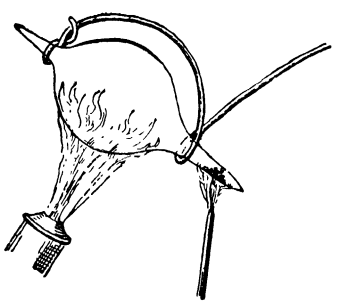


FIG. 438

Temperature Radiation of Gases. — There was at one time a great deal of discussion as to whether gases radiated as a result of temperature alone, and an immense amount of work was done by Pringsheim and others in endeavors to obtain a luminous emission from gases as a result of high temperature alone. Efforts in this direction were almost without exception, in vain, and Pringsheim came to the conclusion that, at least for temperatures which could be commanded in the laboratory, gases remained dark.


FIG. 439

This is due to the fact that these gases have no absorption in the visible spectrum. Absorbing gases such as iodine show a temperature radiation as was shown by Salet and Evershed. Iodine vapor when heated to a temperature of only six or seven hundred degrees gives off a reddish-orange light. The experiment is easily performed by arranging a small spiral of platinum wire, which can be heated by a current, in a test-tube in which a little iodine is vaporized by means of a Bunsen burner. An orange-colored flame is seen to rise from the hot wire. A still better method is to enclose a few crystals

of iodine in an exhausted bulb of fused quartz. The body of the bulb is heated to a high temperature by the flame of a blast lamp, and the iodine in the narrow neck then vaporized by the application of a small flame as shown in Fig. 439. The quartz bulb emits little or no light until the vapor enters it, when it immediately glows with an orange-red light.

Sodium vapor exhibits a similar phenomenon when heated in a long thin-walled steel tube, which has been previously exhausted. The ends of the tube are closed by glass plates. The spectrum is a band spectrum which is also seen in absorption and is now known to be due to the molecule Na_2. It was first photographed by Konen.[1] The author has recently observed the thermic emission of a very beautiful blue light by selenium vapor in a quartz tube. One has only to drop a small speck of selenium into a bit of quartz tubing closed at the bottom, and heat it with an oxy-coal-gas blast lamp. The failure of most gases to emit visible radiations we now know results from the circumstance that the electronic transitions require a greater amount of energy than can be supplied by thermic agitation.

There are other cases in which we have a true temperature emission of a gas. Carbon dioxide, when heated, emits an infrared radiation, the spectrum showing a very sharp band at $\lambda = 4.3\ \mu$. This radiation can be obtained from the column of hot air rising from a Bunsen burner. Paschen [2] found that a layer of the gas 7 cms. thick emitted and absorbed as strongly as a layer 33 cms. thick. This indicated that the radiation from a 7-cm. layer could be regarded as the equivalent of the radiation from a layer of infinite thickness, or in other words, the radiation of a black substance at the same temperature. By heating the gas in a tube, and measuring the intensity of the emitted radiation with a spectrobolometer, Paschen found that for all temperatures between 150° and 500° the intensity of the radiation of wave-length 5.12 μ was only a little below that of a black body at the same temperature. As a black body he used a smoked strip of platinum heated by an electric current. The proof of the law follows from the fact that for a layer from which the radiation is the equivalent of that from an infinitely thick one, we have the relation $E/A = e$, the emissivity of a black body, as we can at once see by comparing the equation $eA = E$, given at the beginning of the chapter with the relation previously deduced, namely, that as the thickness of an emitting absorbing layer increases, the intensity of the radiation leaving it approaches the value E/A as a limit.

[1] *Wied. Ann.*, *65*, 256.
[2] *Wied. Ann.*, *51*, 1, 1894.

Temperature Excitation of Line Spectra. — More recently bright line spectra of many elements have been obtained by A. S. King in his electric furnace. There seems now to be no doubt about the purely thermic excitation in this case, though at one time it was argued by some that the possibility of electrical excitation has not been excluded. Lines resulting from transitions from the upper levels of least energy appear first and as the temperature rises other lines corresponding to higher energy values make their appearance. More caution must be displayed in interpreting results obtained with flames, in which chemical processes are taking place. F. H. Müller, however, obtained results with a Bunsen flame at 1550° C., indicating agreement with the energy levels of the alkaline metals. With Li he obtained only the resonance line, while K gave its two most easily excited lines, Rb four and Cs six. By feeding the flame with oxygen lines of higher energy value can be brought out.

Temperature Radiation of Solids and Liquids. — The radiation of solids and liquids is especially adapted to the proof of Kirchhoff's law, since in these cases we can be sure that it is the result of temperature alone.

That the emission of light by heated substances is proportional to the absorption can be easily shown by heating a fragment of a piece of decorated china in a blast lamp. The design emits much more light than the white background, owing to its stronger absorbing power.

The more powerfully a body absorbs, the more powerfully will it emit when heated, this relation holding for every individual wave-length. Black bodies then give out the most light when heated. The fact that a white block of lime is far more luminous than a carbon rod when heated in the oxyhydrogen flame is not usually cited in support of this law(!), while the fact that the most luminous body of all, the Welsbach mantle, is also quite white is equally unsatisfactory as an illustration, for white bodies are in reality transparent, that is, they are made up of masses of small transparent particles, and transparent bodies ought not to emit at all. It is of course necessary to define just what we mean by transparency in this case, and it may be well to consider first a somewhat analogous case. The absorption which is accompanied by high emissivity is true absorption, and not selective reflection — which is sometimes confused with absorption. A highly reflecting polished metal surface is a poor radiator, but by proper disposition of its surface we may give it the power to absorb and emit. A bundle of polished steel needles with their points all turned towards the source of light reflects scarcely any light at all, the rays undergoing

multiple reflections between the conical ends of the needles. It is black for the same reason that black velvet is blacker than ordinary cloth. Such a bundle of needles should emit much more powerfully than a polished steel surface, for each needle, seen end-on, sends not only emitted light to the eye, but reflects rays coming from its neighbors. The surface formed by the points of the needles can be regarded as an absorbing surface, which absorbs in virtue of its structure: it is analogous to the hollow "black bodies" which will be considered presently. Highly reflecting metals such as silver when in a finely divided state are black. The case is analogous to that of the needles except that the action of the small particles is a scattering of the light rather than reflection.

The point to be emphasized is that such a surface, which absorbs but little in virtue of its molecular nature, may be also a powerful radiator, the mechanism by which its radiating power has been increased being as indicated above. Suppose now we take a perfectly transparent body, which like a perfect reflector has no emitting power. A bead of microcosmic salt (sodium pyro-phosphate) heated in a blast lamp is a good example. Though the platinum wire which supports it glows with vivid incandescence, the bead remains quite dark. A glass bead, however, emits a good deal of light, since its transparency is much less at high temperatures, a very common behavior of transparent substances. The microcosmic salt on cooling solidifies and at this instant becomes traversed by hundreds of cleavage planes, and suddenly glows with a dull red light. On reheating it it emits light strongly until it finally fuses into a transparent drop, when it instantly becomes dark again.

We have perhaps a better illustration, and one which is easier to handle, in the case of mica. A small spot on a thin strip of mica is calcined by directing a small pointed flame against it. The heated spot becomes silvery white and quite opaque, the change resulting from the formation of a multitude of reflecting laminae. If held at the polarizing angle it becomes more transparent, owing to the refusal of the thin laminae to reflect at this angle components of the vibrations which are parallel to the plane of incidence.

If now the strip of mica is held in the hot gases above a small Bunsen flame a region can be found where the temperature is high enough to cause the calcined spot to glow on the comparatively dark background of the transparent mica.

The discussion of this case is simpler than that of the bead of microcosmic salt, as the cleavage planes are parallel to the surface. The increment in reflecting power is of no help, as high reflection means low emission. The author's suggestion that the emission of

transparent bodies might be stronger at the surfaces than in the interior was favorably considered by Jeans, who suggested that the electron accelerations would be greater at the surface. It seems more probable, however, that the change produced by the cleavage planes is the result of release of radiation imprisoned by total reflection, in much the same way as in the case of the fluorescent film in contact with a white diffusing surface, mentioned in the Chapter on Fluorescence. A large proportion of the radiation originating within a flat plate or sphere, is prevented from escaping by total reflection, being for the most part absorbed before reaching the edge, which is its only outlet for escape in the case of a plate. In the case of the mica laminae it is doubtful whether total reflection can occur in such thin films, which are moreover very imperfect optically, forming a flat mosaic of small thin crumpled plates. Even if total reflection occurs, an edge is reached before much absorption has taken place.

These cases of increased emission due to cleavage planes were mentioned with some speculations as to the physical process in-

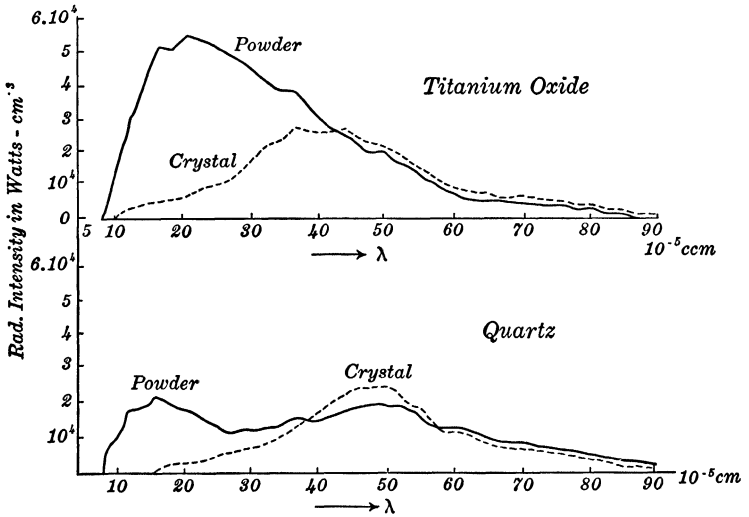


FIG. 440

volved in the earlier editions of this book. Their explanation is less obvious than in the case of the polished needles, and some have doubted their real existence. The recent determinations by Skaupy [1] and Schmidt-Reps [2] of the temperature emission of a sapphire crystal with and without cleavage planes showed an

[1] *Phys. Zeit.*, *28*, 842, 1927.
[2] *Zeit. für Tech. Phys.*, *6*, 322, 1925.

enormous increase of intensity over the whole range from 10 μ to 90 μ in the case of the crystal with numerous cleavages. Curves for titanium oxide crystal and fused quartz are shown by solid lines, and for rods made of the powdered substances by dotted lines in Fig. 440. The effect of powdering the substance is to make its curve of emission approach more nearly that of a black body at the same temperature.

The radiation of a thin layer of crystalline quartz powder (particle size 1.3 μ) has been examined by A. H. Pfund.[1] A very strong *maximum* was found between 8 and 10 μ, the region in which Rosenthal found a *minimum* emissive power for a heated quartz plate. This is analogous to the increased emitting power of a highly reflecting metal when reduced to a very fine powder, platinum black, for example. The form of the curve was similar to the reflection curve shown in Fig. 442.

It is now a matter of common observation that the intensity of the light emitted by a substance when introduced into a flame, is not as great in the case of a black substance such as a carbon rod, as with a white material such as lime. This can be easily shown by heating the two materials in succession in an oxyhydrogen flame. The Welsbach mantle which is nearly white emits a luminous radiation of still greater intensity, while a quartz rod emits practically nothing. This difference in behavior can be illustrated by a hydraulic analogy, which though rather far-fetched, may be of help in visualizing what is taking place.

Hydraulic Analogy of a Radiating Body. — The radiator is represented by a tall hollow cylinder, open at the top and closed at the bottom, where there are a number of outflow pipes of different sizes. Water flows into the cylinder at a certain definite rate from a horizontal pipe or flume, the height of which (T_2) above the base of the cylinder represents the temperature of the flame. Obviously the level of the water in the cylinder will rise until the rate at which the water flows out exactly equals the rate at which it flows in. This height (T_1) is the temperature which the radiator acquires in the flame. The jets of water which issue from the tubes represent radiation of different wave-lengths, the small jets representing the short waves. Their velocity corresponds to *intensity* of radiation. We will first suppose our hydraulic radiator to represent a black body, say a lump of carbon. In this case all of the pipes at the bottom are wide open and we have the maximum outflow of all wave-lengths for any given temperature, *i.e.* for any given height of the fluid within the cylinder. (If we take the cylinder empty, and plunge it into water, jets will squirt into it through the pipes;

[1] *J. O. S. A.*, *23*, 270, 1933.

that is, it is a perfect absorber for all wave-lengths.) With all of the pipes open, however, the level of the water within the cylinder will not rise to any great height, owing to the limited rate at which water flows in from the horizontal pipe. This means that the lump of carbon in the flame does not rise to a very high temperature because it radiates energy at a high rate. At the low temperature there is comparatively little visible light in the radiation, for the shorter waves only appear in quantity at high temperatures. We can imitate this condition in our hydraulic model if we choose by putting valves on the inside of the tubes, those on the small tubes opening only at high pressures.

To make our model imitate the quartz rod we plug up all of the pipes. The cylinder now represents a transparent body. If immersed in water it absorbs nothing through the pipes, and no matter how high the level of the water rises in it when water is poured in there is no emission of fluid, in other words no radiation. The body rises in temperature until the temperature is equal to that of the flame, but there is no radiation. Take next the case of the lime in the oxyhydrogen flame. It is a partially transparent substance, and we can imitate it by plugging the tubes with glass beads or cotton. Owing to the lesser rate at which the water now flows out through the tubes, the level rises much higher than when the tubes are all open, and owing to the greater pressure (temperature) we have liquid jets through the small tubes (short-wave-length radiation). The inferiority in the emissivity is more than made up for by the higher temperature which the body can acquire. We are now ready for the Welsbach mantle. It has been conclusively shown by Rubens that the peculiar brilliancy of the thorium mantles, caused by a small trace of cerium, is due to the fact that the cerium makes the thorium selectively absorbing for the short waves at high temperatures. If we wave a Bunsen flame over a mantle in a brilliantly lighted room, it will be seen to turn yellow at a temperature a little below a red heat. In other words it becomes a strong absorber for the short waves. It is, however, transparent for the long waves, consequently it does not emit energy at anything like the rate at which a black body does, and in consequence can rise to a high temperature in the flame, exactly as a pure thorium mantle does. Its band of absorption in the blue region enables it to pour out visible radiations nearly as powerfully as those which a black body at the same temperature would emit, hence its enormous brilliancy. Our hydraulic model with all of its tubes plugged with cotton represents the mantle of pure thoria, while to transform it into the Welsbach mantle we have only to pull out the porous plugs from some of the smaller tubes. In this condition, owing to the impeded

flow in the large tubes, the water will rise in the cylinder to a great height, and we get very powerful jets from the small tubes which we have opened, much more powerful than in either of the previous cases considered. Of course with all of the tubes open we could get equally intense small jets if we poured the water in at the top at a sufficient rate. There is a limit to this rate, however, for it is obvious that the rate at which the water is poured in at the top corresponds to the rate at which the flame can pour energy into the radiating body, a circumstance which depends on the conductivity of the body for heat and other circumstances.

Lambert's Cosine Law. — This law states simply that the intensity of the radiation from a single point on the surface of a body which is luminous either by diffuse reflection or by temperature radiation falls off in intensity as we pass from perpendicular to grazing emission, being proportional to the cosine of the angle. As a result of this circumstance the intrinsic brilliancy of the surface will be independent of the angle under which it is seen. There are two cases of radiation, however, which do not follow this law, to which attention was drawn by the author many years ago. Surface fluorescence of glass, considered for the radiation *within* the glass, and X-ray emission from a flat plate, in which cases the intrinsic brilliancy of the source increases enormously with increasing obliquity. This can be shown by illuminating one face of a crown glass right-angle prism with the light of an aluminum spark, and viewing the fluorescent surface through the other face at different angles — as described under fluorescence of solids.

If Lambert's law held exactly a white-hot metal sphere should have the appearance of a uniformly luminous disk, but the law has been found to hold only in the case of black body radiation, *i.e.* the radiation from the interior of an enclosure at uniform high temperature.

Uljanin,[1] in 1887 pointed out that Lambert's law was in conflict with the relation of emission to Fresnel's reflection laws, basing his conclusions on earlier observations of Provostaye and Dessair (1847) on glass, and Möller (1885) on platinum, as well as his own observations on black glass at 4 μ and platinum, at 3 μ.

Variation of Intensity and Polarization with Angle of Emission. — The temperature radiation from polished surfaces of metal and other materials varies both as to intensity and state of polarization with the angle of emission. This means that the intrinsic brilliancy of the source will vary according to the angle at which it is viewed, and if viewed through a Nicol prism set to pass either the perpendicular or parallel component of the electric vector further

[1] *Ann. der Phys.*, *62*, 528.

alterations of intensity will be observed. This phenomenon is the result of the relation between the coefficient of emission and the reflecting power of the material, the latter quantity depending upon the angle of incidence.

These relations are shown by the curves of Fig. 441 from determinations made by Worthing [1] for visible light radiated from a molybdenum surface at 1880° C.

Ordinates represent the intrinsic brilliancy of the surface as seen at various angles for unpolarized radiation (upper curve) and for the two polarized components, ⊥ and ||, referring to the electric vectors' position with respect to the plane of emission, and not to the polarization plane in the sense used in the paper.

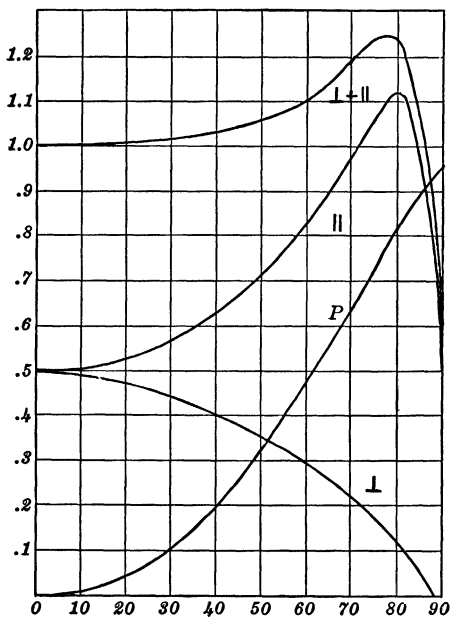


FIG. 441

$$P=\left(\frac{L\|-L\perp}{L\|+L\perp}\right).$$

We see that there is a slight increase with increasing angle of emission reaching a maximum at 78° and then a rapid fall, in the case of unpolarized radiation. The curve for the || component has a similar form, while the ⊥ component drops from .5 to 0 as we pass from 0° to 90°. The upper curve is the sum of the two lower. Bauer and Moulin [2] measured the emission coefficient for polished platinum for red light and total radiation as a function of the angle of emission. Their values follow.

ANGLE	E. CAL. (RED)	E. OBS. (RED)	E. TOTAL
0	1	1	1.
25	1	0.97	.97
50	1.045	1.0	1.03
70	1.167	1.1	1.35
80	1.159	1.13	1.77
85			1.90

[1] *J. O. S. A.*, *13*, 636, 1926.
[2] *Comptes Rendus. 190.* 167, 1910.

Their work indicates that the total emission at grazing emergence is nearly double that at normal emergence.

The inference from these curves is that a white-hot metal ball should appear with two "polar caps" of greater brilliancy than the rest of the sphere when viewed through a Nicol prism, the line joining the bright caps being parallel to the short diagonal of the Nicol. Experiments with a bicycle ball, while showing a trace of the effect, were not very convincing probably owing to oxidation of the surface, but a fused bead of pure silver gave a beautiful result. It was melted on a block of charcoal (graphite would probably be better) with a blow-pipe flame and viewed through a Nicol. The bright spots rotated with the Nicol. A still better way is to fuse the tip of a platinum wire in an oxy-coal-gas flame, or electric arc, forming a small spherical bead. This is heated to incandescence in a small vertical blast-lamp flame, and a greatly enlarged image projected through a Nicol prism with a short focus lens (an aplanatic hand magnifier for example).

