TEXT CROSS WITHIN THE BOOK ONLY

UNIVERSAL
LIBRARY

OU_162301

UNIVERSAL
LIBRARY

OSMANIA UNIVERSITY LIBRARY

Call No. 523 / R82T Accession No. 19189

Author Rosseland. S

Title Theoretical Astrophysic

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

1936.

THE

INTERNATIONAL SERIES

OF

MONOGRAPHS ON PHYSICS

GENERAL EDITORS

R. H. FOWLER AND P. KAPITZA

THE INTERNATIONAL SERIES OF MONOGRAPHS ON PHYSICS

GENERAL EDITORS

R. H. FOWLER
F.R.S., Fellow of Trinity College, Cambridge, Professor of Applied Mathematics in the University of Cambridge.

P. KAPITZA
F.R.S., Moscow, lately Fellow of Trinity College, Cambridge, and Royal Society Messel Research Professor.

Already Published

CONSTITUTION OF ATOMIC NUCLEI AND RADIOACTIVITY. By G. GAMOW. 1931. Royal 8vo, pp. 122.

THE THEORY OF ELECTRIC AND MAGNETIC SUSCEPTIBILITIES. By J. H. VAN VLECK. 1932. Royal 8vo, pp. 396.

WAVE MECHANICS. ELEMENTARY THEORY. By J. FRENKEL. 1932. Royal 8vo, pp. 286.

WAVE MECHANICS. ADVANCED GENERAL THEORY. By J. FRENKEL. 1934. Royal 8vo, pp. 533.

THE THEORY OF ATOMIC COLLISIONS. By N. F. MOTT and H. S. W. MASSEY. 1933. Royal 8vo, pp. 300.

RELATIVITY, THERMODYNAMICS, AND COSMOLOGY. By R. C. TOLMAN. 1934. Royal 8vo, pp. 518.

ELECTROLYTES. By HANS FALKENHAGEN. Translated by R. P. BELL. 1934. Royal 8vo, pp. 364.

CHEMICAL KINETICS AND CHAIN REACTIONS. By N. SEMENOFF. 1935. Royal 8vo, pp. 492.

THE PRINCIPLES OF QUANTUM MECHANICS. By P. A. M. DIRAC *Second Edition*. 1935. Royal 8vo, pp. 312.

RELATIVITY, GRAVITATION, AND WORLD-STRUCTURE. By E. A. MILNE. 1935. Royal 8vo, pp. 375.

THE QUANTUM THEORY OF RADIATION. By W. HEITLER. 1936. Royal 8vo, pp. 264.

THEORETICAL ASTROPHYSICS

ATOMIC THEORY AND THE ANALYSIS OF STELLAR ATMOSPHERES AND ENVELOPES

BY

S. ROSSELAND

OXFORD
AT THE CLARENDON PRESS
1936

OXFORD UNIVERSITY PRESS
AMEN HOUSE, E.C. 4
LONDON EDINBURGH GLASGOW NEW YORK
TORONTO MELBOURNE CAPETOWN BOMBAY
CALCUTTA MADRAS
HUMPHREY MILFORD
PUBLISHER TO THE UNIVERSITY

PRINTED IN GREAT BRITAIN

PREFACE

THE aim of this work is to formulate a programme of theoretical astrophysics.

This volume contains what may be called a bird's-eye view of our knowledge of stellar atmospheres and envelopes, projected on to atomic physics as a background.

A second volume is planned to give a similar view of the internal structure of stars, analysed in terms of nuclear physics and hydrodynamics.

I would like to express my best thanks to Mr. G. Randers for his indefatigable help in the final revision of the manuscript and in proof-reading; to Mr. Steensholt for his valuable assistance in the preparation of the first version of the book; and to the staff of the Oxford University Press for their care and attention in the printing.

S. R.

BLINDERN, near OSLO, NORWAY,
April, 1936.

CONTENTS

INTRODUCTION xi

I. ANALYTICAL DYNAMICS

1. The Equations of Lagrange and Hamilton 1
2. The Hamilton-Jacobi Integration Theory 5
3. Motion in an Electromagnetic Field of Force 7

II. STATISTICAL MECHANICS

4. The Theory of Gibbs 10
5. Dissociative Equilibrium in Gaseous Systems . . . 13

III. QUANTUM MECHANICS

6. The Wave Equation 18
7. The Quantum Conditions 21
8. The Einstein Probability Coefficients and the Theory of Thermal Radiation 22
9. Wave-mechanical Calculation of B_{ik} 24

IV. SOLUTIONS OF SPECIAL PROBLEMS

10. The Free Particle 31
11. The Harmonic Oscillator 33
12. Central Fields 35
13. The Rotator 42
14. The Balmer Formula 44
15. Atomic Dimensions 46

V. ATOMS WITH ONE OUTER ELECTRON AND AN ELECTRONIC CORE

16. The Rydberg-Ritz Formula 49
17. The Spinning Electron 53
18. The Alkali Spectra 57

VI. THE PERIODIC SYSTEM

19. Pauli's Exclusion Principle 59
20. Influence of Electronic Screening 63
21. Relative Importance of Sub-groups 65
22. The Problem of Argon 66
23. The Problem of the Iron Group 66
24. The Fifth, Sixth, and Seventh Periods 67

VII. THE THEORY OF MULTIPLETS

25. Nomenclature 70
26. Displaced Terms 74
27. Probabilities of Transitions between Multiplet Levels . . 75

VIII. THEORY OF ABSORPTION AND EMISSION

28. Energy of the Radiation Field 78
29. Quantization of the Radiation Field 80
30. Structure of an Emission Line 83
31. Probability Distribution Law of Weisskopf and Wigner . . 87
32. Dispersion Theory 89
32. 1. Classical Theory of Dispersion 92
32. 2. Quantum Theory of Dispersion 97
33. Influence of the Doppler Effect 100
34. Experimental Tests 102

IX. TRANSFER OF RADIATION IN A STAR

35. The Equation of Radiative Transfer 105
36. The Flux of Radiation 106
37. Digression on the Stellar Interior 107
38. The Case of an Atmosphere 110
39. The Formation of the Continuous Spectrum . . . 113
40. Formation of Absorption Lines. General Considerations . . 120
41. Coherent Scattering 122
42. Absorption and Scattering Combined 126
43. More General Solutions 128
44. An Integral Equation for J 130
45. Solutions by Numerical Methods 132

X. PROFILES OF ABSORPTION LINES

46. Observations of Line Profiles and their Interpretation . . 134
47. Theory of Cycles 140
48. Cycles in Dilute Temperature Radiation 145
49. Woolley's Theory of the Hydrogen Lines 146
50. Subordinate Lines. 153
51. The Equation of Ionization 153

XI. THE TOTAL INTENSITY OF ABSORPTION LINES

52. Theory of Total Intensities 160
53. Test by Observational Data 166
54. Analysis of Line Intensities by Adams and Russell . . 167

XII. THERMAL EXCITATION OF STELLAR ATMOSPHERES

55. Preliminary Considerations 170
56. Electrically Neutral Atoms 172
57. Intensity of Arc Lines 175
58. Intensity Maxima of the Lines 175
59. Intensity of Spark Lines 177

60. The Rate of Fading-out of the Lines 179
61. The Electron Density 180
62. Giants and Dwarfs. Difference in Effective Temperature . . . 181
63. Determination of Temperature from Intensity Maxima . . . 182
64. Abundance of Elements in the Stars 182