Emission of Polarized Infra-Red Rays. — The ratio of the two polarized components as calculated from the optical constants of the metal agrees much better with the observed value for wavelengths in the remote infra-red. Czerny has employed an oblique platinum strip at 850° C., set to utilize radiation at nearly grazing emergence, as a source of polarized infra-red radiation. The values which he found for the ratio for different angles of emergence [a], and for $\lambda = 6.76\ \mu$ are given in the following table:

θ	$\cos^2\theta$	Calculated	Observed
60°	.25	.266	.277
70	.117	.130	.138
75	.067	.078	.084
80	.030	.038	.045
84	.012	.016	.019
87	.0027	.006	.008

The last value in the table shows that for grazing emergence the radiation of 6.7 μ is almost completely plane-polarized with the electric vector parallel to the plane of emission, or perpendicular to the radiating surface, as is the case with light transmitted through a very oblique plate of glass. Discrepancies between the observed and calculated values were attributed to imperfections in the surface.

Relation between Emission and Reflection. — In the case of substances which are not transparent, as we have seen, a portion of the incident energy is reflected and a portion absorbed. If the intensity of the light is 1, the absorbed fractional part A, and the reflected part R, we have $1 = R + A$ or $A = 1 - R$. Substituting

this value in our formula $E/A = e$ gives us an expression for Kirchhoff's law, in which the relation between emissivity and reflecting power is established:

$$\frac{E}{1-R} = e, \text{ the emissivity of a black substance.}$$

This means that a high value of R is accompanied by a low value of E, the emissivity, a relation which holds in the case of polished metals of high reflecting power, of which more will be said presently.

An especially instructive case is that of quartz which is highly transparent over the visible and near infra-red, but has a band of absorption at 2.5 μ which should give moderate emission, and another band at 8.5 μ which is so strong that it gives metallic reflection, and hence a low emitting power at this point. This subject was investigated by Rosenthal [1] who studied the emission and reflection of quartz, mica and glass with a spectrometer and thermo-element, and compared the results with the values calculated from the above formula. The low emissivity at regions of the spectrum corresponding to those of the maxima of the reflection curve is clearly shown in Fig. 442, and the close agreement between the observed values and those calculated from the above formula may be regarded as a most excellent proof of the law. The curve indicates that the absorption on both sides of the band of metallic reflection is sufficiently strong to give the quartz an emission equal to that of a black body.

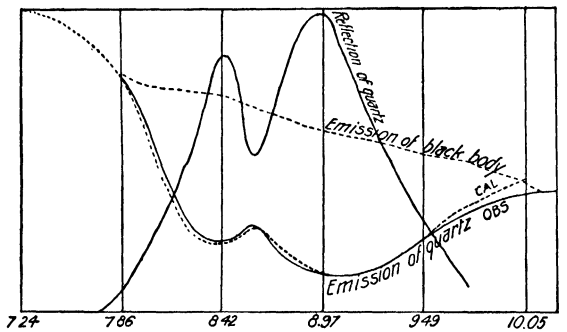


FIG. 442

Emission of Substances Having Absorption Bands. — The influence of absorption bands on the emission of otherwise transparent substances is well brought out in some recent experiments of the author.[2]

If a thin rod of very pure fused quartz, containing no small particles of foreign matter, is heated in the flame of a Bunsen burner, there is practically no emission of light, in conformity with

[1] *Wied. Ann.*, *68*, 783.

[2] Selective Radiation of Colored and Pure Fused Quartz. *Phys. Rev.*, *38*, 487, 1931.

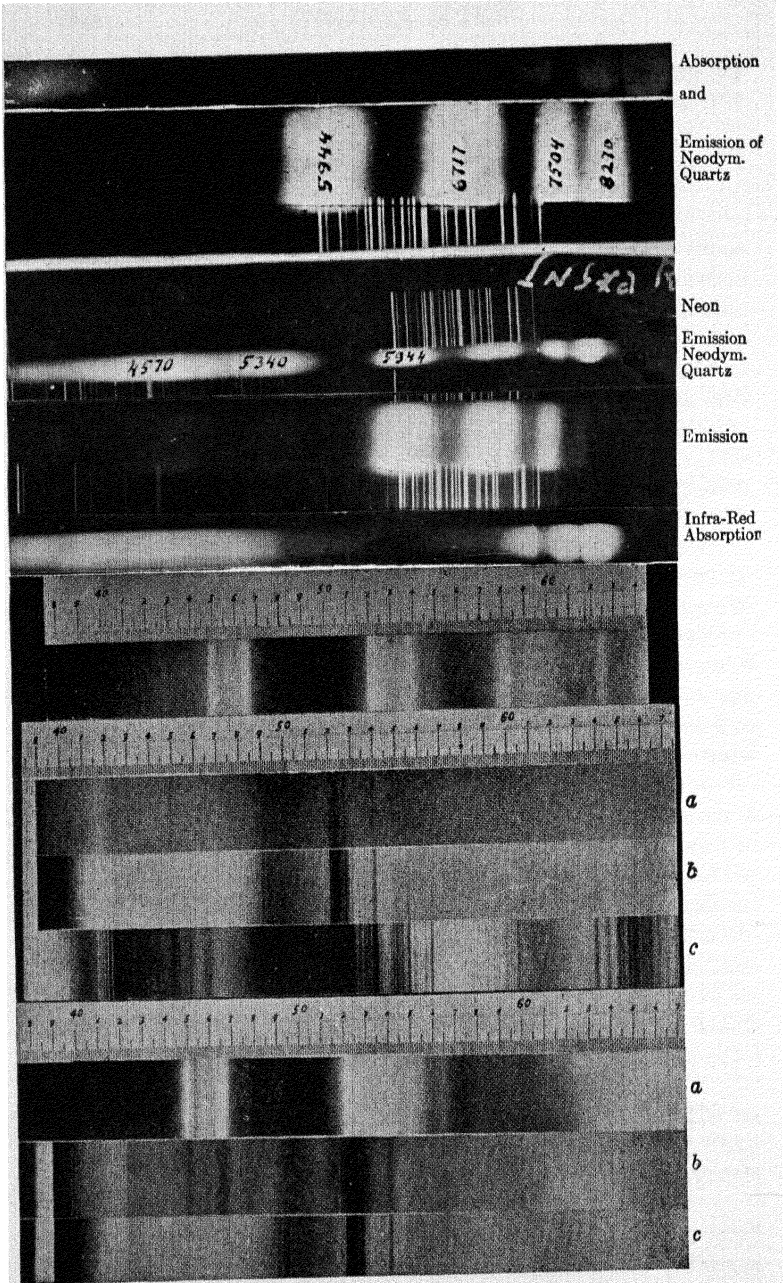


PLATE 16. EMISSION AND ABSORPTION OF QUARTZ CONTAINING A TRACE OF NEODYMIUM. ABSORPTION AND EMISSION OF RARE EARTHS.

Kirchhoff's law, which predicts that an absolutely transparent body will not radiate, no matter how high its temperature. If heated in the oxy-coal-gas flame, it becomes vividly incandescent. The blue white luminosity fades away on removal from the flame without passing through the orange and red stages exhibited by practically all substances shining by thermal radiation.

This indicates clearly that quartz exhibits selective thermal radiation because of the circumstance that, at high temperatures the ultra-violet absorption band migrates towards the region of longer wave-lengths. Radiations in the violet and near ultra-violet preponderate owing to the higher value of the absorption coefficient in this region.

This selective emission was clearly brought out by a spectrogram made in the following way: a fragment of graphite was placed in a small quartz capillary tube and the tube was fused down to a cylindrical rod, with the graphite at its centre. The rod was mounted vertically, and heated with a blast lamp operated by only sufficient oxygen to bring the graphite to a bright yellow heat. An image of the glowing fragment was focussed on the centre of the slit of a quartz spectrograph, and a series of exposures made with increasing oxygen supply to the flame. When the rod reached full incandescence and was just beginning to soften, it was apparent that its image on the slit was much whiter than that of the graphite fragment. Also the spectrogram showed that while the graphite was the brighter in the red and yellow region the radiation from the rod was more intense in the violet and ultra-violet. This might at first sight be interpreted as short wave-length luminescence since no substances can emit a more intense thermal radiation in any region of the spectrum than a black body. The real explanation appears to be that the graphite is at a slightly lower temperature than the quartz owing to its more rapid dissipation of energy in the form of long-wave radiation. Its temperature is not necessarily that of the quartz, depending upon the rate at which energy can be fed in from the quartz in comparison to the rate at which it is lost by radiation.

A solid solution of neodymium oxide in fused quartz was drawn out into thin rods and fibres which, when heated in the flame of a Bunsen burner and viewed through a direct vision prism, exhibited a most beautiful discontinuous spectrum consisting of a red, and orange-yellow and a green band, separated by perfectly dark intervals.

Photographs of the emission spectrum are reproduced on Plate 16 with comparison spectra of neon. The wave-lengths of the band centres are indicated.

The absorption of the neodymium quartz was next investigated. Its absorption coefficient is so low, thin rods appearing perfectly transparent, that an estimated layer of about two centimetres thickness would be required. No facilities for preparing a large mass of the material being available, a large fused bead was drawn out into a rod 2 cms. long by 1 mm. in diameter. This was mounted in a horizontal position by passing it through holes in two pieces of asbestos board and the light of a distant mazda lamp focussed on one end of the rod. The light thus traversed the entire length of the rod by internal reflection, and an image of the other end was focussed on the spectrograph slit. During the exposure the quartz rod was heated to a temperature sufficient to cause the emission of a feeble yellow light. It is important to secure a record of the absorption at a temperature not far from that at which the emission spectrum was secured, since the absorption bands all move toward the region of longer wave-lengths with rising temperature.

The absorption spectra, secured in this manner are reproduced in coincidence with the emission spectra at the top and bottom of the figure.

A few words as to the preparation of the quartz may be of help to any wishing to prepare specimens for demonstration. Neodymium oxide can be prepared from the neodymium ammonium nitrate, obtainable from the Welsbach Light Company of Gloucester City, Pennsylvania, by precipitating the oxalate with oxalic acid, and heating the dried precipitate (after washing) in a platinum crucible to a white heat with constant stirring.

A speck, the size of a pinhead is introduced into the middle of a quartz tube, of say, 3 mms. external and 1 mm. internal diameter. The spot containing the oxide is fused in the oxy-coal-gas flame, and twisted and drawn out over and over again. The rapid twisting of the fused bead, combined with drawing out and reforming into a bead, eventually gives a homogeneous mass. In the earlier stage we find the bead filled with spiral streaks of a bluish color. After the bead is uniform it is drawn out into a thin rod or fibre, of say, 0.5 mm. diameter, and bent at a right angle to the rod which serves as a support for holding it vertically in the Bunsen flame.

A study of the absorption and emission spectra of the oxides of some of the rare earths has been made by Anderson [1] who found that the absorption spectrum, obtained by the "body-color" method, that is by illuminating the powdered or fused substance with white light, was very different at different temperatures. The emission spectrum of heated neodymium oxide showed

[1] *Astrophysical Journal*, xxvi, No. 2, 1907.

very broad and hazy band at $\lambda = 58$–60, the spectrum being almost continuous. If the oxide existed as an impurity in erbium oxide the bands were much narrower, as shown on Plate 16, Fig. 1. When mixed with calcium oxide, however, the bands were similar to those obtained with the pure oxide. The absorption spectrum obtained by illuminating the substance with white light varied with the nature of the surface, as is shown on Plate 16, Fig. 2, in which *a* is the spectrum obtained with the powder; *b*, with a rod of the substance heated for 100 hours in a Bunsen flame; and *c*, with a rod the surface of which was fused with the oxyhydrogen flame. The bands are much blacker and more numerous in the case of the fused oxide, a circumstance which may be ascribed to the penetration of a greater thickness of the substance by the light in this case. In Fig. 3 we have spectrum *a*, the emission of incandescent erbium oxide, *b* its absorption at a high temperature, and *c* its absorption at room temperature.

As will be seen the absorption spectrum at a high temperature is complementary to the emission spectrum.

Emission of Polarized Light by Crystals. — Certain crystals, tourmaline for example, have an absorbing power which differs according to the plane in which the vibrations are taking place. Suppose the crystal to be so oriented that its absorptive power is greatest for horizontal vibrations. We might expect, on heating the crystal, to find a preponderance of horizontal vibrations in the emitted light. This was found to be the case by Kirchhoff, who heated a crystal in a Bunsen flame and found that, on viewing it through a double-image polarizing prism, one of the images was distinctly brighter than the other.

A quantitative proof of Kirchhoff's law in the case of glowing tourmaline has been made by Pflüger [1] who measured with a spectro-photometer the absorption and emission of the crystal at the same temperature and for the same wave-length. If J is the intensity of the incident light, JD that of the transmitted light, and JR that of the reflected, the intensity of the absorbed light $A = J(1-R-D)$. R was calculated from the reflection formula $R = [(n-1)/(n+1)]^2$, while D and E were observed with the spectrophotometer. Designating by E_o and E_ϵ the emissivity for vibrations parallel respectively to the vibrations of the ordinary and extraordinary rays, and by A_o and A_ϵ the corresponding absorptive powers, we have, if Kirchhoff's law can be extended so as to include the direction of the vibration,

$$\frac{E_\epsilon}{A_\epsilon} = \frac{E_o}{A_o} \text{ or } \frac{E_\epsilon}{E_o} = \frac{A_\epsilon}{A_o}.$$

[1] *Ann. der Phys.*, 7, 806, 1902.

After eliminating all sources of error, Pflüger obtained as final values for the two ratios,

$$\frac{A_\epsilon}{A_o}=.650,\ \frac{E_\epsilon}{E_o}=.641,$$

a very beautiful verification of the law as applied to anisotropic media.

Proof of Kirchhoff's Law and Black-Body Radiation of a Hollow Enclosure. — The following proof of Kirchhoff's law and the fact that the radiation within a uniformly heated enclosure is the equivalent of the radiation of a perfectly black substance at the same temperature, regardless of the material of which the enclosure is made, was given many years ago by E. Pringsheim. It is based on the Carnot principle that heat cannot pass from a body at any given temperature to one of higher temperature without the expenditure of work.

The most logical and concise treatment is due to Pringsheim.[1] Consider a ball κ composed of any material enclosed in a hollow vessel, opaque to radiation of all wave-lengths, and uniformly heated to any given temperature. The ball emits in unit time the total radiation E, while there falls upon it from the walls in the same time the amount e, of which the fraction Ae is absorbed. Since by Carnot's principle the temperature cannot change, the amount of radiation emitted by the ball must equal the amount absorbed, so that $E=Ae$. If the ball is made of a conglomerate of different substances, some parts of its surface may absorb more strongly than others. Suppose we rotate the ball: the amount of energy e falling upon it will only be changed by an infinitely small amount, since only the part of the radiation which came originally from the body, and is reflected back from the walls, can be responsible for the change: the amount of this which falls across the body is of course very small. If, however, the radiation from the walls is not uniform, *i.e.* if it has especial states of polarization, or is more intense in certain directions than in others, the amount of heat absorbed by the body would be changed by its rotation.

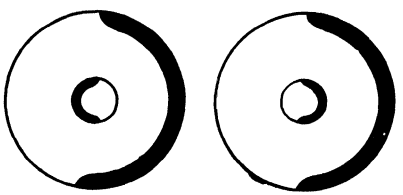


FIG. 443

Suppose, for example, the enclosure was black on one side and white on the other, and that the same thing was true for the ball.

[1] *Verh. d. deutsch. physik. Ges.*, *3*, 81–84, 1901.

If the radiation is more intense from the black wall of the enclosure than from the white, more heat will be absorbed by the ball in position *b* than in position *a* (Fig. 443).

That the intensity of the radiation from the inner surface of a hollow vessel is independent of the nature of the material can be shown by placing a fragment of decorated china in a porcelain crucible heated over a Bunsen burner. If the cover of the crucible is put on, a small opening being left through which the interior can be viewed, and the flame of a second burner be directed upon it so as to bring the whole to a nearly uniform temperature, the decorations on the china will be quite invisible, the radiations from them being equal to the radiation from the rest of the surface. As we have seen, if the china is heated in the open air the dark portions radiate more strongly, the design appearing brighter than the background. The cause of the equality in the case of an enclosed radiator can be very simply stated. The radiation is made up of two parts, the emitted and the reflected, the latter coming from the heated walls. Dark portions of the material emit more powerfully than white portions, since their power of absorption is greater; on the other hand, they reflect scarcely any of the radiation from the walls. The white portions, which emit feebly, reflect powerfully, and, owing to the proportionality between emission and absorption, a perfect balance is secured.

The Perfect Black Radiator. — While the principle that the radiation within a closed space at a uniform temperature is identical with the radiation of a perfectly black body had been recognized for many years, Wien and Lummer [1] were the first to actually prepare radiators acting on this principle, and make use of them in experimental work.

For studying the intensity of the radiation at low temperatures and the distribution of energy in the spectrum of the radiation, a hollow cylinder of brass, blackened on the inside, can be used. The cylinder is provided with a small aperture, and is surrounded by a steam jacket, or imbedded in a mixture of sodium and potassium nitrate, and the whole packed in felt. The smaller the size of the hole in comparison to the internal capacity of the cylinder, the more nearly does the emerging radiation compare with that of an ideal black body. For high temperature work a cylinder of platinum or porcelain, electrically heated, can be employed, or even a hollow iron ball heated in a gas furnace. Kayser has proposed a very simple device, which, though superior to an electrically heated strip of blackened platinum, is not as good as a hollow vessel. Two strips of platinum, one provided with a narrow slit. are mounted

[1] *Wied. Ann., 56,* 451, 1895.

opposite to one another and heated to the same temperature by a current. The principle is of course the same as that of the device just considered.

Paschen [1] has proposed still another device. A glowing carbon filament is mounted at the centre of a hollow silvered sphere. Assuming the silver to reflect all of the energy, it can be regarded as a hollow vessel having the same temperature as the carbon filament. The radiation escapes as before through a small hole. This same device has been applied to the bolometer, the absorbing strip being mounted at the centre of a hollow spherical chamber silvered on the inside. All radiation not absorbed at once by the bolometer is returned to it by the reflecting surface. In this way it is possible to prepare a perfectly black bolometer.

Equilibrium between Radiation and Material Bodies. — In the deduction of the laws of radiation the use of a conception due to Bartoli, which, though it cannot be carried out experimentally, leads to important laws which can be verified in other ways. The idea in brief is to apply the principles of thermodynamics to radiation, performing a cyclical process similar to Carnot's cycle, employing radiation instead of a gas as the working substance.

The radiation within a hollow vessel can be in equilibrium with the walls or with bodies in the interior, only when it is of the same nature as the radiation emitted by the walls or the bodies contained within the vessel.

To get an idea of exactly what we mean by equilibrium between radiation and a material body we will consider the following case:

Suppose we have a hollow vessel the walls of which are perfect reflectors, which contains only ether. If we fill this cavity with monochromatic radiation, say that of the sodium flame, by opening a door in the wall and allowing the light to enter, which, of course, can be done perfectly well in theory, the radiation will, if we close the door, be reflected back and forth within the vessel forever. It will neither change in intensity nor alter its wave-length; in other words, it is in equilibrium with the reflecting walls. We shall now prove that a perfect reflector is the only body with which this radiation can be in equilibrium, with the exception of the flame which originally emitted the light. Suppose we introduce a small fragment of absorbing matter within the cavity of the reflecting vessel. It will immediately absorb the monochromatic sodium radiation as fast as this radiation falls upon it, and in a very short space of time the monochromatic waves will have vanished completely. The temperature of the absorbing body will be slightly elevated, and it will emit long heat-waves, the energy being distributed over a wide

[1] *Wied. Ann.*, *60*, 791, 1897.

range of wave-lengths, the range and distribution depending on the temperature of the body. This radiation will now fill the cavity in place of the sodium radiation, and it will be in equilibrium with the absorbing body, *i.e.* a permanent state is speedily reached, after which there is no further change.

Pressure of Radiation. — The radiation within the vessel exerts a pressure upon the walls and upon the surface of the absorbing body. As we are to make use of this pressure in the derivation of laws it will be well to investigate it somewhat in detail.

Maxwell, in his electro-magnetic theory, showed that radiation must exert a pressure when it falls upon a reflecting or absorbing surface. As this pressure is the foundation upon which the laws of radiation have been built, we will briefly consider the phenomenon.

Maxwell showed that when plane electro-magnetic waves fall in a normal direction upon a perfectly absorbing surface, the pressure exerted on unit area is equal to the energy contained in unit volume of the vibrating medium.

That a pressure is exerted by heat (or light) waves may be proven by making use of the idea of Bartoli. Consider a cylinder, composed of some material which reflects perfectly, closed at the ends by black plates at temperatures $T_1 > T_2$ (Fig. 444). Introduce a screen S, made also of a reflecting material, which divides the cylinder into two compartments. The body at temperature T_1 will fill the upper compartment with radiation of energy corresponding to its temperature. T_2 (at a lower temperature) will fill the lower compartment with radiant energy of less density. Let B represent a movable reflecting diaphragm, provided with a sliding door, which, when open, allows the energy from T_2 to fill the middle compartment. Now close the door and raise the diaphragm or piston. The volume of the middle compartment is decreased, and the density of the radiant energy "trapped" within it is increased. On removing the screen S laterally, which we can do without performing work, the diaphragm will drive the radiant energy above it into the body T_1. We have thus taken energy (or heat) from a body at low temperature and carried it to one of high temperature, which by the second principle of thermo-dynamics is impossible, unless mechanical work is done in the operation. This work can have resulted only from the overcoming of a pressure exerted upon the diaphragm, the vibrating medium resisting compression in the same way that a gas does. This pressure becomes greater as the volume is diminished owing to the increase in the energy density. In the case of the

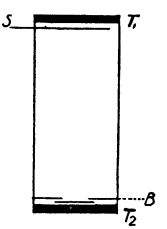


FIG. 444

compression of a gas, the molecules rebound from the moving piston with increased velocity, consequently the force of each blow, and the number of blows per second, are increased.

In the case of compressed radiation the mechanism is not so easy to follow; as we shall see presently, reflection from a moving diaphragm decreases the wave-length by an amount proportional to the distance through which the diaphragm moves (provided the rest of the vessel is reflecting). This means that the number of waves which strike it per second will be increased. The amplitude, as we shall see presently, remains the same, and the increase of energy density is due solely to the fact that more waves are present in unit length of the train after the compression than existed before the motion of the diaphragm. Such a process as that described above cannot be even approximately realized experimentally. It is no less valuable, however, as our inability to carry it out is due solely to mechanical difficulties and our inability to obtain a substance which reflects perfectly.

The pressure due to light was, for a long time, sought for in vain. The disturbing effects of "radiometric action," or the reaction pressure of gas molecules rebounding from the surface heated by the radiation, completely masked the very small effect which was looked for. As early as 1754 an attempt was made by DeMairan and Du Fay to detect the pressure of light. This was of course in the days of the corpuscular theory, and the looked-for pressure was that due to the arrest of the flying corpuscles. Fresnel, Zöllner, Bartoli and Crookes also searched in vain for evidences of the pressure, the experiments of the latter, however, resulting in the discovery of the radiometer. The pressure was first observed by Lebedew [1] in 1900 and by Nichols and Hull independently at about the same time. Though the latter investigators were anticipated by Lebedew by some months, their investigation was conducted with greater care, and the errors due to gas action were more carefully eliminated.

Experiments of Nichols and Hull.[2] — The inability of previous observers to measure the pressure due to radiation was due to the apparent impossibility of separating the effect from the so-called "radiometer" action. Thin vanes were employed to detect the pressure, and the radiation warmed the side on which it fell. When such a condition exists the gas exerts a greater pressure on the warm than on the cold side, and in general this pressure is vastly greater than the true radiation pressure (Fig. 445*a*).

Nichols and Hull finally succeeded in eliminating the gas action

[1] *Rapp. près. au Congres de Phys.*, *2*, 133, Paris, 1900.
[2] *Proc. Am. Acad.*, xxxviii, April, 1903; *Phys. Rev.*, xiii, 293, 1901.

by employing a suspended vane made of two circular disks of thin glass silvered on one side. By employing a reflecting surface the pressure is double that exerted upon a black surface, and the heating is reduced to a minimum. By measuring the deflections when the glass and silver sides were illuminated in succession the gas action could be calculated, for the silver surface is the one heated in both instances. This is due to the fact that the radiation, before its impact upon the vane, has passed through a number of lenses and plates of glass, and is consequently robbed by absorption of all rays capable of heating a glass surface. It is at once apparent that when the radiation falls upon the glass surface the gas pressure and the light pressure are opposed, while when the silver surface is illuminated they act together, *i.e.* in the same direction. Larger deflections are of course observed in the latter case than in the former. To still further eliminate gas action, the ballistic method was adopted; it had been observed that some seconds' or even minutes' exposure to the radiation were required before the gas pressure reached its maximum, while the radiation pressure is of course instantaneous. Very short exposures were consequently given, and the ballistic deflection of the vane was observed by means of a mirror and scale.

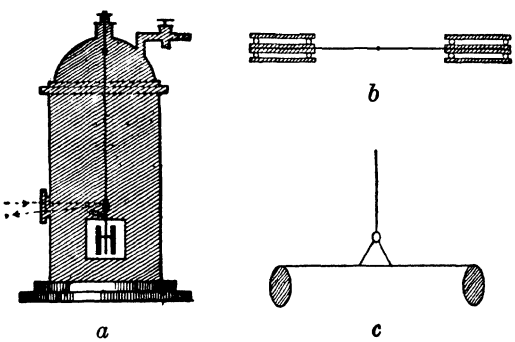


FIG. 445

By an elaborate series of experiments the investigators determined the most suitable pressure for the air in the chamber in which the vane was suspended, the pressure, in other words, at which the gas action was at a minimum. This pressure proved to be about 16 mms. of mercury. After measuring the value of the radiation pressure, the energy of the radiation was determined, by allowing it to fall upon a blackened silver disk. The rise of temperature of the disk was determined by means of iron-constantan thermo-junctions imbedded in the disk. From these energy measurements the pressure to be expected was calculated.

The radiation employed was that of an arc-lamp, either with or without absorbing screens. After correcting for all possible sources of error, the following values were obtained:

Radiation	Pressure in 10^{-6} Dynes Observed	Pressure Calculated from Energy Measurements
Through air only	$7.01 \pm .02$	$7.05 \pm .03$
Through red glass	$6.94 \pm .02$	$6.86 \pm .03$
Through water cell	$6.52 \pm .03$	$6.48 \pm .04$

These experiments can be regarded as establishing in a quantitative manner the existence of the Maxwell-Bartoli pressure, which measured in dynes per sq. cm. is equal to the energy contained in unit volume of the radiation. A reflecting surface doubles the energy density in the medium in front of it by superposing the reflected beam upon the incident.

The gas action was subsequently eliminated by Hull [1] by enclosing the reflecting and absorbing surfaces in thin glass cells, as proposed in the earlier paper. The silvered side of a thin cover-glass was placed in contact with the blackened side of a similar glass, and the whole enclosed by means of two other thin glasses as shown in Fig. 445*b*. Two cells of this description were mounted upon opposite ends of a torsion arm suspended in a receiver from which the air could be removed. When the light falls upon the blackened surface and is absorbed, the temperature of the two outer glass surfaces of the cell are the same, since they are separated from the heated surface by equal thicknesses of glass and air. The gas action should therefore be equal on the two surfaces. Any gas action occurring within the cell will produce no effect, owing to the equality of action and reaction. Hull found that the ratio of the deflections obtained when the silvered and blackened surfaces were illuminated in succession, agreed with the calculated ratio to within 2%, showing that the "radiometer" action had been practically eliminated.

Tangential Component of Radiation Pressure. — An interesting experiment was described by Poynting at the Cambridge Meeting of the British Association in 1904, in which gas action is completely eliminated.

When radiation is incident upon an absorbing surface in an oblique direction, the pressure has a component parallel to the surface. In the case of a reflecting surface this tangential force cannot be detected, since the incident and reflected beam give rise to equal and opposite forces parallel to the surface. The magnitude of the force, when E is the energy density, μ the fraction reflected, and α the angle of incidence, is given by

$$F = \frac{E}{2}(1-\mu) \sin 2\alpha.$$

[1] *Phys. Rev.*, May, 1905.

The existence and magnitude of the force was observed with the apparatus shown in Fig. 445c. Two thin glass disks were mounted on the ends of a fine glass rod, the system being suspended by a quartz fibre in a brass box provided with glass windows. One of the disks was silvered, the other blackened, and the pressure within the case was reduced to 1 cm. Sunlight, or the beam from an arc-lamp, was directed against the black disk at an angle of 45°. Gas action due to heating will give rise to a pressure normal to the surface, but there will be no tendency to rotate the suspended system. The tangential component of the radiation pressure, on the other hand, will produce a deflection, the magnitude of which can be read with a mirror and scale. E was calculated from the observed deflection, and was found to be $5.8 \cdot 10^{-6}$ dynes, while a direct measurement of E, by the heating of a silver plate, gave the value $6.5 \cdot 10^{-6}$.

Pressure of Radiation on an Absorbing Gas. — The repulsion of the tails of comets by the sun has been explained as the result of the pressure exerted by the solar radiation. If we reduce the size of an obstacle its mass becomes less in proportion to its surface, and the pressure of the radiation may eventually become greater than the attraction of gravitation. It has been shown by Schwartzchild, however, that there is a critical size at which the ratio of pressure to gravitational attraction has its greatest value. In other words, if we make the particles too small the radiation no longer exerts any pressure on them, for they no longer act as obstacles, or diffract light. Now the spectroscope shows us that the tail of a comet is gaseous, and the gas molecule is very much smaller than the smallest obstacle capable of feeling the pressure of radiation. There is this difference, however. The gas molecule may be capable of stopping the radiation by resonance, *i.e.* the gas may absorb, and it seems quite probable that radiation may exert a measurable pressure on the molecules of a gas, in spite of the fact that a cloud of material particles, each one of which is vastly larger than the molecule, experiences no pressure at all. This question has been very successfully attacked by Lebedew,[1] who has measured the pressure which the radiation from a Nernst lamp exerts upon various absorbing gases, such as CO_2, methane, butane, propane, etc. The gas was enclosed in a cell provided with fluorite windows, and was set in motion by the pressure of the radiation. This motion was communicated to a very light torsion balance made of magnesium, by which it was rendered visible.

The Stefan-Boltzmann Law. — An empirical law was deduced by Stefan from observations made by other observers on the in-

[1] *Ann. der Phys.*, April, 1910.

tensity of the total radiation from bodies at different temperatures. The law states that the complete emission S of a black body is proportional to the fourth power of the absolute temperature T, or

$$S=\sigma T^4,$$

in which a is a constant.

This same law was subsequently deduced from theoretical considerations by Boltzmann,[1] who availed himself of the ingenious conception by which Bartoli proved that radiation must exert a pressure. Consider a hollow cylinder of unit cross section, the walls of which are black, and of infinitely small heat capacity. The ends of the cylinder are also black, but of infinitely large heat capacity. Within the cylinder is a frictionless black piston, in contact with the left-hand end-plate of the cylinder, which has an absolute temperature T_0. The opposite end-plate has a lower temperature T (Fig. 446). The radiant energy exerts a pressure on the piston, which in the case of plane-waves parallel to the surface is equal to the radiant energy in unit volume of the ether. Let $\Psi(T)$ be the energy in unit volume. Since the energy is travelling in all possible directions, the pressure on unit surface will not be $\Psi(T)$ but $\frac{1}{3}\Psi(T)$. (Compare with the calculation of the pressure due to molecules moving in all directions, in the Kinetic Theory of Gases.) We have then $\frac{1}{3}\Psi(T)=f(T)$, the pressure at temperature T.

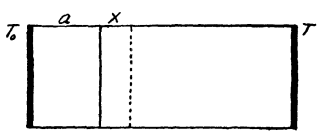


FIG. 446

Now let the piston move forward a distance a, under the influence of the pressure of the radiant energy coming from the high temperature plate T_0. The heat energy leaving T_0 is partly spent in doing the work $af(T_0)$ on the piston, and partly in filling up the volume "a" with radiant energy. This process obviously corresponds to the isothermic expansion of the Carnot cycle, the filling of the space a with energy corresponding to the heating of the gas. The amount of heat which leaves the plate T_0 is $a[\Psi(T_0)+f(T_0)]$.

We will now introduce a screen impervious to heat immediately in front of T_0, which prevents further radiation into the space to the left of the piston. This corresponds to placing the cylinder in Carnot's cycle upon an insulating stand. The pressure to the left of the piston is greater than that to the right, owing to the higher temperature of the plate which filled this portion of the cylinder with radiation. The piston will therefore move forward until the energy per unit volume is the same on both sides. Allowing this

[1] *Wied. Ann.*, *22*, 291, 1884.

adiabatic expansion to take place we have $d[(a+x)\Psi(T)]=-f(T)dx$, the characteristic equation for an adiabatic process (see any Thermodynamics). In this expression T is of course variable. During this process the volume to the right of the piston has been still further diminished, and an amount of heat energy represented by $(a+x)[\Psi(T)+f(T)]$ due to diminishing the volume and work done enters the plate at temperature T.

Since the process is reversible we have, by the second law of thermodynamics,

$$\frac{(a+x)[\Psi(T)+f(T)]}{T}=\frac{a[\Psi(T_0)+f(T_0)]}{T_0}=C,$$

in which x and T are variables. From this can be derived

$$\tfrac{4}{3}\Psi(T)dT=\tfrac{1}{3}Td\Psi(T),$$

$$\frac{d\Psi(T)}{\Psi(T)}=4\frac{dT}{T},$$

$$\Psi=\sigma T^4.$$

Proof of Stefan's Experimental Law. — The law was first deduced empirically from observations made on the rate of cooling of a blackened thermometer bulb.

Lummer and Pringsheim [1] proved the law over a range of temperatures included between 100° and 1300° C. by measuring the intensity of the radiation from a hollow chamber (black body) by means of the bolometer.

The constant a has been determined in absolute measure by Kurlbaum,[2] who heated the bolometer strip (screened from the radiation) by means of an electric current of known strength to the same temperature to which it was raised by the radiation. The radiation was thus determined in absolute units by calculating the Joule heat developed by the current. The value found was

$$a=1.71\cdot 10^{-5}\frac{\text{erg}}{\text{sec.}}=0.408\cdot 10^{-12}\frac{\text{gr. cal.}}{\text{sec.}}.$$

Optical Pyrometers. — Various types of pyrometers have been designed for measuring high temperatures by optical methods. Féry's instrument is based upon the law of total radiation. It consists of a telescope with a fluorite objective, in the focus of which

[1] *Wied. Ann.*, *63*, 395, 1897.
[2] *Wied. Ann.*, *65*, 746, 1898.

is mounted a sensitive thermo-couple as shown in Fig. 447. To use the instrument one has only to point it at the object, the temperature of which is to be measured, *e.g.* the interior of a blast-furnace, and focus the image upon the thermo-junction by means of the eye-piece, which is moved with the latter by means of the rack and pinion wheel at *B*. The temperatures are read with a galvanometer.

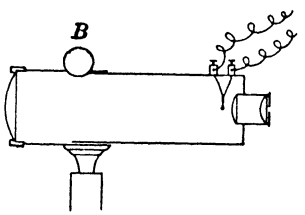


FIG. 447

Other optical photometers have been devised, based upon the laws which we are about to study.

Temperature of the Sun. — The sun's temperature has been computed by measuring the total radiation. Assuming the solar disk to be a black body, and taking for the value of the solar constant 3 gr. cal. per minute, the computed temperature comes out a trifle over 6000°.

Change in the Spectrum of a Black Body with the Temperature. Wien's Laws. — Making use of a conception similar to the one by means of which Boltzmann deduced Stefan's law, but extending it by the introduction of the consideration of the change in wave-length which occurs when radiation is reflected from a moving mirror, Wien [1] arrived at a formula which expressed the change in the spectrum of a heated black body with its absolute temperature. As is well known, when a solid or liquid is heated the longer heat-waves appear first, then red light, and finally at still higher temperatures the violet and ultra-violet. If we measure the energy at different points in the spectrum with the bolometer and plot these values as ordinates, with the wave-lengths as abscissae, we obtain the energy curve for the emission at the temperature in question. The maximum of this curve moves towards the region of the shorter waves as the temperature is increased, but there is an increase in the height of every ordinate; in other words, the curve does not move bodily down the spectrum.

Wien's treatment enabled him to account for the change in the distribution of energy in the spectrum as the temperature was raised, provided the energy curve for any given temperature was known. This relation is expressed by what are known as Wien's Displacement Laws, which give the change in position in the spectrum of the point of maximum energy with the temperature.

$$\lambda_m T = \text{Constant}$$

[1] *Wied. Ann.*, *46*, 633; *52*, 132.

in which λ_m is the wave-length for maximum energy and the energy E_m at the maximum.

$$\frac{E_m}{T^5} = \text{Const.} = 4.16 \cdot 10^{-12} \text{ in watts and cm.}$$

The cycle of operations performed by Wien with his imaginary cylinder with its reflecting pistons and trap doors will be found in earlier editions of this book, but is omitted now for lack of space.

Lummer and Pringsheim have tested these two laws by measuring the energy curves of a heated black body over a range 621–1650 Absolute. Their results are given in the following table:

T	λ_m	E_m	$A = \lambda_m T$	$B = E_m T^{-5}$
1650	1.78	270	2928	$2246 \cdot 10^{-17}$
1260	2.35	69	2959	2176
1094	2.71	34	2956	2166
908	3.28	13.6	2980	2208
723	4.08	4.3	2950	2166
621	4.53	2.03	2814	2190

Neither of the two equations, however, give us any information regarding the actual distribution of energy in the spectrum of a black body.

To express this we require an expression which represents E as a function of λ and T.

Complete Radiation Formulae. — Wien [1] by the consideration of a peculiar type of radiator, deduced a formula connecting E with λ and T. The radiator is considered as a hollow vessel filled with a gas mixture capable of emitting waves of all lengths. He assumes that every molecule emits only a single wave-length, which depends on its velocity, the intensity of which wave is a function of this velocity. Further, the intensity $\Phi(\lambda)$ of the radiation between the limits λ and $\lambda + d(\lambda)$ is proportional to the number of molecules vibrating with periods corresponding to wave-lengths within this range. From these assumptions he derived the formula

$$E = \frac{c}{\lambda^5} e^{-\frac{\beta}{\lambda T}}.$$

This formula represents the energy distribution very well if it is not applied to too long waves.

[1] *Wied. Ann.*, *58*, 662.

Planck deduced a radiation formula of different form from quantum considerations:

$$E = \frac{c\lambda^{-5}}{e^{\frac{c^2}{\lambda T}} - 1}.$$

For short wave-lengths this reduces to the Wien formula.

This formula has been confirmed in a remarkable manner by the work of Rubens and Kurlbaum [1] who measured the intensity of the radiation of $\lambda = 8.8$, 24, 31.6, and 51 μ emitted by a black body over a temperature range comprised between 85 and 1773 Absolute.

T	E Obs.	E Cal. (Wien)	E Cal. (Planck)
85	−20.6	−107	−21.9
193	−11.8	− 48	−12
293	0	0	0
523	+31	+ 63	+30.4
773	64.5	96	63.8
1023	98.1	118	97.2
1273	132	132	132
1523	164	141	166
1773	196	147	200
∞	—	194	∞

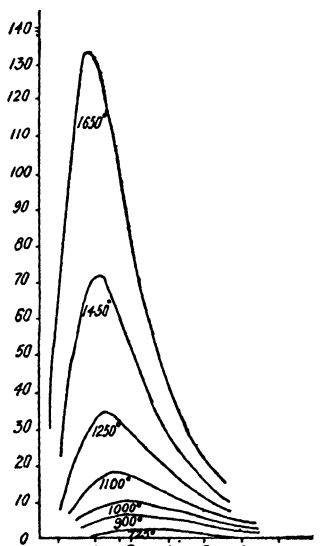


FIG. 448

Their results are given in the preceding table, for $\lambda = 51$ μ (residual rays from rock salt) together with values calculated from the formulae of Wien and Planck.

This table shows not only the close agreement between the observed values and those calculated from Planck's formula, but also that at a temperature of 1773 a value of E was obtained, larger than the limiting value 194 for infinite temperature, calculated from Wien's formula.

For short waves Wien's formula is perfectly satisfactory. It is questionable, however, whether it is anything more than an empirical formula, for many objections have been raised against the methods employed in its deduction. The energy curves for a black body are represented in Fig. 448 for temperatures between 725 and 1650. The shift of the maximum towards the region of

[1] *Ann. der Phys.*, *4*, 649, 1901.

shorter wave-lengths with increasing temperature is clearly brought out. These curves were made from observations by Lummer and Pringsheim.