XIII. THE OPACITY

65. Various Causes of Opacity 184
66. Theory of the Photoelectric Effect 189
67. Differential Effects of Opacity 194

XIV. ROTATING STARS

68. Widening of Line Profiles by Rotation 202
69. Intrinsically Narrow Lines 203
70. Incipient Rotation Effect, and the Determination of the Intrinsic Line Width 205
71. Harmonic Analysis of Observed Profiles 207
72. Eclipsing Binaries 210
73. Observational Results 211

XV. THE EFFECT OF MAGNETIC AND ELECTRIC FIELDS ON ABSORPTION LINES

74. The Zeeman Effect 216
75. Zeeman Effect in the Sun 218
76. The Stark Effect 220

XVI. MOLECULAR COMBINATIONS IN STARS

77. Spectra of Molecular Compounds 223
78. Rotatory Oscillation Bands 224
79. Electronic Bands 225
80. Multiplet Structure of the Bands 228
81. Occurrence of Bands in the Spectral Sequence . . . 230
82. Isotope Effects in Band Spectra 232

XVII. APPLICATION OF THE THEORY OF MOLECULES TO THE DETERMINATION OF EFFECTIVE TEMPERATURE

83. Intensity in Bands and the Determination of Temperature . 234
84. Temperature of the Sun from Bands 235
85. Temperature of the Ionosphere 236
86. Temperature from the Relative Intensities of Bands in a Sequence. 237
87. Hutchisson's Theory of Vibrational Transition Probabilities . 238
88. *N*-type Spectra 242
89. Bands in the Spectra of Comets 243
90. Intensity of Auroral Bands 244

XVIII. DISSOCIATIVE EQUILIBRIUM OF STELLAR COMPOUNDS

91. The Reaction Isochore 245
92. Expressions for the Relative Abundance of Molecules in an Atmosphere 248
93. Abundance of H and N in Late-type Stars . . . 251
94. Carbon Compounds in Late-type Stars 254
95. Stellar Compounds in General 257
96. Molecular Compounds in Planetary Atmospheres . . . 257
96.1. The Earth 257
96.2. Venus and Mars 259
96.3. The Major Planets 259

XIX. THE SOLAR ENVELOPE

97. Extended Atmospheres 261
98. Chromosphere and Corona 261
99. Theories of the Chromosphere and the Corona . . . 265
100. A Calcium Chromosphere Supported by Light Pressure . . 270
101. A Chromosphere Supported by the Emission of Corpuscular Rays . 274
102. Influence of Magnetic Fields 278
103. Kiepenheuer's Theory of the Corona 282

XX. STARS WITH EXTENSIVE ENVELOPES

104. Envelopes of Giant Stars. 284
105. Bright-line Stars 286
106. Rotation Hypothesis for *Be* Stars 289
107. Expanding Envelopes. *WR* Stars and Novae . . . 291
108. Envelope Expanding with Constant V 293
109. Accelerated Expansion 295
110. Decelerated Expansion 297

XXI. FORBIDDEN TRANSITIONS

111. Transitions from Metastable States 301
112. Metastable States of Astrophysical Significance . . . 303
113. Origin of the Luminosity in Gaseous Nebulae . . . 307
114. Spectra of Gaseous Nebulae 310
115. Forbidden Lines in Spectra of Novae and Other Sources . . 313

XXII. THEORY OF RADIATIVE TRANSFORMATIONS IN NEBULAE

116. Ionization Theory of Nebulae 316
117. The Trifid Nebula 319
118. The Temperature in Nebulae 320
119. Transformation of Radiation in a Nebula . . . 322
120. Theory of a Three-State Cycle 325
121. The Boundary Conditions 327

122. The Resonance Line 328
123. Radiation Field in an Expanding Nebula 331
124. Excitation of Permitted Lines of O III and N III . . . 333
125. Excitation of Forbidden Nebular Lines 335
126. Chemical Composition of Nebulae 336

XXIII. ABSORPTION BY INTERSTELLAR GASES

127. Detached Lines 337
128. The Galactic Rotation and the Intensity of Interstellar Lines . 338
129. The Density of Interstellar Calcium and Sodium 341
130. Cosmic Dust and the Reddening Effect 343

APPENDIX 346

INDEX 349

INTRODUCTION

AMONG natural sciences astronomy is unique as the mother of all other sciences, the oldest by far in years, and probably still the one which is dearest to the heart of man. It was originally inextricably linked up with religious thought and practice, and it will always have something to say about the general outlook of man on the universe, which must take place, so to speak, through the astronomer's telescope.

It is the most poetical of sciences, and many a beautiful human dream was woven into a celestial language. But with oncoming age astronomy has taken on a sterner countenance. Poetry may still be there, but it has been forced into a different garb. The machine age has revolutionized astronomy like everything else. There is the ever-insistent demand for the mass production of observations, for increased precision in measurements, for further detail in arithmetical reductions, and for more satisfying theoretical interpretations. As a consequence an astronomical observatory of to-day looks more like a factory plant than an abode for philosophers. The poetry of constellations has given way to the lure of plate libraries, and the angel of cosmogonic speculation has been caught in a cobweb of facts insistently clamouring for explanations.

Who has not experienced the mysterious thrill of springtime in a forest, with sunbeams flickering through the foliage, and the low humming of insect life? It is the feeling of unity with nature, which is the counterpart of the attitude of the scientist analysing the sunbeams into light quanta and the soft rustling of a dragon-fly into condensations and rarefactions of the air. But what is lost in fleeting sentiment is more than regained in the feeling of intellectual security afforded by the scientific attitude, which may grow into a trusting devotion, challenging the peace of a religious mystic. For in the majestic growth of science, analytical in its experimental groping for detail, synthetic in its sweeping generalizations, we are watching at least one aspect of the human mind, which may be believed to have a future of dizzy heights and a nearly unlimited perfectibility.

The analytic nature of science has revealed itself most clearly in the constantly growing importance of the atomic conceptions; and astronomy is the science which, after physics and chemistry, has

benefited most from this development. It is at present the avowed goal of these sciences to build up a complete theoretical self-contained structure, based on atomic theory, sufficient for the adequate description of all physical and chemical processes in the universe.

The historical development leading up to this ambitious prospect may be followed step by step in the literature. It is not necessary to follow uncertain roads for thousands of years back into history; it has all developed before our own eyes.

It is true, of course, that the idea that matter is ultimately atomistic in structure goes back to the very beginning of science; but the saturation of nearly all physical and chemical theories with atomic conceptions is of more recent origin. Chemistry has always been the stronghold of atomism, and rightly, because the limited number of chemical elements, and the formation of chemical combinations out of pure elements mixed in constant multiple proportions, should suffice as an irrefutable proof that the ultimate laws of nature are atomistic, and hence not reconcilable with a description of nature in purely continuous terms.

But physical experiments have the advantage of bringing the atoms nearly within the grasp of human fingers. In the Brownian motion we observe directly the reflections of atoms bouncing back from a small obstacle, and in Faraday's law of electrolysis the counting of separate atoms moving through an electrolyte is nearly tangible. Still more is this the case in Millikan's experiments on the electric charge on fine oil-drops.