In the case of the sun's spectrum the point of maximum energy is the yellow-green region, corresponding to a temperature of 6000°.

The proportion of the total energy radiated by a black body which is comprised in any spectral region can be read in percent-

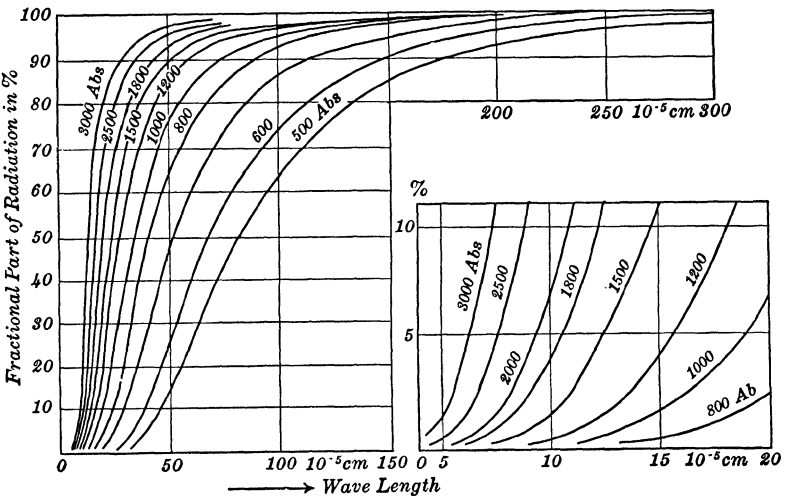


FIG. 449

ages from the very useful set of curves published by Holladay [1] reproduced in Fig. 449.

For example, at 1000° Abs. 27.5% of the radiation is in the region $3 \cdot 10^{-4}$ cms. and 72.5% in the region of longer wave-length. By formation of differences one can compute the per cent in any restricted region.

The Rayleigh-Jeans Formula for Spectral Distribution. — According to kinetic gas theory the average energy K of the molecules at temperature T is

$$K = \frac{\frac{3}{2}RT}{N} = \frac{3}{2}kT$$

in which k is the Boltzmann gas constant or R/N, R being the same as in equation $pv = RT$, and N the number of molecules per mole.

[1] *J. Opt. Soc.*, *17*, 329, 1928.

The molecules of a gas have three degrees of freedom, if we consider motion of translation only, and the law of the equipartition of energy states that the energy of each degree is $kT/2$.

Lord Rayleigh applied this law to the modes of vibration possible in the ether, for the derivation of a formula expressing the distribution of energy in the spectrum of black-body temperature radiation, and the theory was extended some years later by J. H. Jeans.

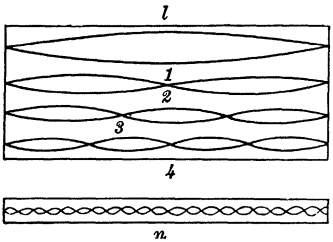


FIG. 450

We must first see what we mean by "modes of vibration" and their relation to "degrees of freedom." In a cylindrical enclosure (Fig. 450) of length l, with plane-waves reflected back and forth between the ends (wave-motion in one dimension) stationary waves can be formed of wave-lengths equal to $2l$ divided by 1, 2, 3, $4 \cdots n$ or frequencies c, $2c$, $3c \cdots nc$, each divided by 2 in which c is the velocity of light.

The number of possible systems of stationary waves of frequency less than ν is given by $$n = \frac{2l}{c}\nu.$$

For a wave-motion in two dimensions the number is $[(2l/c)\nu]^2$ since a wave along the diagonal of a square reflector xy can be considered as the resultant of two waves travelling parallel to the sides, and we can thus combine each wave in the one dimensional problem, considered as moving parallel to x with the total possible number of waves as given above moving parallel to y.

In three dimensions we have $$\left(\frac{2l}{c}\nu\right)^3 = 8\frac{\tau}{c^3}\nu^3$$

in which τ is the volume. A more rigorous treatment shows that the number of possible standing wave-systems between the frequency limits ν and $\nu \pm \delta\nu$ is

$$4\pi\frac{\tau}{c^3}\nu^2\delta\nu. \qquad (1)$$

This is for polarized radiation, and since natural or unpolarized radiation must be regarded as due to the composition of two waves of the same period and phase at right angles the above expression must be multiplied by 2, giving $8\pi(T/c^3)\nu^2\delta\nu$ which may be regarded as the number of degrees of freedom of the ether for the radiation in question. Now experiment has shown that the fre-

quencies and energies of the constituents of radiation coming out from an enclosure through a small hole are independent of the size or nature of the walls, so long as some non-reflecting matter is present.

If a small amount of absorbing matter at temperature T is placed in a box with perfectly reflecting walls, it radiates "pulses" of different type, each one of which can, by Fourier's analysis, be resolved into an infinite series of harmonic terms, and only those terms will be present (when equilibrium is reached) which conform to the stationary wave condition. Or we may regard the matter as containing oscillators of all possible frequencies. Ascribing the usual kT units of kinetic and potential energy to each degree of freedom ($kT/2$ to the kinetic and $kT/2$ to the potential), the energy of the radiation in the frequency range between ν and $\nu+\delta\nu$ is $(8\pi\tau/c^3)\nu^2kT\delta\nu$ ergs for the whole enclosure or $(8\pi/c^3)\nu^2kT\delta\nu$ ergs per ccm. To transform this to wave-length we have, by putting $\nu=c/\lambda$ and $\delta\nu=(c/\lambda^2)\delta\lambda$

$$\frac{8\pi kT}{\lambda^4}\delta\lambda \qquad \text{ergs per ccm.} \qquad (2)$$

This is the Rayleigh-Jeans formula. It indicates that the energy increases indefinitely as the wave-length decreases, which is contrary to experimental evidence. For long wave-length and high temperatures the formula conforms to experiment, however, but it breaks down for small values of λT.

The Rayleigh-Jeans formula for the distribution of energy in the spectrum of black-body radiation indicated a steady increase of energy with decreasing wave-length as indicated by the dotted curve of Fig. 451 whereas the curve as determined by experiment is indicated by the full line, having a maximum at a certain value of λ depending on the temperature of the black body.

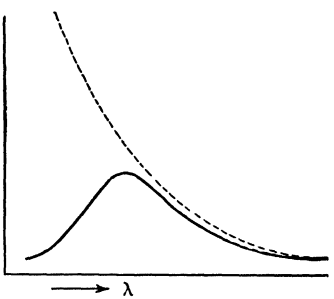


FIG. 451

Planck showed that the difficulty lay in accepting the classical laws of mechanics in the case of atomic phenomena, and abandoning these, laid the foundations of what is now known as the quantum theory. He assumed that the centres of radiation were oscillators of the nature of dipoles, *i.e.* a positive ion with a negative electron, held in an equilibrium position by a quasi-elastic force and capable of vibration to-and-fro along a line. Such a

system is termed a linear oscillator and on the classical theory the amplitude of the vibration may have any value we please within a certain limit.

Planck, however, made the assumption that the amplitude of the vibration could assume only certain discrete values, *i.e.* for frequency ν the energy of a linear oscillator could take only the values o, $h\nu$, $2h\nu \cdots nh\nu$, the energy being the product of the frequency and a constant h, known as Planck's constant. Energy is radiated or absorbed only in definite amounts by each oscillator, corresponding in magnitude to the difference in the energy values of the oscillator as it changes — say — from state, $5h\nu$ to state $4h\nu$. By the introduction of this conception of the vibration of an oscillator, for which there was no physical explanation, Planck was able to deduce a radiation formula for energy distribution in the spectrum which conformed to the results of experiments.

The simplest way of showing how fundamental a difference results from adopting Planck's conception, is by considering the quantum theory as applied to specific heat. Experiments have shown that the specific heat of a polyatomic gas is a function of

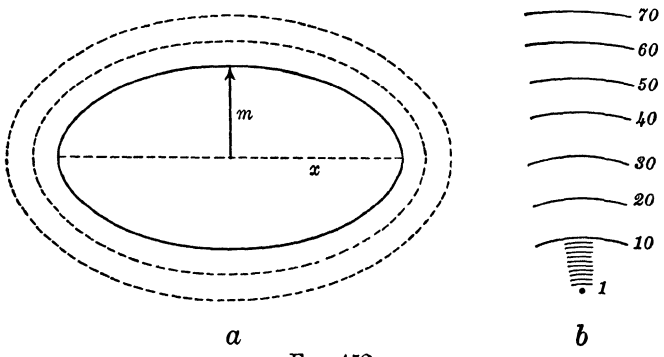


FIG. 452

temperature, being less at low than at high temperatures. There is no explanation of this circumstance on the classical theory, but the theory of Planck explains it fully. Consider the gas to be made up of diatomic molecules, the atoms capable of linear vibrations similar to Planck oscillators. The vibrations of a simple pendulum are governed by the same law, and the energy of the pendulum can be ascertained as follows: If x is the distance of the bob from the equilibrium position, and m its momentum, and we plot m against α we obtain a closed curve, which can be shown to be an ellipse, the area of which multiplied by the frequency giving the energy $E = A\nu$ in which A is the area and ν the frequency.

If now we assume that the pendulum can swing only in discrete amplitudes of say $\sqrt{1}, \sqrt{2}, \sqrt{3}, \ldots \sqrt{n}$ cms. we have the analogy of the Planck oscillators, with energies represented by the series of ellipses of Fig. 452*a*. If we substitute nh for A in the above formula we have Planck's formula $E=nh\nu$, n being an integral number. We can now represent the possible vibrational energies of one of our molecules as indicated in Fig. 452*b*, in which the curves 1, 10, 20, 30, etc., represent portions of the ellipses, corresponding to Planck's permitted amplitudes. On classical theory we could have also all intermediate ellipses as for example 1, 2, 3, 4, 5, $\cdots$ In the case of a gas at low temperature the molecules have only low states of vibrational energy, and we will see what happens when we supply heat to it by mixing with it a monatomic gas at a higher temperature. This method of supplying heat is chosen as we can easily follow the physical processes involved, for the heat energy of a monatomic gas resides entirely in the motion of translation of the atoms.

The only real definition that we can give to temperature is to say that two bodies have the same temperature if there is no exchange of energy (or flow of heat) when they are brought into contact. If energy passes from one to the other we say that the former has the higher temperature.

We will simplify the problem by considering that the molecules of the gas at low temperature have vibration energy only, that is, we will imagine them fixed in space like the atoms in a crystal lattice, but with sufficient room between them to permit the atoms of the monatomic gas (at higher temperature) to penetrate freely between them. We introduce the hot gas and it immediately gives up thermal energy to the stationary system of molecules, until the exchange ceases, when we say that both have the same temperature,[1] the molecular gas having its vibrational energy increased by the impacts with the atoms of the hotter gas, the energy given being equal to the specific heat $\times$ temperature change. The amount of energy given up is proportional to the measure of the specific heat of the molecular gas and we shall see that its calculated value for low temperatures is quite different on the quantum theory from that predicted classically.

On classical assumptions an atom with sufficient energy, colliding with a molecule of energy value 10 in Fig. 452*b* may raise it to energy value 34; but on the quantum theory this energy content is not permitted and the best that the atom can do is to raise the

[1] The term temperature here refers to vibrational energy alone. If the molecules are released from the lattice, a part of this energy will pass over into energy of translation.

molecule to 30. The atom thus gives up less energy to the molecule on the quantum theory than on the classical. The same holds true for all other collisions, and finally thermal equilibrium is reached both gases having the same temperature.

If our molecular gas is at low temperature, the vibrational energy is small, and we will suppose that the greater part of the molecules have energy values of 10 in Fig. 453 in which an allowed quantum transition of 5 is figured instead of 10 as in Fig. 452*b*. The atoms of the hot gas are moving with different velocities (Maxwellian distribution) and some will have sufficient energy to raise the molecules with which they collide to energy value 15. On the classical theory the slower moving atoms would raise molecules to values of 14, 13, 12 and 11, while on the quantum theory they would give up no energy at all, since they must increase the vibrational energy from 10 to 15 or be without influence. We thus see that less thermal energy passes to the colder gas on the quantum than on the classical theory. The state of equilibrium, according to the two theories, is illustrated by Fig. 453 in which the energy conditions of twelve molecules are represented. With quantized energies only three have had their energy increased, while with unquantized, *all* have been affected.

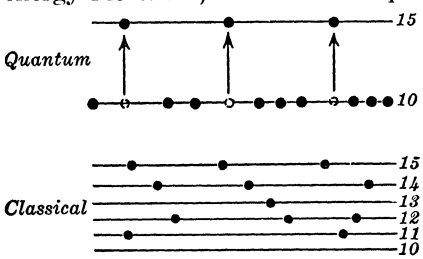


FIG. 453

The mean energy is much higher in the lower than in the upper figure.

If we start with the gas at a higher temperature we must consider that the upper levels are populated, *i.e.* we have molecules in the higher vibration states allowed by the quantum theory, or on the classical theory also in an infinite number of intermediate states. The atomic gas which we introduce must have a higher temperature than our molecular gas and there will consequently be atoms with sufficient energy to raise molecules, say from the ground level to the 34th, but as this level has been excluded on the quantum theory, the 30th only can be reached by collisions with these atoms.

But the percentage difference in energy between a transition 0−30, and 0−34, is much smaller than in the case of the gas at the very low temperature in which we have assumed that the molecules must either remain on the 0 level or be raised to level 10.

The gas is capable of taking up more energy on the classical than on the quantum theory, at all temperatures, but at the

high temperature the difference is relatively less than at the low temperature which means that the specific heat increases with the temperature on the quantum, but not on the classical theory.

We will now take up the application of Planck's quantum theory to temperature radiation. Suppose the walls of the black body to contain the equivalent of harmonic oscillators of all possible frequencies. We mean by this simply that their vibrations radiate "pulses" which can be decomposed by Fourier analysis into an infinite number of frequencies. For this treatment we shall, however, assume oscillators of all frequencies as actually present, as was done by Planck.

On the classical theory each oscillator will radiate waves of a definite frequency, the amplitude dying down with time. An oscillator can start off with a large or small amplitude according to its degree of excitation. If we consider it excited by coming into contact with a hot monatomic gas with Maxwellian distribution of velocities, the condition can be represented by Fig. 454 in which a low frequency vibrator has been excited to large amplitude by an atom of highest velocity, the amplitude dying down as energy is lost by radiation. Any lesser amplitude may be initially excited by the atoms of lower velocity, as indicated by the infinitely close energy levels to the left of the oscillator. In the actual case, the oscillators are to be considered as excited by the thermal agitation of the molecules and atoms.

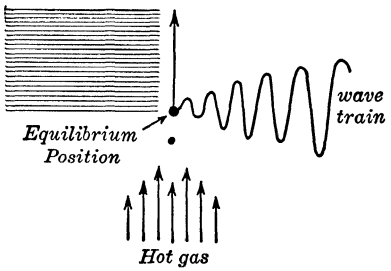


FIG. 454

The spectral distribution of the radiation will obviously depend on the number of degrees of freedom in the ether present for each possible frequency. We have seen that this number increases with the frequency [ν^2 in (1)] but that not all participate as will be shown presently. It is not easy to explain this in words, but perhaps the simplest way is by employing an idea due to Debye regarding the actual physical processes which are going on in the radiator.

Debye has raised objections to all treatments which postulate characteristic vibrations of the separate atoms of a monatomic solid. He regards the actual phenomenon which is responsible for thermal radiation as identical with acoustic vibrations, the body "vibrating" in a fundamental with overtones. He calculates the

number of degrees of freedom in much the same way that the number of frequencies possible in a hollow reflecting cube were calculated by the Rayleigh-Jeans method. The higher frequencies will predominate, but in this case there will be an upper limit to ν while no upper limit can be assigned in the case of the ether if it is regarded as structureless. He showed that the number of degrees of freedom cannot exceed $3N$, N being the number of atoms in one gram atom of the substance. The higher frequencies will predominate since there are more possible combinations giving rise to them, than in the case of low frequencies.

If the ether is structureless and continuous, waves of all lengths from zero to infinity are possible, in other words, it will have an infinite number of degrees of freedom. The number of degrees of freedom of matter on the other hand is finite, owing to its molecular structure, and if we apply the law of equipartition of energy for the equilibrium condition between radiation and matter, we come to the conclusion that all of the energy will exist in ether and none in the matter, in other words, a small mass of a hot substance placed within a perfectly reflecting vacuous enclosure will give up *all* of its energy in the form of radiation to the ether and retain none itself. Jeans has given as an analogy corks bound together by elastics and set into violet vibration on the surface of water in a large basin. The corks will eventually come to rest, having given up their energy to the water, which can be regarded as structureless in comparison to the coarse-grained system of corks and elastics. The energy goes first into the form of waves and finally into heat. This same condition will hold for all of the oscillators of increasing frequency up to the highest.

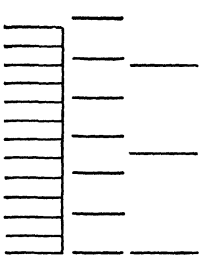


FIG. 455

As Jeans states the case "in all known media there is a tendency for the energy of any systems moving in the medium to be transferred to the medium and ultimately to be found, when a steady state has been reached, in the shortest vibrations of which the medium is capable. This tendency can be shown to be a direct consequence of the Newtonian laws. It is not observed in the crucial phenomenon of radiation; the inference is that the radiation phenomenon is determined by other than the Newtonian laws."

On the quantum theory only discrete amplitudes are permitted as in Fig. 455 and the energy levels will be more widely separated as the frequency increases.

Therefore in a single oscillator there will be, for a given temperature, less energy than called for by the classical theory, if the frequency is high. Above a certain frequency practically no oscillator will be excited since the minimum amount of energy which it can take is too large.

The Nature of White Light. — Present theory regards white light as made up of irregular pulses, which are not however of arbitrary form, since if this were the case, as Lord Rayleigh showed, there would be no distinction between radiation from sources at different temperatures, the character of the disturbance (*i.e.* the shape of the pulse) being fixed by the distribution of energy in the spectrum into which a grating or prism resolves it. The formation of a train of waves from a single pulse by a grating is easy to understand since the secondary pulses which spread out from the lines of the grating form wave-trains in oblique directions of frequency depending on the angle at which they are observed. We have the acoustical analogy in the case of the echo of a sound pulse, made by clapping the hands, from a flight of steps, which we hear as a faint squeak of rather high pitch.

The production of a continuous spectrum by a prism is not as easy to follow. It can be shown that, in a dispersing medium, a pulse suffers periodic inversions, *i.e.* if the pulse enters the medium as a crest it will, as it advances, develop a trough in its rear, the trough increasing in amplitude and the crest decreasing until only the trough remains, the process then repeating itself. It is now clear that if a plane-pulse is incident on the medium in an oblique direction the refracted pulse-front will be corrugated, as the crests and troughs will alternate along parallel lines, since the disturbance along each line has traversed a different distance in the medium from that traversed by the disturbances along the adjacent lines, the pulse-front thus having a structure not unlike the structure that would be impressed on it by a laminary diffraction grating. A full account of the various theories of the nature of white light will be found in the earlier editions of this book in the Chapter on White Light.

CHAPTER XXIV

THE RELATIVE MOTION OF ETHER AND MATTER

Aberration of Light. — The discovery was made by Bradley, in 1728, that the apparent direction of the stars was modified by the motion of the earth through space. To understand just how this results, let us take the case of a gun on shore which has sent its projectile through the hull of a ship. If the ship is at rest, the position of the gun could be determined by sighting through the shot-holes made by the entrance and exit of the ball. If, however, the ship is moving at high speed, it will have advanced a certain distance during the time occupied by the projectile in passing through the hull, and the point of exit will be further aft than in the previous case. A line drawn through the two holes will not, in the present instance, determine the true direction of the gun, as can easily be seen by constructing a diagram. The gun's position, as determined by this method, will appear to have shifted in the direction of the ship's motion, through an angle, the tangent of which is the ratio of the ship's velocity to that of the projectile. This angle is called the angle of aberration. Consider now the case of light-waves entering the object-glass of a telescope. The lens transforms them into concave waves, and we will assume that the telescope is so pointed that they come to a focus on the cross hairs of the eye-piece. If the earth were at rest, a line drawn from the point of intersection of the cross hairs through the centre of the lens would give the true direction of the star. But the earth and the telescope are in motion, and while the waves are travelling down the tube the tube is being carried forward. The focus point will in this case fall a little behind the point at which the rays would have met if the telescope had been at rest, and if the star image is now brought upon the intersection of the cross hairs it is clear that the telescope is pointing a little ahead of the star's true position. The amount of the shift due to the earth's motion can of course be determined only by extending the observations over an entire year; the total change in the star's position will clearly be double the true angle of aberration, for the shift is in opposite directions when the earth is on opposite sides of its orbit around the sun. The case is analogous to that of a ship steaming around in a circle, the crew of which are endeavoring to locate the position of a gun on shore by sighting through the shot-holes.

Bradley found the total angle of aberration to be 40.89 seconds of arc, or that the actual shift due to the earth's motion in its orbit was 20.44 seconds. The velocity of light in space, which was given by dividing the earth's velocity by the tangent of this angle, agreed well with the value found by Römer from observations of the eclipses of Jupiter's satellites.

The phenomenon of aberration apparently indicates that the medium which is transmitting the undulations must be at rest with respect to the telescope. If the ether in the tube were carried along with it, the point at which the waves came to a focus would be wholly uninfluenced by the motion of the tube, and there would be no aberration.

As we shall see presently, however, no experiments have yet shown any evidence of a relative motion between the earth and the ether, all experiments performed with terrestrial sources of light giving results consistent with the hypothesis that the medium in which the undulations are propagated is carried along with the earth in its rush through space.

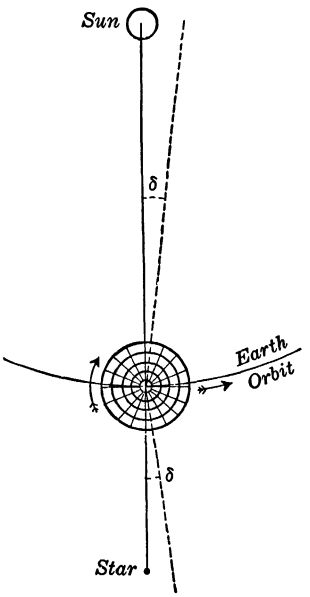


FIG. 456

The Klinkerfuess Experiment. — If the velocity of light could be diminished between the object-glass of the telescope and the cross hair we should expect an increase in the angle of aberration. W. Klinkerfuess in 1867 introduced a tube filled with water between the objective and eye-piece, in an attempt to detect this effect. Instead of observing the position of a star throughout the year he originated a method designed to show the effect with four observations made on two successive days, at a time when the earth's axis is perpendicular to its orbit as shown in Fig. 456. If the earth were at rest in its orbit, the time-interval between transit of the sun and the star, across the meridian, as determined with a meridian circle, would be equal to the half period of the earth's rotation. Owing to orbital motion, the sun and star are seen in the advanced positions indicated by the dotted lines and the difference of time between the two transits will be less than the earth's half period (for clockwise rotation of the earth). If the water tube is interposed, the aberration will be increased and

the time-interval still further lessened. Klinkerfuess thought that he had found an increase of the aberration angle of 7.1 seconds, in agreement with 8 seconds calculated. This was, however, a coincidence of an error of observation with a predicted effect, as all experiments by subsequent observers have shown that the aberration angle is not changed by the introduction of a retarding medium.

Airy's Experiment.[1] — The same question was investigated six years later by G. B. Airy, the Astronomer Royal, with a telescope so constructed that the entire instrument could be filled with water. With this he carried out a long series of observations obtaining the same aberration angle as had been previously found with air-filled telescopes.

To explain this we may assume that the water carries the contained ether along with it, not with its full velocity, for in this case there would be no aberration, but with a velocity sufficient to compensate for the change resulting from the diminished velocity of the light. That something analogous to this dragging along of the ether actually occurs, was proved experimentally by Fizeau, and subsequently by Michelson and Morley.

The Fresnel Entrainment Coefficient and Fizeau's Experiment. — In 1818 Fresnel developed an expression for the phase velocity of light in a medium of refractive index n, moving with a velocity v in the direction of the light. He found that the ether was carried along by the medium, not with its full velocity but with a fractional part $\epsilon = 1 - (1/n^2)$. For water $\epsilon = .434$, that is water in motion imparts to the light traversing it a velocity increment nearly half as great as its own velocity.

Thirty-three years later Fizeau performed his celebrated experiment on the velocity of light in moving media. He arranged an apparatus in which two beams of light were caused to traverse a system of tubes through which water could be forced at a high velocity. A system of interference-fringes was formed by the union of the two beams, and the effect of the motion of the fluid upon the position of the fringes was studied. The arrangement of the apparatus is shown in Fig. 457. Light from a slit at S' after reflection from a plate of glass is made parallel by a collimating lens, and divided into two portions which traverse tubes containing running water. It is clear from the diagram that each interfering beam traverses the same thickness of ponderable medium, for each ray is obliged to pass through the entire tube system. This is accomplished by focussing the rays upon a plane-mirror, the effect of which is to interchange the paths. Moreover,

[1] *Proc. Roy. Soc.*, *20*, 35; *21*, 121, 1871–72.

it will be seen that one ray is travelling always with the current, the other against it. On emerging from the apparatus the rays are brought to a focus at S behind the plate (a portion at S' also), where a system of interference-fringes is formed. A shift of the fringes was observed when the water was put in motion, which could be doubled by reversing the direction of the current. His tubes were 150 cms. long and 5.3 mms. in diameter and the water-velocity 700 cms./sec. The entrainment coefficient calculated was .437 which should have given him a shift of .404 fringe width. He found .46.

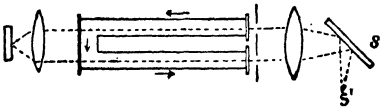


FIG. 457

Let c be the velocity of light in vacuo, v the velocity in water and V the velocity of the water. Assume that the ether is carried along by the water with a velocity $V\theta$, in which θ is a fraction. The velocity of the two interfering beams will be $v-V\theta$ and $v+V\theta$, and if l is the total length of the water path, the difference in time over the two paths will be

$$\frac{l}{v-V\theta}-\frac{l}{v+V\theta}.$$

This experiment was repeated in an improved form by Michelson and Morley.[1] In Fizeau's arrangement the distance between the slits which divide the beam into two portions is necessarily large, and the fringes are in consequence extremely close together and require very high magnification, with its accompanying loss of light. Michelson's arrangement permitted the use of an extended source of light such as a gas flame, and any desired distance between the tubes. Light from a source at S (Fig. 458) is divided at a half-silvered surface at A, and sent around the water-tube system in opposite directions, as shown in the diagram. With tubes six metres long and a velocity of eight metres per second, the displacement observed on reversing the direction of

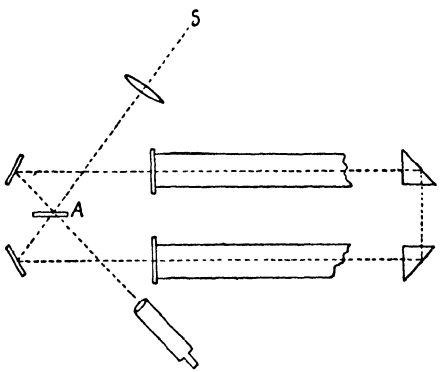


FIG. 458

[1] *Am. Journal of Sci.*, xxxi. p. 377, 1886.

the current amounted to less than the width of a single fringe. The results obtained were fairly concordant, however, the value .434 being found for θ. They also experimented with an air current moving with a velocity of 25 metres per second, but the effect in this case was too small to measure.

The expression for θ which has been given above was developed by Fresnel from the following considerations: He regarded the refractive index as the square root of the ratio of the ether density in the medium to the ether density in vacuo, the refraction being due to the condensation of the ether within the pores of the medium. Consider a transparent plate, of ref. index θ, moving with velocity V, and let D_1 be the ether density within it, and D the density in vacuo. Then $D_1 = \mu^2 D$. If there is no flow of the ether around the edges of the plate, the same amount must enter the front surface in unit-time as leaves the back surface, or $DV = D_1(V - \theta V)$, which gives us, if we substitute $\mu^2 D$ for D_1,

$$\theta = 1 - \mu^{-2} = .438 \text{ for water.}$$

This amounts to saying that the condensed ether within the plate is carried forward with a velocity such that the *excess* of the ether in the body over that in the corresponding free space is carried along with the full velocity of the plate. We may, however, regard the condensed ether as a part of the medium, in which case we can say that the ether proper is entirely uninfluenced by the motion of the medium. If c is the velocity of light in a vacuum, and c/μ the velocity in a medium at rest, the absolute velocity of the light in a medium moving with a velocity V is

$$\frac{c}{\mu} \pm V(1 - \mu^{-2}),$$

the plus or minus sign being used according as the light travels in the same or opposite direction as that in which the medium moves. Commenting on the derivation of the above formula, Lord Rayleigh remarks: "Whatever may be thought of the means by which it is obtained, it is not a little remarkable that this formula and no other is consistent with the facts of terrestrial refraction if we once admit that the ether in the atmosphere is at absolute rest. It is not probable that the ether in moving refracting bodies can properly be regarded as itself in motion, but if we knew more about the matter we might come to see that the objection is verbal rather than real. Perhaps the following illustration may assist the imagination:

"Compare the ether in vacuum to a stretched string, the transverse vibrations of which represent light. If the string is loaded

(say with beads) the velocity of propagation is diminished. This represents the passage of light through stationary refracting media. If now the loads be imagined to run along the string with a velocity not insensible in comparison with that of the waves, the velocity of the latter is modified. It appears that the suggested model would lead to a somewhat different law of velocity from that of Fresnel; but in bringing it forward the object is merely to show that we need not interpret Fresnel's language too literally."

Lorentz's Treatment. — The subject of the entrainment coefficient was investigated in 1895 by H. A. Lorentz who developed the formula on the electron theory and showed that μ in the c/μ of Fresnel's treatment should have the value corresponding to the frequency of the light relative to the moving medium.

Zeeman's Experiments. — An experimental proof of the Lorentz formula was undertaken by Zeeman in 1914–16 employing the Michelson method, but using monochromatic light, photographic registration, and improved methods of determining the water velocity. For the green mercury line he obtained for the entrainment coefficient the value .451 (.454 calculated) in contrast to the value .439 calculated from the Fresnel formula.

Zeeman also made experimental determinations with rods of quartz and glass 1 metre in length, oscillated at high frequency in a direction parallel to the long axis, the light being transmitted only for brief intervals corresponding to the moments of maximum velocity (10 metres per sec.).

The Michelson-Morley Experiment. — The history of this experiment, perhaps the most celebrated ever performed, since it shook, to its very foundations, the wave-theory of light, and gave rise to discussions which are not yet finished, is as follows. Clerk Maxwell, in a letter to Michelson in 1880 expressed the belief that the motion of the earth through the ether would so alter the relative velocity of light with respect to the various parts of optical apparatus as to give rise to effects, too small however to be detected experimentally. Michelson in his reply stated his belief that the effects could be found with suitably designed optical instruments and at once commenced work on the subject. These expected effects depend upon the square of the ratio of the velocity of the earth in its orbit to the velocity of light, a term which can be neglected in all experiments involving such small velocities as occur in experiments such as that of Fizeau. The theory of these celebrated experiments, about which so much discussion has occurred, is as follows.[1]

Consider a system of interference fringes formed by a Michelson

[1] Michelson and Morley, *Phil. Mag.*, xxiv, 449.

interferometer, the three mirrors of which occupy the positions A, B, C (Fig. 459) at the moment when the incident beam SA strikes the first plate. While the light is travelling from the mirror A to the mirrors B and C and back again to A, assume the whole apparatus carried forward by the earth in the direction of the incident light to the position $A'B'C'$. The ray reflected from B, which interferes with a given ray reflected from C, along the line $A'D'$, we must consider as travelling along $AB'A'$, the angle BAB' being equal to the angle of aberration. It must not be thought, however, that the path of the reflected ray is altered by the motion of the mirror. The change of path merely indicates that the ray which we are utilizing, and which strikes the mirror A in its second position at the point where the ray BA would have met it had the apparatus been at rest, is a ray reflected at the angle indicated The mirror B, at the moment when reflection occurs at its surface, has moved only one-half of the distance between A and A', from which it follows that the angle BAB' is equal to the angle of aberration, the tangent of which is the ratio of $BB'/2$ to AB. The dotted lines in the figure are not quite correctly placed, as is obvious.

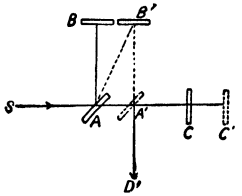


FIG. 459

Suppose that the ether remains absolutely at rest, and let c = the velocity of light, and u = the velocity of the apparatus, *i.e.* of the earth in its orbit.

Further, let T = the time occupied by the ray in passing from A to C (located at the point to which it has been carried), and T' = the time in returning from C to A'. At the moment of reflection from C the mirror will occupy a position midway between C and C'.

Call D the distance AB or AC, then

$$cT = D + uT,$$
$$cT' = D - uT',$$

so that

$$T = \frac{D}{c-u}, \quad T' = \frac{D}{c+u},$$

the whole time being given by

$$T + T' = 2D\frac{c}{c^2-u^2}.$$

If the system is at rest, $T = 2D/c$, which is identical with the above equation when $u = 0$.

The actual distance traversed in the time $T+T'$ is obviously given by multiplying the time of transit by the velocity c, or

$$\text{Path} = 2D\frac{c^2}{c^2-u^2} = 2D\left(1+\frac{u^2}{c^2}\right),$$

which we obtain by simple division, neglecting u^4/c^4 and terms of higher order.

The length of the other path $AB'A'$ is

$$2D\left(1+\frac{u^2}{c^2}\right)^{\frac{1}{2}} = 2D\left(1+\frac{u^2}{2c^2}\right) \text{ (approx.), since } \frac{AA'}{AB} = \frac{2u}{c}.$$

It is thus seen that the effect of the motion of the apparatus is to slightly increase both paths, the increment being twice as great, however, along the path parallel to the earth's motion. The path-difference which was originally zero is now

$$2D\left(1+\frac{u^2}{c^2}\right) - 2D\left(1+\frac{u^2}{2c^2}\right) = D\frac{u^2}{c^2}.$$

We may, if we choose, look at the thing from a slightly different point of view. Imagine an observer moving through space with a velocity v in the direction of the x axis, Fig. 460. Suppose that he sets off a flash of light which originates a spherical wave, which moves out from its origin with a constant velocity in all directions. As the wave spreads out, the observer is moving along towards the right, drifting away from the point where the wave originated, and reaching the point B at the moment when the wave has the radius c. It is clear that if he were unaware of his own motion, but had some means of following the motion of the wave, he would come to the conclusion that the velocity of the wave in the direction of $+x$ is $c-v$, and in the reverse direction $c+v$. In the direction

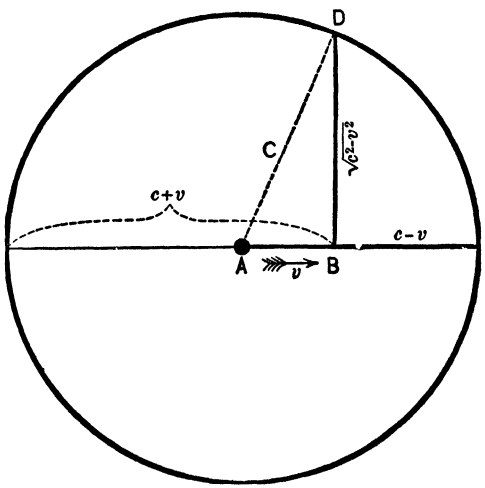


FIG. 460

perpendicular to the x axis he will also find the velocity altered, for he will, as he moves along, be dealing with the velocity with which that portion of the spherical wave which lies in the direction $\perp$ to x from his viewpoint is moving away from him. It is clear from the figure, that when he is at the point B he will regard the distance BD as the distance which the wave has travelled, and this distance is less than c, being in fact $\sqrt{c^2-v^2}$, as can be seen from the right triangle ADB. We can now imagine the observer furnished with an interferometer, and if he uses these new values for his light velocity in the two directions he can consider the two paths in his instrument as constant. If D is, as before, the distance between the mirrors, the time occupied by the light in making its to-and-fro excursion between the half-silvered mirror (which can be regarded as the source of light) and the mirror lying in the direction of motion will be

$$\frac{D}{c-v}+\frac{D}{c+v}=\frac{2D}{c}\left(1+\frac{v^2}{c^2}+\cdots\right),$$

while the time over the other path will be

$$\frac{D}{\sqrt{c^2-v^2}}+\frac{D}{\sqrt{c^2-v^2}}=\frac{2D}{c}\left(1+\frac{v^2}{2c^2}+\cdots\right).$$

The time-difference over the two paths is obviously $(D/c)(v^2/c^2)$ and the path-difference $D(v^2/c^2)$ which is identical with our former expression.

If now we rotate the whole apparatus through 90°, the path $AB'A'$ will be the one which receives the larger increment, and a shift in the position of the fringes should result.

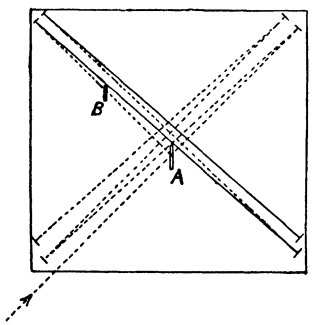


FIG. 461

In the first experiments tried, the expected shift amounted to only about $\frac{1}{20}$ of the distance between the fringes; moreover, it was found impossible to rotate the apparatus without introducing strains, which caused slight changes in the position of the fringes. As a result no very definite conclusions could be drawn from the observations. The experiment was then repeated with improved apparatus. By means of multiple reflections the path D was increased to 11 metres. The mirrors, 16 in number, were mounted on a heavy slab of stone which was floated on mercury. The ap-

paratus was kept in slow rotation while the observations were taken, which did away with the strains which always occurred when it was brought to rest. A diagram of the apparatus is shown in Fig. 461, the number of mirrors having been reduced by one-half, however. The beam of light is divided at the half-silvered plate A. The compensating plate is located at B.

The value of u^2/c^2 is 10^{-8}, if the earth's orbital motion is alone considered, while D measured in wave-lengths of sodium light was 2×10^7. If the ether remains at rest relatively to the earth, we should expect a displacement of the fringes equal to

$$4\times10^7>10^{-8}=.4 \text{ of a fringe width.}$$

The actual displacement observed was certainly less than $\frac{1}{20}$ of the expected, and probably less than $\frac{1}{40}$, from which the conclusion was drawn that since the displacement is proportional to the square of the velocity, the relative velocity of the earth and the ether is probably less than one-sixth of the earth's orbital velocity. This amounts to saying that the earth drags the ether in its vicinity along with it, a circumstance which cannot be reconciled with the phenomenon of stellar aberration, to account for which we must assume the ether at rest with respect to the earth.

An explanation of the absence of any fringe-shift in the experiment of Michelson and Morley was suggested simultaneously by Fitzgerald and Lorentz. This explanation was based upon an assumed change in the linear dimensions of matter resulting from its motion through the ether; a contraction of the base upon which the mirrors are supported occurring in the direction of motion would compensate for the increment of optical path due to the motion of the apparatus.

Attempts have been made to detect this hypothetical effect, but thus far all have been unsuccessful. If the effect occurs it might very well happen that its magnitude would vary with different materials. Morley and Miller therefore repeated the experiment under conditions such that the distance between the mirrors could be made to depend upon the length either of a metal rod or a pine stick. The result was the same in each case, however.

It occurred to Lord Rayleigh that the contraction in the direction of motion, if it existed, might give rise to double refraction, but he was unable to detect anything of the kind; and a subsequent experiment by Brace, performed with the greatest care, has established conclusively that no trace of double refraction occurs as a result of the motion of transparent media through the ether.

It is obvious that the failure of the earth's motion to influence in any way phenomena occurring in optical systems located wholly

upon the earth are at once explained if we assume that the velocity of light is influenced by the motion of the source, *i.e.* if the velocities are additive, as would be the case on the corpuscular theory, for in this case the moving observer would remain at the centre of the wave. This hypothesis was not seriously considered, for it is incompatible with the wave-theory.