In the phenomena associated with the passage of electric currents through rarefied gases the electrical properties of the ultimate particles of matter reveal themselves still further and lead to the discovery of the electron as an integral part of matter. This immediately raises the question as to the nature of the corresponding positive entity, which was nearly solved by Lenard about the beginning of this century. Nearly, but not quite. There is a minimum 'Reizschwelle' in such matters, which in the present case was first surmounted by Rutherford in 1911, when he showed how the experiments of Geiger and Marsden on the scattering of alpha-particles by thin metallic foil unambiguously pointed to the conclusion that the positive electricity in a molecule is concentrated practically at points, the *nuclei*. That each atom should contain but one nucleus could not be definitely concluded from the experiments in question,

but is an outcome of later developments. By studying the cases when the Couloumb law of force between an alpha-particle and a nucleus breaks down, the dimensions of the latter were found to be of the order 10^{-12} cm., or even less, being thus a hundred thousand times smaller than the atoms themselves.

This success was relentlessly pursued, so as to determine the number of electrons in an atom, or what is the same thing, the number of unit positive charges on the nucleus. This could be determined by studying the deviations suffered by alpha-particles passing through matter, and it was found that this number was, apart from minor exceptions, equal to the ordinal number of the element in the periodic table, the so-called atomic number, the fundamental importance of which was first realized by van den Broek. J. J. Thomson did the same thing previously by counting the number of electrons in the atom, using X-rays.

And then the theoretical physicists discovered that they were in for hard times. For it was now only too evident that ordinary mechanics and electrodynamics cannot be taken seriously inside the atoms. This might have been realized at a much earlier stage, and on one occasion it was so realized when Gibbs declared (*Elementary Principles of Statistical Mechanics*, 1900) that the failure of the specific heat of diatomic gases to show the vibrational degree of freedom was a severe indictment against the whole application of mechanics to atomic problems. And how *can* a composite gas retain its chemical properties intact indefinitely, seeing that its molecules suffer billions of violent collisions per second? But the physicists refused to become shocked until the plain facts left them no alternative, as was the case with the picture of the atoms drawn up by Rutherford. For, since the forces holding the atoms together were purely electrostatic in nature, the atoms could not possibly remain stable. Static configurations are unstable, and moving electrons will radiate energy, and not stop until they coalesce with the nucleus. Moreover, spectroscopy would demand electrons performing harmonic oscillations, and, even more rigidly, one electron for each individual spectral line which could not be explained as a higher harmonic of some other line. This would demand thousands of electrons in a single hydrogen atom, which J. J. Thomson had proved to contain but one.

The quantum theory, which first burst out in full blossom after Rutherford's discoveries in 1911, originated in September 1900 when

Max Planck contributed his first paper on the quantum theory of temperature radiation to the Prussian Academy. Considering the prevailing tendencies in physics it is no wonder that it got only a half-hearted start, and that Planck spent the better part of the next fifteen years in attempting to smooth out the apparent hardnesses in his original assumptions.

It was Einstein who, by his work on specific heats and the photo-electric effect, gave to the quantum theory its first concrete actuality. Incidentally Einstein then formulated the idea of light quanta, a paradoxical notion which did not lose its sting until twenty years or more later, when it was realized that this paradox contained a good part of the real truth.

However, the man who started the quantum theory on its road to success was undoubtedly Bohr, who showed how the principal laws of spectroscopy may be interpreted from the quantum-theory point of view, provided the Rutherford view of atomic constitution is accepted. It was at this time that it was first realized that quantum theory aimed at a complete re-interpretation of the basic notions of classical physics. By the two quantum postulates, which Bohr made his starting-point, certain limitations of ordinary mechanics and electrodynamics were recognized to be a necessary corollary for the introduction of quantum conceptions, and hence to be the essentially new feature of the theory.

It is not true, however, that all subscribers to the quantum theory realized the radical implications of these postulates from the start. But there came constantly new impetuses in this direction, one of the most effective probably being Einstein's theory of temperature radiation of 1917, where Planck's original result was derived simply by assuming Bohr's postulates to hold, and assuming the elementary atomic processes to go by chance, just as the radioactive processes appear to do. This point of view was carried still farther by Bohr in his *Correspondence Principle* of 1918, which was a first attempt at the creation of a separate quantum mechanics, a self-sufficient system of mathematical rules for the description of atomic phenomena, just as ordinary mechanics plays the part of such a system for the motion of the planets of the solar system. And then, half a dozen years later, the idea behind this attempt was carried out literally by a pleiad of young scientists, Heisenberg, Dirac, de Broglie, and Schrödinger, giving rise to the matrix mechanics on one hand and

the wave mechanics on the other, which we now know to be but different aspects of one and the same theoretical structure, the quantum mechanics.

Just as many scientists of the last century believed it possible to describe the universe adequately in purely mechanical terms, so we are now expecting the physical universe to yield its secrets when attacked with the artillery of quantum-theory conceptions. In pursuing this course we are liable to make some of the errors of our forefathers all over again. But there is nevertheless ground for some confidence, since there is a much greater margin of safety this time, the quantum theory having already been subjected to many severe tests, which show that its range of application extends all over the field of terrestrial physics and chemistry. Moreover, during the last fifteen years astronomy has become permeated by atomic theory, and it is by now abundantly clear that further progress in theoretical astrophysics will depend intimately on the astronomers understanding how to express their observed facts in the language of atomic theory.

The beginning of this development goes back to the foundation of astronomical spectroscopy about 1860. Already Secchi's first survey of the stellar spectra showed that the stars are made from the same chemical elements as the earth and the sun, indicating very strongly that terrestrial physics and chemistry, properly administered, must apply throughout the starry heaven. This result has been further fortified by every new improvement in laboratory technique, leading to a match of previously unidentified lines in stellar spectra by lines in terrestrial sources.

The first really sensational case of this sort was the discovery by Lockyer and Ramsay (1893) that the yellow doublet discovered by Lockyer in the chromospheric spectrum of the sun (1868) was really due to helium, which was previously unknown on the earth. Through the work of Bohr and Fowler (1915) it became clear that the so-called 'secondary hydrogen lines' first discovered by E. Pickering in the spectrum of ζ Puppis were also due to helium, this time in its singly ionized state. This was the first strict proof that the elements in stellar atmospheres are in a state of ionization. It is true that Lockyer many years before (cf. *Chemistry of the Sun*, or *Sun's Place in Nature*) had urged the idea that the change in the spectra of the elements on passing from the arc to the spark corresponds to a progressive change in the atoms themselves due to the increased temperature. But the

notions of atomic structure entertained at those times were too crude for Lockyer's ideas to assume a more concrete form. Yet it must be admitted that he was right in his contentions in this particular respect, although their value is slightly diminished by the fact that in his writings the finer points are sometimes obscured by controversies.

Although Bohr must thus be considered the pioneer in the field, it was the Indian physicist Megd Nad Saha who (1920) first attempted to develop a consistent theory of the spectral sequence of the stars from the point of view of atomic theory. Saha's work is, in fact, the theoretical formulation of Lockyer's view along modern lines, and from that time the idea that the spectral sequence indicates a progressive transmutation of the elements has been definitely abandoned. From that time dates the hope that a thorough analysis of stellar spectra will afford complete information about the state of the stellar atmospheres, not only as regards the chemical composition, but also as regards the temperature and various deviations from a state of thermal equilibrium, the density distribution of the various elements, the value of gravity in the atmosphere and its state of motion. The impetus given to astrophysics by Saha's work can scarcely be over-estimated, as nearly all later progress in this field has been influenced by it, and much of the subsequent work has the character of refinements of Saha's ideas.