Kennedy used an improved apparatus, in which one-half of the reflecting surface of one mirror was raised $\frac{1}{20}$ of a wave-length above the other, the two halves meeting along a sharp line, to which the fringes were made perpendicular.

With this instrument Kennedy was able to detect a shift as small as .004 of a fringe-width, while the expected shift, according to Miller's results, if occurring in his instrument would have been .016. No trace whatever of a shift was found.

Illingworth [1] employing Kennedy's method reached an accuracy which would have shown an ether-wind of 1 kilometre a second. Neither in Pasadena or at the top of Mt. Wilson was any shift observed.

The last and most elaborate experiment was made by G. Joos.[2] The rectangular frame carrying the mirrors was built of fused quartz and the entire system mounted in a vacuum. Continuous photographic registration of the fringes was secured automatically while the entire apparatus was in slow rotation. There was no evidence of a shift of the fringes due to the earth's motion in space. The total length of the light path was 20 metres and an "ether-wind" of less than 1.5 kms./sec. could have been detected if it existed, according to his calculations.

Miller's Experiments. — The most exhaustive series of observations extending over a period of thirty years have been made by D. C. Miller. The experiments were made with an instrument in which the total light path was 64 metres, in various localities and under different conditions. Plotting the small fringe shifts observed from day to day over a period of years in the form of a curve and analyzing the curve with an harmonic analyzer, he computed the velocity and direction of the earth's absolute motion in space, on the assumption that the observed effects were real. Astronomical observations indicate that the solar system is moving with a velocity of 19 kms./sec. with respect to the brighter stars towards the constellation Hercules. Miller's results showed an absolute motion in the *opposite* direction of 208 kms./sec., which, if real, would mean that the nearer stars are moving in the same direction with a velocity of 230 kms.

[1] *Phys. Rev., 30,* 692, 1927.
[2] *Ann. der Phys. 7.* 385. 1930, 1931.

His fringe shifts amount to only about $\frac{1}{20}$ of that to be expected for a stagnant ether, but he was able to deduce from them not only the part of the earth's motion due to the motion of the solar system but also the much smaller part due to its orbital velocity. A very full and convincing account of these laborious observations and calculations will be found in *Review of Modern Physics*, Vol. 5, No. 3, July, 1933.

Lodge's Experiment. — The experiment of Michelson and Morley indicating that the earth drags the adjacent ether along with it, it occurred to Lodge [1] to investigate directly the effect of moving matter upon the ether. Two steel disks were mounted side by side and close together upon a common axle, and two interfering beams of light were passed in opposite directions around the annular space between the disks, by means of a system of mirrors. The disks could be rotated at a high speed, and if they dragged the ether wholly or in part the effect would be noticeable in a shift of the interference-fringes, since one beam of light is travelling in the direction of rotation, the other in the opposite direction. No effect was observed at even the highest possible speeds. Thinking that perhaps the mass of the moving matter entered as a factor, Lodge substituted for the disks an immense spheroid of iron weighing half a ton, provided with a narrow circular crevasse along its equator, around which the luminous beams were reflected. The spheroid could be magnetized by means of a coil of wire, since it appeared possible that magnetization of the moving medium might have some effect. As in the previous case, the results were all negative, proving that at such speeds as can be handled in the laboratory the ether remains practically at rest. This is in agreement with all of the other experiments except the one performed by Michelson and Morley.

Influence of the Earth's Motion on Rotatory Polarization. — Lorentz developed a formula which apparently indicated that a change of one part in ten thousand in the rotation of the plane of polarization by active substances such as quartz was to be expected when the polarimeter, set parallel to the earth's orbit, was turned through 180°. Larmor, in his *Aether and Matter*, criticised this result, and concluded that no effect was to be expected. Lorentz [2] replied, defending his position, and maintained that Larmor was in error. The subject was then attacked experimentally by Lord Rayleigh [3] who found that the change, if it occurred, was less than 1/100,000 part of the total rotation. He used the five

[1] *Trans. Roy. Soc.*, 189, 1897.
[2] *Proc. Amsterdam Acad.*, May, 1902.
[3] *Phil. Mag.*, 4, 215, 1902.

quartz blocks which had been prepared for Tait's rotation spectroscope, each block 5 cms. thick, the battery producing a rotation of over 4000 degrees for sodium light. As the difference in the rotations for D_1 and D_2 amounted to 11°, it was impossible to secure complete extinction with sodium light, and the helium tube was consequently employed, which gave an abundance of yellow monochromatic light. The apparatus was mounted on a horizontal stand, which could be rotated on a pivot. No change whatever was observed, however, which was in agreement with the predictions of Larmor. Still more recently Bruce [1] has shown that any change must be less than 1/10,000,000 of the whole.

Effect of the Earth's Motion upon the Intensity of Terrestrial Sources. — Fizeau came to the conclusion that if the ether was at rest with respect to the earth, the earth's orbital motion ought to affect the intensity of the light emitted by terrestrial sources, the light emitted in the direction of the earth's motion being less intense than that emitted in the opposite direction. This conclusion was reached by the following reasoning: Let A be a lamp (Fig. 462), B and C two screens, upon which the light falls, the whole apparatus moving with the earth through a stationary ether in the direction of the arrow, with a velocity v. Let the screens be at distance s from the lamp. To reach B the light has to traverse not the distance s, but the slightly smaller distance $sc/(c+v)$ (in which c is the velocity of light) since the screen is approaching the source during the passage of the light. To reach the screen C the distance traversed is $sc/(c-v)$. The intensity being inversely as the square of the distance, it follows that if J_0 is the intensity at each screen when the apparatus is at rest, or turned perpendicular to the direction of the earth's motion, the intensity when the apparatus is flying through the ether in the direction indicated will be

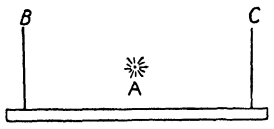


FIG. 462

$$J = J_0\left(1 \pm \frac{2v}{c}\right) = J_0\left(1 \pm \frac{1}{5000}\right).$$

Fizeau proposed measuring the total intensity of the radiation at two points equidistant from a lamp by means of a pair of thermo-elements opposed to each other, so as to secure compensation and no current. On rotating the apparatus through 180°, a feeble current would result from the slight change of temperature at the two points, due to the exchange of the slightly different intensities

[1] *Phil. Mag.*, September, 1905.

of radiation. Strangely enough, the experiment was not tried until 1903, when Nordmeyer,[1] working in Kayser's laboratory at Bonn, conducted a very careful series of experiments embodying the ideas of Fizeau.

The question was carefully considered from a theoretical standpoint by Bucherer, working in the same laboratory, his results appearing in a paper immediately preceding the one referred to above.

There are a number of points which must be carefully considered. A change of amplitude results from the difference of path, as shown above; in addition to this we have, however, a change in the wavelength of the emitted light due to Döppler's principle, and, as we have seen in the Chapter on Radiation Laws, a decrease of wavelength, the amplitude remaining the same, means an increase in the energy in unit volume of the medium. Moreover, if the intensity is measured by the thermopile, it is represented by the energy absorbed by the blackened surface of this instrument. If the instrument is moving against the light-waves, work must be done owing to the pressure of the radiation, the equivalent of this work appearing as heat in the body receiving the radiation. Bucherer considers all these questions carefully, and comes to the conclusion that Maxwell's theory leads to the conclusion that the intensity is increased by the amount $1+(2v/c)$, or in the same amount as found by the more elementary treatment given above. This is a first order effect, while all other experiments pertaining to the relative motion of the earth and the ether depend upon v^2/c^2, or on second-order effects. Taking into account the pressure of radiation, however, and calculating the amount of energy absorbed by a moving screen, he finds that the common motion of the screen and source is without influence upon the intensity as measured by the heating of the absorbing surface. This result is obtained by first calculating the amount of energy absorbed by a screen moving in the direction in which the light is travelling. The relative velocity is then $c-v$, and the energy absorbed will be the $(c-v)/x$ part of that which would be absorbed if the body were at rest. From this is to be subtracted the heat equivalent of the work done by the pressure of the radiation, which is equal to the product of the pressure and the distance moved in unit time. This work can be done only at the expense of the absorbed energy.

Lorentz has also shown in a different manner that the earth's motion is without influence upon terrestrial sources of light.

The experimental investigation was conducted by Nordmeyer in the manner proposed originally by Fizeau, but with all the refine-

[1] *Ann. der Phys.*, *11*, 284, 1903.

ments employed at the present time. The results were purely negative, and it was established that the intensity was not changed by one part in 300,000 by the rotation of the apparatus, thus confirming the conclusions arrived at by Lorentz and Bucherer.

The experiment cannot, however, be regarded as proving that the ether is at rest with respect to the earth, since the same results are to be expected on the hypothesis of a moving ether. It is worthy of remark, however, that every optical experiment, with the exception of the one performed by Michelson and Morley, is in accord with the hypothesis of a stationary ether.

The Principle of Relativity. — So far as the testimony of all experiments which have been tried up to the present time goes, the motion of the earth does not affect optical phenomena exhibited with terrestrial sources of light. So far as we have been able to find, the measured velocities of light parallel and perpendicular to the direction of the earth's motion are the same. It is true that we have never made independent observations in both directions, though it is highly probable that even the results of such determinations would be open to question, but the Michelson-Morley experiment has shown, beyond any doubt, that we must remodel our views regarding the stationary ether assumed by Lorentz. The Theory of Relativity is an attempt in this direction. It was first definitely formulated in its entirety by Einstein, though the work of Lorentz may be regarded as having paved the way for it. It has been placed on a substantial mathematical basis by Minkowski, and may be regarded as partially proved by the experiments of Kaufmann and of Bucherer upon the mass of the moving electron, and the negative results of all ether-drift experiments. The Fitzgerald-Lorentz hypothesis of a contraction of matter in the direction of its motion, which was formulated as a sort of loophole of escape from the difficulty arising from the Michelson-Morley experiment, will be seen to be the logical outcome of the much more general postulates established by Einstein, which denies at the outset the possibility of ever measuring or even detecting absolute motion through space.

The Theory of Relativity starts out with two postulates.

The first of these states that the uniform motion of translation cannot be measured or even detected by an observer stationed on the moving system from observations confined to the system. This amounts to saying that motion through the ether (if the ether exists at all) will be wholly without influence upon all optical experiments made with terrestrial sources of light.

The second postulate is that the velocity of light in space is a

constant, independent of the relative velocity of the source and the observer.

Combined with the first postulate this leads us to the extraordinary conclusion that the *measured* velocity with which light passes a moving observer is the same as for a stationary observer. Going back now to the conception which we made use of in the treatment of the Michelson-Morley experiment, of a moving observer who sets off a flash of light, and then drifts away from the centre of the wave, we see at once that, on the theory of relativity, so far as any measurements which he could make, even with imaginary apparatus, are concerned, he will find himself always at the centre of the wave. The Chapter on the Principle of Relativity, which was given in the second edition of this book is omitted for lack of space.

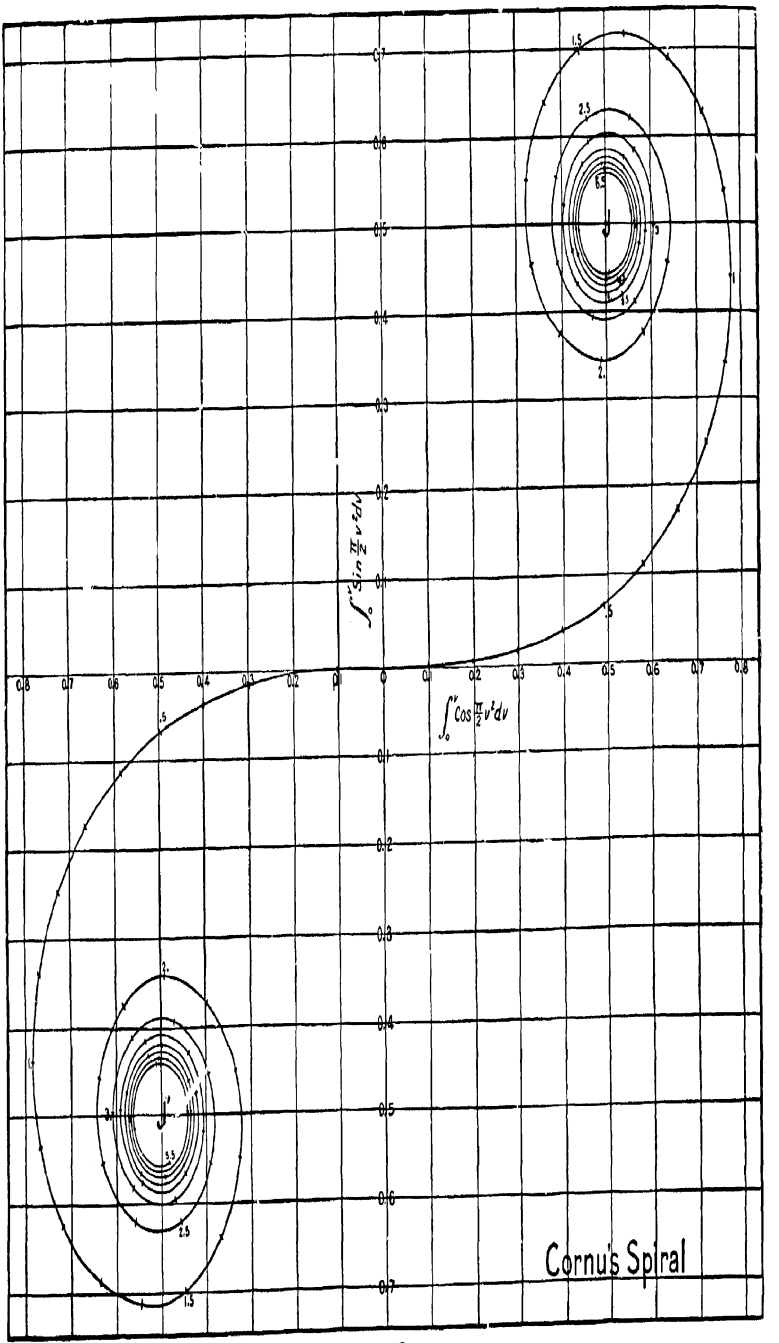

Cornu's Spiral
$\int_0^v \sin \frac{\pi}{2} v'^2 dv$
$\int_0^v \cos \frac{\pi}{2} v'^2 dv$

PLATE 17.

APPENDIX

Plane-Polarized Vibrations in Natural Light. — The experiment on the rotation of plane-polarized vibrations in ordinary or unpolarized light, described on page 362, was inserted after reading the manuscript of the paper, which appears in the *J. O. S. A.* early in 1934 and not 1933 as stated in the footnote. It was, however, impossible to make room for the insertion of the further note that a similar experiment was performed by Stefan in 1865,[1] which had been overlooked by the authors. He, however, used a single plate of quartz cut perpendicular to the axis, which gave a rotation of 90° for wave-lengths in the orange. Covering one-half of the objective of a three-prism spectrograph with the plate, and illuminating the instrument with white light, he found some thousand or more dark lines in the continuous spectrum. These were obviously Talbot's bands, discovered in 1813, though he does not so specify. The important point was that the bands were wholly absent in the orange and faint in the adjoining regions, due to the failure of interference between the rotated and unrotated plane polarized constituents of the original light.

It should be noted that Talbot's bands can be seen with fairly thick plates if a spectroscope of high dispersion and resolving power is employed. The "best thickness" is that which retards the advanced portion of the wave train, coming say from a grating, by an amount sufficient to cause it to advance alongside of the other half, as pointed out by Schuster, and more fully treated in the earlier editions of this book. For a grating of N lines this retardation will be $1/2\ N\lambda$.

Absent Spectra. — There are two cases emphasized by Lord Rayleigh in which only spectra of the first order appear, the first being with a grating acting by opacity, which varies as a sine function.

In the direction of the second order spectrum the path difference between disturbances coming from adjacent elements of the grating is 2λ, and in this direction each element of the grating will produce zero illumination as can be shown by the following graphical method. Consider the grating element subdivided into 38 narrow strips, radiating disturbances of intensities varying as the sines of 90°, 85°, 80° . . . 0° . . . 80°, 85°, and 90°. If we plot the first half of these (which represent disturbances from the most transparent to the most opaque portion) as vectors with a constant phase advance of 360°/19 or 19° (nearly), we obtain a curve somewhat similar in form to the first complete turn of the upper half of Cornu's spiral, except that the terminal point is over the origin. Completion of the curve is

[1] *Sitzungsber. d. Wien Akad. L.*, Abth. II, 380, 1865.

unnecessary as it is obvious that the second set of progressively *increasing* vectors will bring the curve around to the starting point, which means zero illumination. A similar state of affairs exists if the phase varies along the wave-front as a sine function, the amplitude being constant, which condition could be obtained with a reflecting or refracting surface of a wave-form. Interesting experiments verifying the first condition were described by A. B. Porter who made a contact print on glass of a grating of 400 very sharply ruled black lines to the inch, one edge being in contact with the photographic plate while the other was separated from it by a strip of paper causing the lines to have blurred edges. The original grating showed 35 spectra of the sodium flame, while the copy showed only three when the flame was viewed through the edge which was in contact with the grating, and only one when it was viewed through the other edge.

The amplitude of the light transmitted by opaque bars and transparent intervals is represented by a square-topped curve, which by Fourier's theorem may be represented by an infinite series of cosine terms. If the edges of the opaque bars are not sharply defined, *i.e.*, if they are shaded slightly, the analysis shows that the higher harmonic terms in the series are absent, and since each one of these gives a spectrum, the higher orders of spectra are absent. Applying this principle to Abbe's theory of microscopic vision, we see at once that the sharpness of the edges of the images of a series of black lines depends upon the transmission of the spectra of high orders. If only the first order spectra were passed he found that the lines appeared greatly blurred. If four or five orders were passed, the images became sharper and less blurred, but a fine dark line appeared down the centre of each. This was predicted by Porter from curves drawn representing the summation of the first five terms of the Fourier series, and subsequently verified by experiment. The result is rather remarkable in showing that a falsification of the image may result from an improvement of the lens. See Porter's article [1] or page 225 of the last edition of this book for further applications of this principle to microscopic vision.

The Form of the White Light Pulse. — An interesting paper by Jaumann [2] has appeared recently in which the form of the pulse (which Lord Rayleigh had computed as having roughly the shape of the probability curve) is deduced by integrating a continuum of harmonic vibrations the amplitudes of which are as the square roots of the spectral intensities, as given by Planck's formula for the radiation of a black body.

He obtained a curve which was not strictly aperiodic, but one in which the sign changed twice or three times according to two arbitrary phase relations adopted. The general form of the curve in the first case was somewhat similar in form to the dispersion curve running through a narrow absorption line (sign changing twice) and in the second case to the intensity

[1] *Phil. Mag.*, 1905.
[2] *Zeit. f. Phys.*, *72*, 700, 1931.

curve of a diffraction maximum bordered by two faint minima (sign changing three times). He discussed the formation of a spectrum by a prism and the interference of white light, pointing out the fact that the periodic inversion of the pulse by the prism, as treated by Schuster, gives modulation of sine form and that spectra of higher order do not appear.

There appears, however, to be a point which has been overlooked or at least not emphasized in previous treatments, namely that the periodic structure produced by the pulse inversion is not fixed in space along the face of the prism (as it is in the case of a grating) but is moving in a direction perpendicular to the lines of constant pulse form, with a velocity represented by the trace of the oblique wave-front along the face of the prism. It is due to this circumstance that the periodic structure produces a single spectrum only, instead of two, as with a grating in which case the periodic structure is stationary. This point will be more fully brought out in a paper now in preparation in collaboration with K. Herzfeld.

INDEX

Abbe, diffraction theory of microscopic vision, 277.
Aberration in image formation, longitudinal, 59.
spherical, 80.
Aberration of light, 812.
Abraham and Lemoine method, 739.
Absent orders of spectra with grating, 247, 829 (Appendix).
with echelon, 315.
Absorption, by alkali metals, 560.
Bohr's theory of, 144.
filters, 14.
laws of, 101.
of light, 99.
by porous surfaces, 106.
by rock-salt, 530.
by sodium vapor, 146.
Absorption and reflection curves in infra-red, 515.
Absorption bands, density and pressure effects, 527.
electronic, 153.
Fortrat diagram, 155–156.
of HCl, 152.
infra-red, 149.
infra-red of crystals, 515.
photographs of, 100.
vibration-rotation, 150.
Achromatic, lens, 112.
prism, 110.
Achromatism, 109.
by grating, 111.
Achromatization by grating, 111.
Airy, water-telescope, 814.
Alkali metals, optical properties, 558.
spectra of, 139.
Amici principle, 82.
Amplitude and intensity, 363.
Anderson, J. A., on emission of rare earths, 788.
Anharmonic oscillator, 150.
Anomalous dispersion, 118.
bearing on solar phenomena, 121.
of cyanine, 120.
laboratory demonstration, 122.
laboratory flash-spectrum, 122.
of sodium, 492–500.
Anomalous wave propagation, 288.
Anti-Stokes lines, 445.
Aperture, image formation by, 272.
circular, diffraction by, 224, 266.
Argon, dispersion of, 490.
Astigmatism of concave grating, 261 264. Also Plate 5, 262.
Astronomical application of interference, 176.
Atom of Thomson, Rutherford, Bohr, 124.
Atomic hydrogen, 126.
Atoms, fluorescence of, 610.
core of, 142.
resonance radiation of, 587–593.

Babinet compensator, 356.
Babinet's principle, 268.
Balmer, hydrogen spectrum, 127.
Bands (absorption), magneto-polarization effects in vanishing fields, 726.
magnetic rotation spectra, 729.
missing, 635–636, 674–675.
sodium vapor absorption, 616.
Bands, electronic, 153.
Barnes, absorption of quartz to, 130 μ, 522.
Becquerel, J., magnetic rotation in crystals, 715.
Beer's law, 101.
Belopolsky, Döppler principle, 26.
Benzene, absorption and fluorescence, 651.
Biaxial crystal, 367.
Billet's split lens, 187.
Bjerrum, theory of molecular spectra, 147.
"Black-body" radiation, 791.
Black surfaces, 106.
Blue-sky, experiments on, 426.
Body color, 103.
Bohr, theory of absorption, 144.
theory of atomic radiation, 129.
theory of resonance radiation, 596.
Boltzman, law of radiation, 798.
Born and Oseen, rotatory polarization, 581.

Boundary conditions, 403.
Boys, C. V., rain-bow cup, 191.
Brace, circular components in magnetic rotation, 698.
Brackett, infra-red H_2 series, 133.
Breit and Ellett, Larmor precession in A.C. fields, 723.

Cabannes, light scattering, 426, 433.
Canal ray velocity, 17.
Cario, sensitized fluorescence, 612.
Cauchy dispersion formula, 469.
Caustics, 53, 57, 59.
  interference fringes bordering, 187.
Chlorate of potash colored crystals, 198.
Christiansen, anomalous dispersion, 119.
  colors, 115.
Cinematograph diagrams of wave-reflections, 57.
Circular light, 352.
Colloids, scattering by, 440.
Color, body and surface, 103.
  Christiansen's colors, 115.
  subjective, 12.
  and wave-length, 12.
Complex refractive index, 488.
Compound interferometer, 315.
Compton, proof of Einstein's law of photo-electric effect, 764.
Conical refraction, 380–381.
Core of atom, 142.
Cornu prism, 574.
Cornu's spiral, 229.
Cotton, Zeeman effect without spectroscope, 671.
Cotton-Mouton effect, 718.
Critical angle of reflection, 66.
Cuprous oxide photo-cell, 771.
Curved light rays, 90.
Cuthbertson, absorption by argon, 490.
Cyanine, dispersion of, 508.
Czerny, absorption by rock salt, 530.
  dispersion of rock salt, 529–531.
  infra-red bands, 149.
  polarized infra-red emission, 785.

Damping factor in dispersion, 487.
de Broglie, photo-electrical effect with X-rays, 768.
Degenerate vibrations, 461.
de Groot, resonance radiation of neon, 609.
Dichromatism, 104.
Dieke, Raman lines of HCl, 452.
  astigmatism of grating, Plate 5, 262.
Diffraction (*see also* Diffraction grating).
  any number of slits, 242.
  circular aperture, 266.
  circular disc and aperture, 224.
  colors of mixed plates, 279.
  diffraction in relation to images, 271.
  early experiments on, 219.
  echelon grating, 284.
  echelon transformed to prism, 287.
  Fraunhofer class of fringes, 232.
  Fresnel class of fringes, 219.
  graphical solution of problems, 225, 242.
  halos, 269.
  and image formation, 272, 277, 278.
  mathematical treatment, 237.
  narrow-slit, 231, 233, 239.
  narrow-wire, 231.
  particles on mirrors, 270.
  rectangular aperture, 238.
  by straight edge, 220–230.
  Talbot's bands, 280.
  transition from Fresnel to Fraunhofer class, 235.
  by two apertures, 268.
  two parallel slits, 232, 235, 240.
Diffraction grating, absent spectra, 247, 315, 829 (Appendix).
  astigmatism of concave, 261, 264, Plate 5, 262.
  comparison spectrum with concave, 262.
  concave, 260.
  crossed gratings, 259, 325.
  Eagle's mounting, 263.
  echelette for infra-red, 265.
  echelon grating, 284–306.
  errors of ruling: "ghosts," 254.
  focal properties of, 253.
  "ghosts," measuring intensity of, 257.
  graphical method for, 242.
  gratings with predominant spectrum, 264.
  intensifying glass, 253.
  intensity distribution, 247.
  laminary grating, 248.
  number of spectra, 247.
  overlapping spectra, 250.

Diffraction grating, Paschen mounting for concave, 262.
plane grating, 249.
reflecting echelon, 307.
resolving power, 251.
spectra with mercury waves, 245.
testing resolving power and focus, 262.
Dipoles, 483.
Dispersion, anomalous, 118.
of argon, 490.
curves of, 108, 471, 473.
curves of cyanine, 508.
dipoles, natural and induced, 483.
of hydrogen, 491.
of hydrogen electrically excited, 503.
of light, 106.
metal vapors, 492.
Newton's investigations of, 107.
of nitroso-dimethyl aniline, 506.
of quartz (table), 520.
of rock-salt, 525, 529, 532.
of sodium in ultra-violet, 502.
within absorption band, 507.
Dispersion theory, complex refractive index, 488.
damping factor, 487.
dielectric constant, 484.
dipoles, natural and induced, 484.
Drude formula, 486.
effect of density and pressure on absorption bands, 527.
experimental proof of, 489.
"Hacken-Methode," 500.
molecular scattering and dispersion, 476.
radiation damping, 479.
synthetic waves, 480.
Displacement currents in free ether, 400.
Döppler effect in laboratory, 26.
Döppler-Fizeau principle, 23, 28.
Dorfmann magnetic rotation, 713.
Double image prisms, 338.
Double refraction, axes of single ray velocity, 381.
biaxial crystals, 367.
colors of thin crystal plates, 382.
conical refraction, 380, 381.
convergent and divergent polarized light with crystal plates, 382, 385.
convergent circular light, 386.
Double refraction, ellipsoid of elasticity, 373.
Fresnel's theory of, 372.
Huygens's construction in case of, 368.
isochromatic surfaces, 385.
in non-crystalline media, 386.
normal velocity surface, 376.
optic axes, 379.
physical explanation of, 366.
temperature effect on angle between axes, 386.
wave surface in crystals, 377.
wave surface in uniaxial crystal, 367, 370.
Drude, dispersion formula, 486.
Dunoyer, resonance radiation of Na, 593.

Eagle, mounting for concave grating, 263.
Echelon grating, 284–306.
Echelon reflecting, 307.
Edge-wave (diffraction), 273.
Eichenwald, energy flow in total reflection, 421.
Einstein, law of photo-electric effect, 763.
principle of relativity, 826.
proof of law, 764.
Electrical double refraction, 758.
Electro-magnetic theory of light, 397.
Electron spin, 688.
Electrons, dispersion, 490.
Ellett, polarization of resonance radiation, 723.
Ellipsoid of elasticity, 373.
Elliptical orbits, 136.
Emission (thermal) of polarized light, 343.
Energy-flow, along boundary in total reflection, 421.
in system of interference fringes, 170.
Entrainment of ether by matter, 814.
Epstein, theory of Stark effect, 748.
Eriometer of Young, 270.
Ether-drift, effect of earth's motion on intensity of terrestrial source, 824.
Fitzgerald-Lorentz hypothesis, 821.
Joos's experiment, 822.
Kennedy's experiment, 822.
Lodge's experiments, 823.

Ether-drift, Michelson-Morley experiment, 817.
Miller's experiments, 822.
principle of relativity and ether drift experiments, 826.
rotatory polarization and contraction of matter, 823.
Ether vapor, scattering by, 432.

Fabry and Buisson, Döppler effect in laboratory, 27.
and Perot interferometer, 309.
Fermat's principle, 62.
law, 75.
Filippov, sodium vapor dispersion, 503.
Filter, potassium for ultra-violet, 565.
Filters for absorption, 15.
Fish-eye camera, 67.
views, 68.
Fizeau, experiment on velocity of light in moving media, 814.
star diameters by interference, 176.
velocity of light, 16.
Flames, absorption by, 775.
Flash spectrum, 122.
Fluorescence, of crystals, 660.
efficiency and Stokes's law, 653.
influence of concentration, 658.
intensity variation with emission angle, 655.
investigation of, 654.
polarized, 659.
of rare earths, 661.
Fluorescence of gases, of atoms, 610.
of mercury destroyed by temperature, 641.
mercury, Lord Rayleigh's experiments, 643.
of mercury vapor, 636.
sensitized, 612.
sodium and mercury, theory, 645.
of sodium vapor, 618.
time interval between excitation and fluorescence, 640.
Fluorescence of liquids and solids, benzene, 651.
effect of solvent, 657.
inorganic compounds, 657.
line fluorescence of Weardale fluorite, 665.
organic compounds, 650.
Stokes's law, 653.
of sugar activated with dyes, 666.
Fluorite, absorption and dispersion, 525–526.
Fluorite dispersion curve, 472.
Focal lines, 59.
Focal-isolation method, 523.
Fortrat diagram, 155.
Foster, J. S., Stark effect, 745.
Foucault, velocity of light, 18.
Fowler, nelium series, 134.
Franck and Cario sensitized fluorescence, 611–613.
and Condon principle, 645.
and Grotrian 340 μ radiation of mercury vapor, 525.
and Wood, transformation of resonance to band spectrum, 635.
Fresnel, explanation of rotatory polarization, 572.
class of diffraction fringes, 219.
direction of revolution, rotatory polarization, 361.
integrals, 228.
rhomb, 354.
theory of double refraction, 372.
Fresnel-Arago laws of polarized interference, 346.
Frilled films on metal mirrors, colors of, 206.
Fry, T. C., and Ives, photo-electric emission of electrons, 767.
Fuchtbauer, step-up excitation of spectra, 599.
density effect on absorption lines, 529.

Gases, thermal radiation of, 776.
Gaviola fluorometer, 741.
Gaviola and Pringsheim curve for percentage of polarization, 343.
polarization of only one D line in magnetic field, 725.
Gehrcke, interference points of crossed plates, 325.
"Ghosts," theory of, 256.
Glasses, colored by colloidal metal, 442.
Gold, colors of granular films, 438.
optical constants, 566.
Goudsmit and Uhlenbeck, spinning electron, 689.
Gouy, anomalous wave propagation, 288.
Gratings (*see* Diffraction gratings).
"Green-ray," 84.

Grimaldi, discovery of diffraction, 219.
Group-velocity, 21.
Gudden and Pohl, theory of phosphorescence, 666.
Gundelach, dispersion of KBr, 526.

"Hacken-Methode," 500.
Hackett, F., magnetic rotation spectrum, 729.
Hagen and Rubens, temperature and emission of metals, 569.
Haidinger's brush, 343.
fringes, 207.
Hale, Zeeman effect on sun, 673.
Halos, 394.
diffraction, 269.
Hanle, observation of Larmor precession, 722.
Hanle and Richter, polarization in step-up excitation of fluorescence, 727.
Hardy and Silverman, fine structure of selective reflection bands, 522.
Hayner, Lucy, imprisonment of resonance radiation, 598.
Heat waves, longest, 523.
Helium spectra, 133.
resonance radiation in infra-red, 608.
Helmholtz, dispersion theory, 473.
ophthalmometer, 72.
Hertz, photo-electric effect, 760.
Herzfeld, dispersion, phase lag of vibration, 481.
Heurlinger, electronic bands, 156.
Houston, W. V., compound interferometer, 315.
Huygens, discovery of polarization, 329.
principle, 29.
Hydraulic analogy of radiating body, 781.
Hydrogen, atomic, spontaneous incandescence of metals in, 127.
dispersion of, 491.
dispersion and absorption of electrically excited, 503.
spectrum, production of, 125.

Iceland spar, double refraction of, 365.
Ilberg, Kerr effect, 738.
Images formed by reflection, 43–59.
Imes, HCl band spectrum, 152.
Infra-red, photography by, 424.
Inhomogeneous waves in absorbing media, 555.
Interference, 157.
astronomical applications, 176.
colors of thin plates, 188.
colors of mixed plates, 279.
composition of vibrations, 159.
corresponding points of source, 167.
distribution of illumination, 160.
disturbances of arbitrary phase, 161.
Fresnel biprism, 166.
Fresnel mirrors, 163.
frilled films, 206.
kinetic energy of vibration, 158.
light beats, 180.
Lippmann color photographs, 214
Lloyd's mirror, 165.
multiple reflecting planes, 198.
multiple reflections, 192.
Newton's colors, preparation of films for, 203.
Newton's rings, 190.
Newton's rings seen through prism, 202.
refractometer, 173.
silvered films, 194.
stationary waves, 210.
strata, elementary, in mica and soap films, 191–205.
thick plates, 195.
Young's experiment, 162.
Interference fringes, achromatic, 181, 184, 202.
along caustics, 187.
correction of lenses and prisms by, 299.
curves of equal thickness and inclination, 207.
energy flow in, 170.
Haidinger's, 207.
intensity distribution, 160.
limit to number, 167.
multiple reflections and fringe width, 209.
shift by sodium vapor, 184.
shift by transparent plate, 168, 182.
sources in line, 187.
with white light, 168.
Interferometer, adjustment of, 294.
compound or tandem, 315–316.
crossed Lummer plates, 325.

Interferometer, determination of refraction, 300.
determination of wave-length in infra-red, 524.
determination of wave-length with, 311.
effects of films on mirrors, 297.
Fabry and Perot, 309.
Lummer and Gehrcke, 317.
Lummer-Gehrcke plate technique, 319.
Michelson, 292.
Rayleigh, 174.
stellar at Mt. Wilson, 178.
Twyman and Green, 297.
Invisibility of objects, 97.
Iridescent crystals, 198.
Ives, H. E., directional distribution of electrons in photo-electric effect, 767.
Lippmann color photographs, 201–205.
photo-electric effect, 766.

Jaumann, white light pulse, 830 (Appendix).
Jeans, thermal radiation, 804.
Julius, anomalous dispersion and solar phenomena, 121.

Kemble, Raman lines of HCl, 452.
Kerr constants, table of, 735.
Kerr effect, discovery of, 736.
duration of fluorescence by Kerr cell, 740.
fluorometer of Gaviola, 741.
Kerr constant, 737–738.
Kerr effect in liquids, 736.
Abraham Lemoine method, 739.
modulation of light, 742.
velocity of light by Kerr cell, 740.
Kerr magneto-optic effect, 701.
Kinematograph, diagrams of wave-reflection, 57.
Kirn, dispersion of hydrogen, 491.
Kirchhoff's law, 774.
Kistiakowski and Nelles resonance spectrum of benzene, 652.
Klinkerfuess experiment, 813.
Koch, J., method for study of Rest-Strahlen, 517.
König, inhomogeneous waves, 556.
Kuhn, theory of rotatory polarization, 581.
Kuhn, theory of mercury fluorescence, 646.
Kundt, anomalous dispersion, 119.

Ladenburg, E., electron velocity and γ, 763.
Ladenburg, R., determination of magnetic rotation of Na vapor, 707.
paramagnetic rotation, 713.
Ladenburg, R., and Kopfermann, inverse Stark effect, 759.
and Loria, absorption and dispersion of H_2 electrically excited, 503.
and Minkowski, magnetic rotation, 712.
and Rubens, dispersion of water to 18 μ, 509.
Lambert's law of absorption, 101.
law of emission, 783.
Landé's formula, 687.
Landolt's fringe, 344.
Lange, cuprous oxide photo-cell, 770.
Langsdorf and Du Bridge, plane-polarized vibrations in natural light, 362, 829 (Appendix).
Larmor precession, 676.
Hanle's observation of, 722.
precession in strong fields, 695.
Lau and Ritter, tandem interferometer, 316.
Lebedew, radiation pressure, 794–797.
Lenard, phosphorescence, 666.
photo-electric effect, 761.
Lens, achromatic, 112.
aplanatic, 79.
pseudo-, 88.
refraction by, 79.
Lenz, interpretation of resonance spectra, 623.
Light, aberration of, 812.
absorption of, 99.
curved rays, 90.
dispersion of, 106.
elastic solid theory, 4.
electro-magnetic theory, 4.
Lorentz's theory, 5.
Newton's theories, 1, 2.
quantum theory, 5.
rectilinear propagation, 31.
sources for experiments, 13.
velocity of group, 21.
velocity by Kerr cell, 740.
velocity in moving media, 814.

Light, velocity of Fizeau, 16.
velocity of Foucault, 18.
velocity of Michelson, 18, 20.
velocity of Newcomb, 19.
velocity of Romer's method, 16.
white, nature of, 811, 830 (Appendix).
Light-beats, 180, 363.
with Kerr cell, 742.
Light-wave as standard of length, 301.
Lippmann's color photographs, 214.
plates, preparation of, 214.
Lithium spectra, 133–135.
Lochte-Holtgraven, sodium lamp, 597.
Loomis, F. W., iodine resonance spectrum, 632, 635.
theory of sodium fluorescence, 645.
Lorentz, theory of Zeeman effect, 674, 675.
Lummer-Gehrcke, interference points, 325.
interferometer, 317.
technique for, 319.
technique for crossed plates, 326.
Lyman "ghosts," 257.
helium in remote ultra-violet, 134–135.
hydrogen series, 127–133.

Macaluso and Carbino, magnetic rotation in Na vapor, 702.
Magnetic double refraction, 716.
moment of atom, 687.
quantum, number, 685.
rotation (*see* Faraday effect).
rotation spectra, 729.
Magneto-optics, 668.
of band spectra, 729–731.
circular components, 697.
in crystals at low temperature, 714.
direction of rotation, 699, 731.
dispersion electrons, 712.
double R and L quartz prism, 708.
explanation, 697.
Faraday effect, 696.
Faraday effect, theory of, 702, 710.
Faraday effect, time required for, 700.
and field strength, 699.
magnetic rotation spectra, 729.
magnetic rotation in absorption band, 708.
paramagnetic rotation, 712.
Magneto-optics, relation to Zeeman effect, 710.
resolution into circular components, 697.
rotation by sodium vapor, 702.
of sodium resonance radiation, 724.
Magneto-optics of resonance radiation, 718.
alternating fields, 723.
destruction of polarization by earth's field, 718.
mercury vapor, 718.
polarization of one *D* line only, 725.
sodium vapor, 724.
theory, 721.
Magnifying power of prism, 77.
Marcellin, interference by mica and soap films, 191–205.
elementary strata in mica and soap films, 205.
Maxwell, dispersion theory, 470.
electro-magnetic theory of light, 397.
equations of, 397.
theory of light, 4.
Meissner, duration of metastable states, 610.
Mercury vapor, absorption, 637.
arc, simple laboratory, 14.
fluorescence, 636.
long wave radiation, 524.
polarization in step-up excitation, 727.
resonance radiation, 588.
resonance radiation in magnetic field, 718.
selective reflection by, 535.
spectra by step-up excitation, 599–607.
spectrum, diagram of levels, 601.
spectrum forbidden line, 607.
Metals, alkali, optical properties of, 558–566.
Brewsterian angle for potassium, 565.
determination of n and κ, 551.
determination of principal azimuth and incidence, 557.
determination of R from emission, 550.
effects of film thickness, 568.
emission table, 551.
free electrons and optical properties, 542.