Fowler and Milne showed how the analysis could be made much more precise by focusing attention on the maximum intensity of the lines, instead of considering the marginal appearance or disappearance as Saha preferred to do. In this way it was possible to show conclusively that the density in the reversing layer of the stars is several thousand times smaller than the density of ordinary air.

It was further shown by Fowler that the conspicuous contrast between the persistence of hydrogen lines with increasing temperature beyond their maximum and the very rapid decline of metallic arc lines beyond their maximum is a simple consequence of the different atomic constitution of these elements. Milne, on the other hand, showed how varying ionization with height in the atmosphere, and the dependence of general opacity on temperature and density, are reflected in the behaviour of the spectral lines, and are responsible for some of the more striking spectral differences between giant and dwarf stars.

In this initial phase of development no detailed theory of the form of the absorption lines was involved, except that the intensity of a particular line will increase when the number of atoms of the right kind and in the right state above the photospheric level increases. Thus only the most fundamental features of atomic theory were utilized. And it is interesting that doctrines like the existence of definite energy levels forming excited states from which the lines are formed by transitions to other states, can be subjected to such a convincing test by astrophysical observations. One may recall here all the difficult experimental technique which a couple of decades ago was brought to bear on the problem of bringing out the successive lines of a series by increasing the voltage of the exciting electrons. In stellar spectra this process is illustrated on a grand scale, and just as convincingly as in the laboratory.

The next step was obviously to incorporate more detailed theoretical information about the spectral lines into the analysis. Progress here has been achieved through successive approximations as the theoretical notions have become more and more refined. First Unsöld on one hand and Stewart on the other showed how the coefficient of scattering predicted by classical dispersion theory would account for the wide wings of strong solar and stellar lines. Unsöld worked out in this way a programme for analysing the state of an atmosphere from the profiles of the spectral lines. It is true that the applicability of the method was originally, and still is, very limited; partly because so many things remain to be worked out in atomic theory; still more because the theory of transmission of radiation through a stellar atmosphere is only very crudely developed; most of all because the observational determination of line profiles is one of the most difficult tasks in astronomy, which in many cases will demand a wholly new instrumental equipment and technique.

The cruder procedure of observing only the total intensity of a line must therefore remain for a long time to come the principal way of approach. Here conditions are much less uncertain, as was shown by Minnaert and his collaborators. In the case of strong lines the intensity appears to be a simple function of the product of the number of atoms above the photosphere, the oscillatory strength of the line, and, finally, its damping constant. When the intensity of the line diminishes, the intensity becomes gradually more and more independent of the argument, passes through a minimum of

dependence, and when for still smaller concentrations of atoms the intensity begins to fall more rapidly, the damping constant no longer plays any part. In its place the intensity is influenced by the intrinsic width of the line which is due to thermal Doppler effect.

This description refers to an ideal case of a quiescent atmosphere. If the atmosphere is in some state of motion, conditions will be different, and it is here that the study of the lines reveals essentially new information. Rotating stars are thus recognized from their broad and diffuse lines with dish-shaped profiles. Expanding stars, as we observe them in novae, Wolf-Rayet stars, and *O* stars have very broad emission lines which are easy to recognize. Recently Struve and Elvey believe they have found what may be called 'boiling' stars, where the apparent thermal Doppler effect exceeds the one to be expected by a factor of the order ten, and this fact they have interpreted to mean that the atmospheres are in a state of irregular motion. The study of the Doppler effect in interstellar lines, so important for the study of galactic rotation, also belongs to this field. It is in fact not merely a question of a simple bodily displacement of the lines, as in the case of routine radial velocity measurements, but of an accumulated Doppler effect, in which the width of the line increases proportionately to its displacement from the zero position.

Another line of study in which the finer details of atomic theory play a considerable part was opened up by Bowen in 1927, when he removed the veil which up to that time had hidden the secrets of nebular radiation by showing conclusively that the previously unknown nebular lines must be due to forbidden transitions in the ionized atoms of oxygen and nitrogen and some other elements. It goes without saying that this discovery at once stimulated theoretical research into the quantitative theory of forbidden transitions, and astrophysicists became on the alert as regards other possibilities where such transitions might play a part.

Forbidden transitions entered the field of cosmical physics for the first time when McLennan discovered that the green auroral line is really due to a transition from a metastable state in the neutral oxygen atom. In our own atmosphere we have also other instances of forbidden transitions in the well-known red absorption bands of oxygen. Recently Adams and Dunham found forbidden bands of carbon dioxide in the spectrum of Venus. Still more sensational was

the discovery that the absorption bands in the spectra of outer planets could be interpreted as forbidden bands of methane and ammonia, which was brought out by the theoretical work of Wildt and the observations of Dunham.

The discovery by Menzel, Payne, and Boyce, that many lines in novae and *O* stars of previously unknown origin are due to forbidden lines in atoms of ordinary elements, also serves to show that astrophysics is just the field where forbidden transitions may be studied with advantage.

The story of the role of atomic theory in astrophysics has not been told in full, but enough has been said to indicate how it intervenes at strategic points, and how it promises in due time to reveal a full picture of the constitution of stellar atmospheres. Already we begin to perceive in outline the possibilities which are lying ahead. We know enough now to maintain with considerable confidence that the study of the atmospheres is not an isolated field. On the contrary, the structure of the atmospheres is so intimately connected with the internal structure of the stars, that it is possible to derive from the spectra such typical properties of the star in bulk as the mass and the luminosity, which again leads to a knowledge of the distance of the star in space. In fact, stellar spectroscopy has in this way already threatened to supplant many methods of classical astronomy, and is likely to do so much more in the future.

In the present book an attempt has been made to build up the theory of stellar spectroscopy in a systematic way from the elementary principles of quantum mechanics, ready for the interpretation of astrophysical facts. It is fully realized that this can scarcely be done except in the form of a programme for future work, but it has nevertheless been thought worth while to work out a programme of this kind.

I

ANALYTICAL DYNAMICS

1. The Equations of Lagrange and Hamilton

In this section we shall recall some salient features of the canonical form of the equations of motion first given by Hamilton,† and the integration theory of these equations developed by Hamilton† and Jacobi.‡ This theory not only forms the most convenient form of the dynamical equations, but also the most natural bridge from ordinary mechanics to quantum mechanics.

Consider a dynamical system consisting of r mass points. Denote the rectangular cartesian coordinates of the ith point by x_i, y_i, z_i, its mass by m_i, and the corresponding force by X_i, Y_i, Z_i. The Newtonian equations of motion of the system are then

$$m_i \ddot{x}_i = X_i; \quad m_i \ddot{y}_i = Y_i; \quad m_i \ddot{z}_i = Z_i \quad (i = 1, 2, \ldots, r),$$

where a dot denotes differentiation with respect to time. All affixes will be placed as subscripts for the present, while they have no tensorial meaning.

For various reasons it may be convenient to introduce other systems of coordinates in these equations, particularly so when the system involves constraints, so that the number of degrees of freedom of the system is smaller than the number of rectangular coordinates used in its specification, $3r$. We therefore introduce a set of generalized coordinates $q_1, \ldots, q_s$. We shall not here stop to consider the difference between holonomic and non-holonomic systems, which plays a considerable part in rational mechanics. We suppose that the particles are either moving freely in space, or that they are connected in such a manner that it will always be possible to find equations not involving differential relations between the new and the old coordinates.

In order to effect the transformation in question we multiply the Newtonian equations of motion by

$$\frac{\partial x_i}{\partial q_k}, \quad \frac{\partial y_i}{\partial q_k}, \quad \frac{\partial z_i}{\partial q_k}$$

respectively, add the resulting equations, and take the sum over all.

† 'On a general method in dynamics', *Trans. Roy. Irish Acad.*, 1834–5.

‡ *Vorlesungen uber Dynamik.*

The resulting equation is then

$$\sum_{i=1}^{r} m_i\left(\ddot{x}_i\frac{\partial x_i}{\partial q_k}+\ddot{y}_i\frac{\partial y_i}{\partial q_k}+\ddot{z}_i\frac{\partial z_i}{\partial q_k}\right)=\sum_{i=1}^{r}\left(X_i\frac{\partial x_i}{\partial q_k}+Y_i\frac{\partial y_i}{\partial q_k}+Z_i\frac{\partial z_i}{\partial q_k}\right). \quad (1)$$

The quantity $$Q_k=\sum_i\left(X_i\frac{\partial x_i}{\partial q_k}+Y_i\frac{\partial y_i}{\partial q_k}+Z_i\frac{\partial z_i}{\partial q_k}\right)$$

occurring on the right-hand side of the above equation is obviously the q_k-component of the force in the new coordinate system. We further denote the kinetic energy of the system by T:

$$T=\tfrac{1}{2}\sum_i m_i(\dot{x}_i^2+\dot{y}_i^2+\dot{z}_i^2).$$

On account of the identities

$$\frac{\partial x_i}{\partial q_k}=\frac{\partial \dot{x}_i}{\partial \dot{q}_k}$$

and

$$\frac{\partial \dot{x}_i}{\partial q_k}=\frac{d}{dt}\left(\frac{\partial x_i}{\partial q_k}\right)$$

we have

$$\frac{\partial T}{\partial \dot{q}_k}=\sum_i m_i\left(\dot{x}_i\frac{\partial x_i}{\partial q_k}+\dot{y}_i\frac{\partial y_i}{\partial q_k}+\dot{z}_i\frac{\partial z_i}{\partial q_k}\right)$$

and

$$\frac{\partial T}{\partial q_k}=\sum_i m_i\left\{\dot{x}_i\frac{d}{dt}\left(\frac{\partial x_i}{\partial q_k}\right)+\dot{y}_i\frac{d}{dt}\left(\frac{\partial y_i}{\partial q_k}\right)+\dot{z}_i\frac{d}{dt}\left(\frac{\partial z_i}{\partial q_k}\right)\right\},$$

which together with (1) gives the relation

$$\frac{d}{dt}\left(\frac{\partial T}{\partial \dot{q}_k}\right)-\frac{\partial T}{\partial q_k}=Q_k \qquad (k=1,2,\ldots,s). \quad (1\,a)$$

These equations were given for the first time by Lagrange.† If a potential energy function ϕ exists, so that

$$Q_k=-\frac{\partial \phi}{\partial q_k},$$

it is natural to introduce the difference between the kinetic energy and potential energy, the kinetic potential

$$L=T-\phi,$$

into equations (1 a), which then assume the form

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}_k}\right)-\frac{\partial L}{\partial q_k}=0 \qquad (k=1,2,\ldots,s), \quad (2)$$

since ϕ is independent of the velocities. According to the calculus

† 'Théorie de la libration de la lune. Deuxième section, *Nouveaux Mémoires de l'académie royale des sciences et belles-lettres de Berlin, année 1780.*

of variations, equations of this type are necessary conditions for the first variation of the integral

$$J = \int_{t_1}^{t_2} L\, dt$$

to vanish for arbitrary variations δq_k of the coordinates, which vanish at the limits of integration. This is Hamilton's variation principle.

To see this, assume first that the coordinates are varied without varying the time. We have then

$$\delta J = \int_{t_1}^{t_2} \sum_k \left(\frac{\partial L}{\partial q_k} \delta q_k + \frac{\partial L}{\partial \dot{q}_k} \delta \dot{q}_k \right) dt.$$

By partial integration of the second term in the bracket we find

$$\delta J = \sum_k \frac{\partial L}{\partial \dot{q}_k} \delta q_k \Big|_{t_1}^{t_2} + \int_{t_1}^{t_2} \sum_k \left\{ \frac{\partial L}{\partial q_k} - \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}_k} \right) \right\} \delta q_k\, dt.$$

But the first term on the right-hand side vanishes in virtue of the conditions at the limits. Since the variations are arbitrary otherwise, it follows that for δJ to vanish the quantity in brackets must vanish, which means that equations (2) hold.

Suppose next that the time is varied. Account must then be taken of the fact that the velocities $\dot{q}_k$ contain the time. We have then

$$\dot{q}_k + \delta \dot{q}_k = \frac{dq_k}{d(t+\delta t)} = \frac{dq_k}{dt}\left(1 - \frac{d\,\delta t}{dt} \right) = \dot{q}_k - \dot{q}_k \frac{d\,\delta t}{dt},$$

so that the variation suffered by the velocity $\dot{q}_k$ on varying the time is $-\dot{q}_k \dfrac{d\,\delta t}{dt}$.

For the time variation of J we have

$$\delta J = \int_{t_1}^{t_2} \left\{ \left(\frac{\partial L}{\partial t}\, \delta t + \sum_k \frac{\partial L}{\partial \dot{q}_k}\, \delta \dot{q}_k \right) dt + L\, \delta(dt) \right\},$$

which by the use of the expression for the variation in $\dot{q}_k$ becomes

$$\delta J = \int_{t_1}^{t_2} \left\{ \frac{\partial L}{\partial t}\, \delta t + \left(L - \sum_k \frac{\partial L}{\partial \dot{q}_k} \dot{q}_k \right) \frac{d(\delta t)}{dt} \right\} dt.$$

The same condition is imposed on δt as on the other variations, that it must vanish at the limits of integration. By partial integration it

follows then in the same way as before that the term

$$\left(L-\sum_k \frac{\partial L}{\partial \dot{q}_k}\dot{q}_k\right)\frac{d\,\delta t}{dt}$$

in the integrand may be replaced by

$$\frac{d}{dt}\left(\sum_k \frac{\partial L}{\partial \dot{q}_k}\dot{q}_k - L\right)\delta t,$$

without altering the integral. The variation of the time, δt, may now be taken outside the parenthesis, and the condition for the vanishing of the integral, no matter what δt is, becomes

$$\frac{\partial L}{\partial t}+\frac{d}{dt}\left(\sum_k \frac{\partial L}{\partial \dot{q}_k}\dot{q}_k - L\right) = 0.$$

But this equation is identically satisfied in virtue of equations (2). For multiplying these equations by $\dot{q}_k$ and summing, the above equation results by an easy transformation of the terms.