Metals, inhomogeneous waves and rigorous formulae, 554.
interference fringes by K films, 562.
metallic reflection, 547.
optical constants (table), 558.
optical constants in ultra-violet, 557.
principal azimuth and incidence, 551.
reflecting power and conductivity, 548.
reflecting power, formula for, 547.
reflecting power of (table), 550.
silver and gold optical constants, 566.
temperature and film emissivity, 569.
theory of metallic absorption, 544.
Metal films, scattering by granular, 436.
Metal vapors, dispersion of, 492.
Metastable atoms formed by N_2, 605.
state, duration, 610.
McLennan and McLeod, Raman spectra of liquid O_2, N_2 and H_2, 458.
Mica, interference mosaics, 205.
Michelson, echelon grating, 284.
interferometer, 292.
light-waves as standards of length, 300.
star diameters, 176.
velocity of light, 18–20.
visibility curves, 303.
Michelson and Morley, velocity of light in moving media, 815.
Michelson-Morley experiment, 817.
Microscope, refractive index by, 74.
objective, 81.
Microscopic vision and diffraction, 277.
Mie, scattering by colloids, 441.
Miller, ether drift experiments, 822.
Millikan, e/m determination, 489.
photo-electric experiment, 764.
Mirage, 87.
Mirrors, ellipsoidal, 45.
hyperbolic, 50.
parabolic, 47–50.
plane, reflection from, 43.
at right angles, 44.
spherical, 54.
Mixed plate colors, 279.
Mock-suns, 394.
Mohler and Chenault, ionization of gases by light, 769.
Molecular scattering, 476.
Molecular spectra, 146.
Monochromator for separation of close lines, 349.
simple prism, 13.
Morse, line fluorescence of fluorite, 665.

Natural light, state of polarization, 361.
presence of plane-polarization in, 362, 829 (Appendix).
Nebulae, velocity of extra gallactic, 25.
Neon, resonance radiation, 609.
Newton, black spot, 190.
colors, preparation of films, 203.
interference rings, 190.
theory of light, 1–3.
Nichols, E. F., selective reflection of quartz, 514.
Nichols, E. F., and Hull, radiation pressure, 794.
Nichols, E. L., and Merritt, Stokes's law of fluorescence, 653.
Nicol prism, 336.
Nitro-benzene, Kerr constant of, 738.
Nitroso-dimethyl aniline dispersion, 506.
surface color, 510.
reflection of polarized light, 331.
Non-homogeneous media, 82.
Normal velocity surface, 376.

Opals, cause of color, 198.
Ophthalmometer, 72.
Optic axes of crystals, 379.
Optical paradox, 274.

Parhelia, 394.
Paschen, gas radiation and absorption, 777.
infra-red H_2 series, 133.
mounting for concave grating, 262.
resonance radiation of helium, 608.
Zeeman effect in strong fields, 694.
Penetrating orbits, 142.
Penetration of disturbance into second medium in total reflection, 420.
Perrin, F., fluorescence disactivation, 659.

Pfund, radiation of powders, 781.
hydrogen line 7.4 μ, 133.
total reflecting refractometer, 70.
Phase change at polarization angle, 413.
Phase distribution over small source, 186.
Phillips phosphorescence of mercury vapor, 640.
Phosphorescence, 660.
activation of pure substances, 664.
duration of, 662.
fluorite, 664.
nature of, 663.
quenching by infra-red, 665.
temperature effect on, 662.
Phosphoroscopes, 663.
Photo-electric effect, cuprous oxide photo-cell, 771.
directional distribution of electrons, 767.
discovery of, 760.
effects with polarized light, 766.
effects with X-rays, 768.
Einstein's law of, 763.
electron current and wave-length, 764.
electron velocity and frequency of light, 762.
ionization of gases, 768.
methods of investigating, 761.
normal and selective effect, 766.
photo-conductivity, 769.
photo-voltaic effects, 770.
Photography by infra-red, 424.
Pickering, series of spectrum lines, 134.
Pigments, 103.
Pin-hole photography, 272.
Planck, quantum theory of radiation, 805.
radiation, formula, 802.
theory of damping, 479.
Planetary atmospheres, refraction by, 87.
Pohl, absorption by KBr, 649.
Pohl and Hilsch, light absorption in solids, 649.
and Pringsheim, photo-electric effect, 766.
Polarization, aperture of polarizing prisms, 337.
bi-quartz, 339.
Brewsterian angle, 330.
Polarization, circular, direction of revolution, 359.
determination of refractive index by, 332.
discovery, 329.
dispersion of polarization angle, 331.
double image prisms, 338.
by double refraction, 335.
elliptical and circular, 348, 352, 358.
elliptical, constants of, 357, 359.
elliptical polarization by surface films, 414.
Fresnel rhomb, 354.
Malus, law of, 333.
Nicol prism, 336.
by oblique emission, 343.
phase change at Brewsterian angle, 413.
plane of, defined, 329.
polariscope, construction of simple, 333.
quarter wave plate, 352.
Raman lines, polarization of, 467.
by reflection, 330.
by refraction, 332.
rotation of plane by transmission through oblique plate, 335, 412.
Polarization, rotatory, absorption of circular components, 574.
anomalous rotatory power, 578.
by biaxial crystals, 574.
circular double refraction of quartz along axis, 574.
Cornu prism, 574.
Fresnel's explanation of rotation, 572.
influence of solvent on rotatory power, 577.
physical explanation of rotation, 580.
quartz R and L, 571.
rotary dispersion, 571.
rotary dispersion of liquids, 577.
R.P. and change of state, 576.
R.P. in convergent light, 579.
R.P. of liquids, 576.
secondary wave from compound resonator, 584.
stereoscopic model, 583.
theory of rotation, 581, 586.
Polarized light, bi-quartz, 339.
circular and elliptical, amplitude of, 364.

Polarized light, convergent and divergent, 382, 385.
crystal plate transmission, 351.
curves for percentage of by oblique plates, 342.
detection of, 339.
from fluorescent vapors, 634.
interference of, 346.
Landolt's fringe, 344.
partially polarized and natural, 361.
plane-polarized vibration in natural light, 362, 829 (Appendix).
percentage of, 340.
polarized interference fringes, 347.
polarizing monochromator, 349.
Savart's plate, 339.
stationary polarized waves, 344.
Porter, A. B., diffraction and image formation, 278, 830 (Appendix).
Potassium bromide, absorption by, 649.
dispersion of to 18 μ, 526.
Powders, dispersion of, 117.
radiation of, 781.
Poynting, radiation pressure, 796.
Principal azimuth and incidence, 551.
Pringsheim and Terenin, mercury fluorescence, 640.
Prisms, achromatic, 110.
crossed (method of), 119.
cyanine, 120.
direct vision, 112.
magnifying power of, 77.
refraction by, 75.
refractive index of, 76.
resolving power, 113.
total reflecting, 69.
Projection of narrow light beam, 49.
Propagation rectilinear, 31.
Pyrometer, optical, 799.

Quantum numbers, 136–144.
effective, 144.
magnetic, 685.
Quantum theory, 5, 805.
Quarter wave plate, 352.
Quartz, absorption and selective reflection of, 518, 522.
refractive indices of, 520.

Radiating states, optical control, 603.
Radiation (*see* Thermal radiation), pressure of, 793.
Radiometer, 794.
Rainbow, theory of, 388.
Rainbow cup of C. V. Boys, 191.
Raman effect, HCl energy levels and transitions, 453.
non-polar molecules, 457.
polarization of spectra, 467.
Raman's experiment, 445.
relation to infra-red absorption, 451.
relation with resonance spectra, 464.
R.E. of diatomic gases, 451.
R.E. in liquid gases, 458.
selection rule, 454.
Smekal's prediction, 444.
technique for study of, 446, 450.
tetrachlorides, 463.
theory of, 446, 451, 454.
triatomic molecules, 458.
unmodified line and density, 467.
unmodified line, frequency changes, 466.
virtual transitions, 454.
Rare earths, fluorescence, 661.
Rasetti, Raman effect for gases, 452, 457.
Rayleigh, Lord, colors of opals, 198.
group velocity, 20.
interferometer, 174.
optical images and diffraction, 271.
optical paradox, 274.
pin-hole images, 272.
radiation by rotating oscillator, 147.
refractometer, 173.
resolving power of gratings, 251.
scattering of light, 423–425.
testing glass plates, 196.
zone-plate, 37.
Rayleigh-Jeans radiation formula, 803.
Rayleigh, Lord (II), atomic sodium fluorescence, 610.
mercury fluorescence, 643–645.
polarization of resonance radiation, 592.
scattering by gases, 426.
Reflection, from concave sphere, 54.
and image formation, 42.
laws of, derivation, 406.
metallic, 547.
at oblique incidence, 412.
at perpendicular incidence, 410.
of sound waves, 44.

Reflection, from spherical surfaces, 53.
total, 65, 69.
total reflection, penetration of disturbance, 418.
total reflection refractometer, 70.
total, theory of, 416.
by transparent media, laws of, 406.
by unpolished surface, 39.
of waves, kinematograph, 56.
Reflection, selective, by absorbing gas, 534.
by absorbing media, 509.
and absorption curves, 515.
curves for, 512.
fine structure of bands, 522.
influence of refractive index, 510.
points of maximum reflection, 512.
of quartz, 514.
residual rays, 516.
in ultra-violet, 511.
Refraction, astronomical, 83.
atmospheric, 84.
effect on beam width, 77.
laws of, derivation, 406.
by lens, 79.
by non-homogeneous medium, 82.
by plane parallel plate, 72.
at plane surface, 64.
by planetary atmosphere, 87.
by prism, 75.
Snell's law, 64.
by sphere, 80.
spherical waves at plane surface, 73.
at surface of pond, 67.
theory of, 469.
by unpolished surface, 39.
vision under water, 67.
Refractive index, 65.
complex refractive index, 488.
of plate by microscope, 74.
of prism, 76.
Refractometer, Haber and Loewe, 174.
Rayleigh, 173.
total reflection, 70.
Williams, 179.
Relativity, principle of, 826.
Residual blue, 429.
Residual rays, interferometer study of, 517.
Rest-Strahlen, 516.
Resolving power of grating, 251.
Resolving power of grating, prism, 113.
telescope, 267.
Resonance colors of granular metal films, 437.
Resonance radiation, absorption by resonating gas, 591, 595.
Bohr's theory of R.R., 596.
of helium in infra-red, 608.
imprisonment of R.R., 597.
of mercury vapor, 588.
method of excitation, 590.
polarization of R.R., 592.
preparation of Na bulbs, 593.
reflecting power of Na vapor, 595.
separate excitation of *D* lines, 596.
of sodium vapor, 587, 593.
transition from to selective reflection, 539.
Resonance radiation of electrically excited atoms, 607.
neon, 609.
Resonance spectra, of benzene, 652.
discovery, 621.
of iodine, 623.
multiplex excitation, 626.
polarization of light, 634.
preparation of iodine tube, 624.
of sodium vapor, 621.
of solids, 648.
theory of, 628.
transformation to band spectrum, 635.
Rochon prism, 338.
Rock salt, absorption and dispersion, 525, 526, 529.
table of optical constants, 535.
Römer, velocity of light, 16.
Roschdestwenski, "Hacken-Methode," 500.
sodium dispersion, 502.
Rotation of plane of polarization by quartz (table), 572.
Rotatory dispersion, 571.
Rotatory polarization (*see* Polarization, rotatory).
Rowland, concave grating, 260.
Rubens and Aschkinass, 56 μ waves with prism, 521.
and Hagen, reflection by metals, 550.
and Holnagel, residual rays, 517.
and Ladenburg, refraction of H_2O, 509.

Rubens and Aschkinass, and Nichols, residual rays, 516.
and Wood, focal isolation, 523.
Rupp, modulation of light, 742.
conduction by phosphorescent crystals, 667.
Rydberg constant, 127.

Savart's plate, 339.
Scattering, abnormal polarization by absorbing particles, 430.
by colloids, 440.
density fluctuations and scattering, 431.
depolarization and molecular structure, 434.
depolarization of scattered light, 433.
by ether vapor, 432.
by gases, 426.
by granular metal films, 436, 438.
liquids, solids and crystals, 427.
metal glasses, 442.
molecular, theory of, 476.
polarization measurement, 432.
polarization by ultra-microscopic particles, 442.
Rayleigh formula for, 425.
residual blue, 429.
by small particles, 423.
by sodium fogs, 440.
Schaefer, selective reflection in ultra-violet, 511.
Schaefer and Gross, energy flow in total reflection, 422.
Schmidt, theory of sun, 85.
Schuster, diffraction fringes, 222.
scattering and dispersion, 477.
Scintillation of stars, 91.
Selective reflection by mercury vapor, 535.
by crystals, 513.
in ultra-violet, 511.
Selenyi, stationary light waves, 213.
Selmeier, dispersion formula, 470.
Sensitized fluorescence, 612.
effects of foreign gas, 615.
Silver, optical constants, 566.
Silverman, reflection bands, 523.
Smekal, prediction of Raman effect, 444.
Snell's law, 64.
Soap-film, elementary strata of, 206.
Sodium, absorption spectrum, 146.
Sodium, direction of rotation at band lines, 734.
fluorescence of atomic, 610.
magnetic rotation spectra, 731.
molecules, absorption by, 616.
resonance radiation in magnetic field, 724.
separate excitation of *D* lines, 596.
sodium lamp, 597.
vapor, dispersion, 492, 500.
vapor, refractive indices, 498–500.
vapor, shift of interference fringes by, 184.
vapor, in ultra-violet, 502.
Sodium fog, scattering by, 424, 440.
colors of granular deposits, 436.
Solar phenomena and anomalous dispersion, 121.
Sommerfeld, diffraction, 221.
penetrating orbits, 142.
theory of structure of spectrum lines, 136–139.
Soret, zone-plate, 37.
Sound-waves, photography of, 95.
Space quantization, 685.
Sparrow, theory of "ghosts," 256.
Spectrum, of alkali metals, 139–141.
band, 148.
"flash," 121.
hydrogen, 125.
ionized helium and lithium, lines, fine structure of, 139.
molecular, 146.
older theories of, 124.
penetrating orbits, 142.
Pickering series, 134.
rotation bands, 149–154.
series, early study, 127, 129.
sodium absorption, 128, 146.
spectral terms, 137.
vibration-rotation, 151–154.
Spherical surface, reflection by, 53.
Star diameters by interferometer, 178.
Star velocities, 24.
Stark, Döppler effect in canal rays, 27.
Stark effect, alkali metals, quadratic effect, 758.
discovery of, 742.
electrical double refraction, 758.
electron's path in, 748.
energy levels, 754.
Lo Surdo method, 744.

Stark effect, quadratic effect, 756.
stereoscopic model of electron's motion, 751.
theory of, for hydrogen-like atoms, 745.
Stationary light waves, Selenyi, 213.
transition to Lloyd's fringes, 214.
Wiener, 210.
Stationary states, 131.
Stefan, attempt to measure star diameter, 176.
plane-polarization in natural light, 829 (Appendix).
Stefan-Boltzmann law, 797.
Step-up excitation, polarization in, 727.
power relation of intensities, 604.
of spectrum, 599.
Stereoscopic diagram, explaining rotatory polarization, 583.
of reflected components, 407.
Stereoscopic model of electron's motion in Stark effect, 751.
Stern and Gerlach, orientation of atoms, 685.
Stokes's law, 653.
Striae, method of, 93.
Sun, temperature of, 800.
theory of Schmidt, 85.
Zeeman effect on, 673.
Surface-color, 103, 509.
influence of refractive index on, 510.
Surface films and polarization, 414.
Sylvite, absorption and dispersion, 525, 526.
Synthetic waves, 480.

Talbot's bands, 280, 829 (Appendix).
Telescope, resolving power, 267.
reflecting, 51.
zone-plate, 38.
"Terms," spectral, 137.
Testing glass plates for flatness, etc., 196.
Thallium, absorption and fluorescence, 611.
energy diagram, 612.
Thermal radiation, absorption by flames, 775.
"black-body" radiator, 791.
effects of cleavage planes on emission of transparent solids, 779–781.
emission and absorption, 773.
Thermal radiation, emission of polarized infra-red rays, 785.
emission of polarized light by crystals, 789.
emission and reflection relation, 785.
emission of substances with absorption bands, 786.
equilibrium between radiation and matter, 792.
excitation of line spectra, 778.
formulae for radiation, 801.
of gases, 776.
general discussion of, 772.
hydraulic analogy of radiating body, 781.
Kirchhoff's law, 774.
Kirchhoff's law, proof of, 790.
Lambert's law, 783.
neodymium quartz emission, 787.
optical pyrometers, 799.
polarized emission, 783.
pressure on absorbing gas, 797.
pressure of radiation, 793.
quantum theory of, 805, 810.
Rayleigh-Jeans, formula, 803.
Ritchie's experiment, 774.
of solids and liquids, 778.
Stefan-Boltzmann law, 797.
tangential component of radiation pressure, 796.
white light, nature of, 811, 830 (Appendix).
Wien's laws, 800.
Tiede and Tomaschek, phosphorescence, 664.
Töpler, "Schlieren-methode," 93.
Total reflection, 65.
theory of, 416.
Transmission coefficient, 104.
Trowbridge and Wood, resolution of CO_2, 4.3 μ band, 147.

van der Lingen, mercury fluorescence, 639.
Vibrations, amplitude of forced, 481.
composition of, 159.
free and forced, 474.
Visibility curves, 303.
Voigt effect, magnetic double refraction, 717.
method for principal azimuth and incidence, 557.
Voss and Wood.
cence, 639

Water, dispersion of, 509.
Wave, absence of back, 8.
anomalous propagation, 288.
edge waves, 273.
front, 9.
heat-waves, longest, 523.
inhomogeneous waves, 554.
length determination with Fabry and Perot interferometer, 311.
length and frequency, 10.
light-waves as standards of length, 301.
plane, effect on exterior point, 33.
secondary, phase and amplitude of, 3.
secondary wavelets, 30.
sound photographs, 32.
stationary light-waves, Wiener, 210.
stationary polarized, 344.
surface in crystals, 367.
synthetic waves, 480.
table of wave-lengths, 11.
velocity, derivation of, 404.
wave-length and color, 12.
Wawilow, fluorescent efficiency, 653.
Weigert, polarized fluorescence of liquids, 659.
White light, nature and resolution of by grating and prism, 811, 830 (Appendix).
Wien's laws of radiation, 800.
Wiener, stationary light-waves, 210, 211.
stationary polarized waves, 344.
Williams, reflecting echelon, 307.
refractometer, 179.
Wollaston prism, 338.
Wood, continuous absorption at head of Na series, 146.
dispersion of Na vapor, 492, 500.
echelette grating, 265.
hydrogen spectrum, 125.
interference fringes achromatized by Na vapor, 184.
magnetic rotation spectra, 703, 729.
optical properties of alkali metals, 558.
polarized fluorescence, 634.
Wood, selective reflection by Hg vapor, 535.
sodium principal series, 146.
sound-wave photographs, 57.
step-up excitation of Hg spectra, 600.
zone-plate, 38.
Wood and Dunoyer, excitation of *D* lines, 596.
and Ellett, polarization of resonance radiation in magnetic field, 718.
Worthing, thermal radiation, 784.

Young, colors of mixed plates, 279.
diffraction fringes, 218.
discovery of interference, 162.

Zeeman, velocity of light in moving media, 817.
Zeeman effect, anomalous, 686.
anomalous magnetic levels, 691, 694.
for band spectra, 730.
in band spectrum of Na, 735.
Bohr magneton, 687.
Bohr theory, 679.
classical theory of, 674.
of crystals at low temperature, 714.
discovery of, 668.
inverse effect, 672.
Landé formula, 687.
Larmor precession, 676.
Larmor precession in strong fields, 695.
magnetic quantum number, 685
magnitude of separation, 671.
observation without spectroscope, 671.
Paschen-Back effect, 694.
quantum theory of, 682.
relation to Faraday effect, 710.
and spectral series, 669.
on the sun, 673.
Zemansky, imprisonment of resonance radiation, 597.
Zone-plates, 37, 253.
Zone-plate telescope, 38.
phase-reversal, 38–39.
Zones, Fresnel-Huygens, 36.

Snell's resonance radiation, 587.
Soap-film, [illegible] spectra, 621, 623.
Sodium, absorption [illegible]