It may be noticed in passing that this last equation contains the principle of conservation of energy. If the potential energy is independent of the velocities, the expression in parenthesis is just equal to the total energy. For since the kinetic energy is a quadratic function of the velocities, the sum

$$\sum_k \frac{\partial L}{\partial \dot{q}_k}\dot{q}_k$$

will equal twice the kinetic energy. If the potential energy does not contain the time explicitly, the same is the case of L, and the equation reduces to

$$\frac{d}{dt}(T+\phi) = 0,$$

so that the total energy remains constant in time.

From equations (2) we pass over to Hamilton's form of the equations of motion in the following way. We introduce a new series of variables, the generalized momenta p_k, defined by the equations

$$p_k = \frac{\partial T}{\partial \dot{q}_k} = \frac{\partial L}{\partial \dot{q}_k}.$$

In rectangular coordinates these quantities usually reduce to the product of mass and the velocity components, so that the vector with components p_k represents the momentum as ordinarily defined

in mechanics. We have then in virtue of equations (2)

$$\delta L = \sum_k (\dot{p}_k \delta q_k + p_k \delta \dot{q}_k).$$

Consider next the function

$$H = \sum_k p_k \dot{q}_k - L.$$

The first variation of this function is easily brought to the form

$$\delta H = \sum_k (\dot{q}_k \delta p_k - \dot{p}_k \delta q_k)$$

on using the above expression of δL. This simply means, however, that the variables p_k, q_k satisfy the system of equations

$$\dot{q}_k = \frac{\partial H}{\partial p_k}; \qquad \dot{p}_k = -\frac{\partial H}{\partial q_k}, \tag{3}$$

which are just Hamilton's canonical equations of motion.

2. The Hamilton-Jacobi Integration Theory

The general idea underlying this theory is first to find the conditions to be satisfied by a transformation of variables which has the property not to alter the canonical form of these equations.

The next step is then to find a set of new variables which will remain constant during the motion of the system, so that they may be considered a set of integration constants. The relations between old and new variables will then just constitute the integrals of the differential equations which we are looking for.

Consider thus a change of variables in the Hamiltonian equations from p_k, q_k to P_k, Q_k say, the change being by hypothesis of such a kind as to leave the Hamiltonian equations invariant, that is to say

$$\dot{P}_k = -\frac{\partial \bar{H}}{\partial Q_k}; \qquad \dot{Q}_k = \frac{\partial \bar{H}}{\partial P_k}, \tag{4}$$

$\bar{H}$ being some function of P_k, Q_k, and perhaps the time t. This will evidently be the case if the kinetic potential, which enters the variation principle, when expressed in terms of P_k, Q_k is equal to the corresponding L expressed in p_k, q_k, *save for the possible addition of a term which is the total derivative with respect to time of a certain function,* which may be denoted by S. Any term of this latter kind may evidently be added to the kinetic potential in Hamilton's principle because the time integral of such a term depends upon the initial and the final variations of the coordinates and the time only, and those variations are supposed to vanish at the integration limits. The new

and old variables will therefore be connected by the differential relation

$$\sum_k (p_k \dot{q}_k - P_k \dot{Q}_k) - (H - \bar{H}) = \frac{dS}{dt}. \tag{5}$$

The function S may have four different forms. It may depend upon q_k, Q_k only, and not on p_k, P_k, or alternatively on p_k, Q_k; p_k, P_k; or q_k, P_k only. The dependence of S on the time, on the other hand, is by (5) always such that

$$\bar{H} = H + \frac{\partial S}{\partial t}. \tag{6}$$

As equation (5) stands S is supposed to be a function of q_k, Q_k only, and we have the following equations connecting the old and new variables

$$p_k = \frac{\partial S}{\partial q_k}; \qquad P_k = -\frac{\partial S}{\partial Q_k}. \tag{7}$$

The other forms of these equations are easily obtained by writing $p_k dq_k = d(p_k q_k) - q_k dp_k$, etc., in (5) and assimilating the total derivatives $d(p_k q_k)$ into dS.

From equations (4) we notice that the new variables will remain constant in time if the new Hamiltonian function $\bar{H}$ is independent of Q_k and P_k. Since this function by its definition involves an arbitrary constant, it involves no loss of generality to assume $\bar{H} = 0$, or by (6)

$$H + \frac{\partial S}{\partial t} = 0. \tag{8}$$

We have next to replace the p's in H by their equivalents $\partial S/\partial q_k$ according to (7), in consequence of which (8) becomes a partial differential equation for the determination of the transformation function S. Any solution of this equation will, of course, define a certain transformation. But for the transformation to represent the general solution of the dynamical problem in question, it must contain s arbitrary constants, which may be taken to represent all the Q_k. The corresponding P_k are then obtained by simple differentiation according to (7). Since both the Q_k and the P_k are constants of the motion, relations (7) will now represent the integrated equations of motion.

If H does not contain the time explicitly, we may write

$$\frac{\partial S}{\partial t} = -E,$$

where E is a constant, which in this case signifies the total energy of the system.

3. Motion in an Electromagnetic Field of Force

Consider a general electromagnetic field in empty space, defined by a scalar potential ϕ and a vector potential $\mathbf{A}$. These potentials are connected with the electric and magnetic field-strengths $\mathbf{E}$ and $\mathbf{H}$ by the equations of definition

$$\mathbf{E} = -\nabla\phi - \frac{1}{c}\frac{\partial \mathbf{A}}{\partial t},$$

$$\mathbf{H} = \operatorname{curl}\mathbf{A}.$$

The force on unit charge is

$$\mathbf{f} = \mathbf{E} + \frac{1}{c}\mathbf{V}\times\mathbf{H},$$

where c is the velocity of light, and the $\times$ denotes the vector product. The equation of motion of a particle of constant mass m and charge e becomes

$$m\dot{\mathbf{V}} = e\mathbf{E} + \frac{e}{c}\mathbf{V}\times\mathbf{H}.$$

Introducing the expressions for $\mathbf{E}$ and $\mathbf{H}$ into this equation, and carrying out some not too difficult transformations, it may be brought to the form

$$\frac{d}{dt}\left(m\mathbf{V} + \frac{e}{c}\mathbf{A}\right) = -\nabla\left(e\phi - \frac{e}{c}\mathbf{V}\mathbf{A}\right).$$

But the three components of this vector equation are just the Lagrangian equations corresponding to a kinetic potential

$$L = T - e\phi + \frac{e}{c}\mathbf{V}\mathbf{A}.$$

For a system of particles the kinetic potential is simply the sum of the kinetic potentials of the individual particles.

We have above assumed the masses to be independent of the time and the velocities. Assume now that the masses depend on velocity in the manner required by relativity theory

$$m = \frac{m_0}{\sqrt{(1 - V^2/c^2)}}.$$

The equation of motion of a particle moving in an electric field is then

$$\frac{d}{dt}(m\mathbf{V}) = \frac{d}{dt}\left(\frac{m_0\mathbf{V}}{\sqrt{(1 - V^2/c^2)}}\right) = -\nabla e\phi$$

and corresponds to a kinetic potential

$$L = -m_0 c^2 \sqrt{(1-V^2/c^2)} - e\phi.$$

The only difference from the non-relativistic case consists in the replacement of the kinetic energy by $-m_0 c^2 \sqrt{(1-V^2/c^2)}$, and this holds evidently also when magnetic forces are present.

This does not mean that the relativistic expression of kinetic energy is $-m_0 c^2 \sqrt{(1-V^2/c^2)}$, as will be apparent when passing on to the Hamiltonian function of the problem. We now include a magnetic field as well, so the kinetic potential of the problem is

$$L = -m_0 c^2 \sqrt{(1-V^2/c^2)} - e\phi + \frac{e}{c}\mathbf{VA}.$$

The corresponding generalized momentum is easily found to be

$$\mathbf{p} = \frac{m_0 \mathbf{V}}{\sqrt{(1-V^2/c^2)}} + \frac{e}{c}\mathbf{A},$$

and the Hamiltonian function becomes

$$H = \mathbf{pV} - L = \frac{m_0 c^2}{\sqrt{(1-V^2/c^2)}} + e\phi.$$

Since this is the total energy, the kinetic energy must be

$$T = m_0 c^2 \left(\frac{1}{\sqrt{(1-V^2/c^2)}} - 1 \right),$$

where the constant term has been adjusted so as to make T vanish with $\mathbf{V}$. This is permissible since ϕ contains an arbitrary constant.

Expressed in terms of the momentum $\mathbf{p}$, the mass of the particle becomes

$$m = m_0 \sqrt{\left\{ 1 + \frac{1}{m_0^2 c^2} \left(\mathbf{p} - \frac{e}{c}\mathbf{A} \right)^2 \right\}},$$

and consequently the Hamiltonian function becomes

$$H = m_0 c^2 \sqrt{\left\{ 1 + \frac{1}{m_0^2 c^2} \left(\mathbf{p} - \frac{e}{c}\mathbf{A} \right)^2 \right\}} + e\phi - m_0 c^2. \tag{8 a}$$

Removing the square root in this equation gives

$$m_0^2 c^2 + \left(\mathbf{p} - \frac{e}{c}\mathbf{A} \right)^2 = \left(\frac{1}{c} H + m_0 c - \frac{e}{c}\phi \right)^2.$$

The corresponding Hamilton-Jacobi equation is obtained by introducing

$$\mathbf{p} = \nabla S, \qquad H + m_0 c^2 = -\frac{\partial S}{\partial t}$$

in the above equation, giving as a result

$$\left(\nabla S - \frac{e}{c}\mathbf{A}\right)^2 + \left(\frac{i}{c}\frac{\partial S}{\partial t} + i\frac{e}{c}\phi\right)^2 + m_0^2 c^2 = 0.$$

This equation displays the four-dimensional symmetry of space and time, so characteristic of the relativity theory. The quantity $\frac{i}{c}\frac{\partial S}{\partial t}$ plays the part of a fourth component of the four-dimensional momentum vector:

$$\nabla S + \frac{i}{c}\frac{\partial S}{\partial t},$$

while the magnetic and electric potentials together form the four-dimensional vector potential

$$\mathbf{A} - i\phi, \qquad i = \sqrt{(-1)}.$$

II

STATISTICAL MECHANICS

4. The Theory of Gibbs

In ordinary mechanics the task of solving a given problem runs much as follows. At first the nature of the system is specified, which is done by giving its particular Hamiltonian function. Each degree of freedom of the system has then its own specified coordinate and momentum, the rate of change of which is determined by the corresponding canonical equations. By solving these equations the coordinates and momenta are found as explicit functions of the time.

It is frequently of advantage to imagine the state of the system at a given time represented by a point in a space of $2s$ dimensions, s being the number of degrees of freedom of the system, the rectangular coordinates of this space being the canonical coordinates and momenta. This so-called *phase space* is then a union of the *coordinate space* and the *momentum space*. From this point of view the solution of a mechanical problem consists in determining the trajectory of the representative point of the system in its particular phase space as a function of the time. This is what may be called the microscopic point of view in mechanics, it being tacitly understood that the investigator is able to follow up the very smallest details in the motion of the system under observation, and that he cares to do so.

In the mechanical theory of heat we have a case where the investigator does not care to follow up all details of the motion, while he is intensely interested in certain features of the motion which are apt to repeat themselves very frequently. To satisfy this need *the statistical point of view* in mechanics was developed by the leading physicists of the nineteenth century: Maxwell, Clausius, Boltzmann, and Gibbs, to name the most outstanding pioneers in this field.

We then imagine a very large number of systems, mechanically identical in nature but different with regard to *phase* (i.e. configuration and velocities). We assume the representative points of the systems continuously distributed through phase space, so that the probability for a member of the system to have its generalized coordinates inside the intervals q_1 to $q_1+dq_1,\ldots,p_s$ to p_s+dp_s may be represented by an expression of the form

$$d\omega = P\,dW \qquad (dW = dq_1,\ldots,dp_s).$$

The quantity P is called the *coefficient of probability*, and the integral

$$\omega = \int P dW$$

taken for arbitrary integration limits is called *probability in phase*, which is natural, since it measures the probability that an arbitrarily selected system is to be found within the assigned limits of phase space.

The numerical value of this integral is by definition independent of the time if the integration limits are moving with the velocities of the adjacent systems, it being remembered that the component velocities are now $\dot{q}_1, \dot{q}_2, \ldots, \dot{p}_s$. In order that no system be created or annihilated during the motion, the coefficient of probability must satisfy the same equation of continuity as the density in hydrodynamics, which in this case becomes

$$\frac{\partial P}{\partial t} + \sum_k \left\{ \frac{\partial}{\partial q_k}(P\dot{q}_k) + \frac{\partial}{\partial p_k}(P\dot{p}_k) \right\} = 0. \tag{9}$$

For constant P this equation reduces to

$$\sum_k \left(\frac{\partial \dot{q}_k}{\partial q_k} + \frac{\partial \dot{p}_k}{\partial p_k} \right) = 0,$$

which is automatically fulfilled in virtue of the canonical equations of motion. Hence it follows that the integral

$$W = \int dq_1, \ldots, dq_s dp_1, \ldots, dp_s$$

is also independent of the time, provided the integration limits are moving with the adjacent systems through phase space. This is the main content of Liouville's theorem. The importance of this theorem for statistical mechanics is due to the fact that it indicates a rational way of defining statistical probability. Since equal volumes of phase space remain equal during the motion, we may take the *a priori* probability of a certain volume of phase space proportional to this volume, without any appended variable factor. This will no longer be the case if non-canonical coordinates are used.

In order to have a statistical theory of thermodynamics the coefficient of probability must be given as a function of the coordinates in phase space. In statistical equilibrium ($\partial P/\partial t = 0$) equation (9) reduces to

$$\sum_k \left(\dot{q}_k \frac{\partial P}{\partial q_k} + \dot{p}_k \frac{\partial P}{\partial p_k} \right) = 0.$$

This simply means that P is a function that does not change during

the motion of the system. From mechanics we know a few functions of this kind, which are the same for all conservative systems, viz. the total energy, and each rectangular component of total linear momentum and angular momentum. Since thermodynamics makes the express assumption that the equilibrium shall be independent of special properties of a system, we must assume the probability coefficient of an assembly of systems in thermodynamic equilibrium to be a function of the energy and the component momenta of the system in question. P must further be single-valued and real and fulfil the condition that the total probability is unity:

$$\int_{\text{phases}}^{\text{all}} P\,dW = 1.$$

But for the rest P is arbitrary, and it is scarcely possible to give an *a priori* proof of its functional form. Gibbs showed, however, that for a system at rest the so-called canonical distribution, which is given by

$$P = e^{(\psi - E)/kT},$$

is particularly suitable for a statistical interpretation of thermodynamics, and it satisfies at the same time the more trivial conditions enumerated above. Here E is the total energy, T is the absolute temperature, and k the gas constant relative to a single atom, the so-called Boltzmann's constant; ψ corresponds to the free energy of thermodynamics, and is determined by

$$e^{-\psi/kT} = \int_{\text{phases}}^{\text{all}} e^{-E/kT}\,dW,$$

since $\int P\,dW = 1$. The average value of $k \log P$ corresponds to the entropy S:

$$S = k \int_{\text{phases}}^{\text{all}} P \log P\,dW = k \int_{\text{phases}}^{\text{all}} e^{(\psi - E)/kT} \frac{\psi - E}{kT}\,dW.$$

The pressure is then given by the ordinary thermodynamic relation

$$p = -\frac{\partial \psi}{\partial V},$$

where V signifies the volume in ordinary three-dimensional space. When the number of degrees of freedom is very large, such as is always the case when the system consists of atoms or molecules, it is easy to show that the kinetic and potential energy of the system will remain practically constant, in as much as the fluctuations will be

of the order of kT, the average energy of a single particle performing linear harmonic oscillations. These are the principal features of thermodynamic statistics in the form given by Gibbs.

Gibbs's 'canonical distribution' is essentially a special way of introducing the Maxwell-Boltzmann distribution law into molecular statistics. Applied to a swarm of non-interacting particles in a box the canonical distribution becomes identical with the Maxwellian distribution law. Adding a field of force, gravity for example, leads to the Maxwell-Boltzmann distribution law, and so on. The only difference between these different forms of statistical theories lies therefore in the particular mathematical machinery in use, and in the general background of physical arguments leading up to the mathematical formulae.

5. Dissociative Equilibrium in Gaseous Systems.

We shall now apply the theory of Gibbs to the dissociation in a gas, a problem of fundamental importance in astrophysics. We consider a mixture of r different elements in a gaseous state, enclosed in a volume V. These atoms may enter into chemical combinations in many different ways, say p different types. Let, further, x_k be the number of molecules of the kth type and m_i^k the number of atoms of the ith type which enter in a molecule of the kth type. We have then evidently r equations of the form

$$\sum_k x_k m_i^k = X_i \qquad (i = 1, 2, \ldots, r), \tag{10}$$

where X_i is the total number of atoms of the ith sort.

These definitions and assumptions imply certain restrictions on the dependence of the functional form of the total energy on the individual atoms. For the term 'a molecule' to have a definite meaning it is necessary that the forces acting between the molecules do not perturb the internal motions of the molecules in any serious manner, that is apart from the effect of collisions of a transient nature. These conditions are fulfilled in all cases of practical interest, and do not prevent the consideration of cases in which there are considerable deviations from the ideal gas laws. By introducing equations (10) we have, however, excluded the Fermi-Dirac quantum statistics from our considerations, since these statistics, according to Pauli,† require a canonical distribution of the atom numbers themselves.

† *Zeits. f. Physik*, **41** (1927), 81.

Molecules of the same kind, but in different quantum states, will be considered as identical, so that x_l refers to a definite atomic complex. The instantaneous values of a particular x_l are not determinable, since they will vary extremely rapidly from instant to instant. What we observe are the mean values, which in statistical language are given by

$$\bar{x}_l = \int x_l P \, dW,$$

where the integrations are to be extended to those parts of phase space only which correspond to the same number of atoms of different sorts, as demanded by relations (10).

We shall now try to bring the above integral to a form in which the dependence of the mean values $\bar{x}_l$ on macroscopically measurable quantities appears more clearly. For this purpose we first observe that the given integral may be written as a sum of partial integrals, each extended to that part of phase space in which the x_l preserve constant values, that is

$$\bar{x}_l = \sum x_l \int_{x_l} e^{(\psi - E)/kT} \, dW, \tag{10a}$$

where the index x_l appended to the integral sign indicates that the integration is to be carried out for constant x_l-values, and where the sum is extended to all combinations of values x_l which are compatible with (10). In the same manner the free energy of the system is determined by the *partition function*

$$e^{-\psi/kT} = \sum \int_{x_l} e^{-E/kT} \, dW.$$

In order to simplify the expression further it is necessary to consider the functional form of the energy of a molecule more closely. Let m be the mass of the molecule, and u, v, and w the rectangular velocity components of its centre of gravity, ϵ^i its internal energy, and ϵ^x its potential energy relative to all other molecules in the assembly. The system energy is consequently

$$E = \sum \{\tfrac{1}{2} m(u^2 + v^2 + w^2) + \epsilon^i + \epsilon^x\},$$

where the summation is to be extended to all molecules of the assembly. Due to the assumption about the energy mentioned above, ϵ^i is independent of the coordinates of the centre of gravity of the molecule in question, while ϵ^x is independent of the coordinates

determining the internal state of the atom. In consequence of this each integral

$$W_{x_l} = \int_{x_l} e^{-E/kT}\, dW$$

may be split up into independent factors. First of all each molecule contributes a factor of the form

$$f = \int e^{-(\frac{1}{2}m(u^2+v^2+w^2)+\epsilon^l)/kT} m^3\, dudvdwdW',$$

where dW' refers to the internal degrees of freedom of the molecule. The integration over the velocities of the centre of gravity is easily performed. In fact, by the theory of the error integral we have

$$m^3 \int e^{-\frac{1}{2}m(u^2+v^2+w^2)/kT}\, dudvdw = (2\pi mkT)^{\frac{3}{2}}.$$

Hence

$$f = (2\pi mkT)^{\frac{3}{2}} f',$$

where a dash indicates that the partition function is only relative to the internal state of the molecules.

Next comes a factor relative to the coordinates of the centre of gravity of the molecules

$$F = \int e^{-E^x/kT}\, dxdy\ldots \qquad (E^x = \sum \epsilon^x),$$

where E^x denotes the total interaction energy of the molecules. This latter partition function may be simply expressed by the average value of the interaction energy. We have, in fact

$$\overline{E^x} = \frac{\int e^{-E^x/kT} E^x\, dxdy\ldots}{\int e^{-E^x/kT}\, dxdy\ldots} = kT^2 \frac{\partial \log F}{\partial T},$$

and hence

$$\log F = -\int_T^\infty \overline{E^x} \frac{dT}{kT^2} + \log V^{\sum_l x_l}.$$

The integration limits adopted correspond to the requirement that the average interaction energy shall vanish for infinite temperature.

Finally, it is to be remembered that by permutation of atoms in different molecules new states of the systems are obtained, states which are not counted in the preceding integrations. The total number of permutations of atoms of the ith kind is $X_i!$ Permuting all atoms *inter se* we obtain in all $\prod_i X_i!$ new states. But it must be remembered here that permutations within a homogeneous group of *molecules* do not really lead to new distinguishable states, and these permutations must consequently be removed. This number is

obtained by considering two different processes. First we permute all homogeneous molecules *inter se*. This gives in all $\prod_l x_l!$ permutations. We next permute all atoms of the same sort within each molecule. This gives an additional number $\prod_{il} (m_i^l!)^{x_l}$ of permutations. The final correct number of relevant permutations is consequently

$$\prod_{il} \frac{X_i!}{x_l!(m_i^l!)^{x_l}}.$$

Thus far we have made no particular assumptions about the numbers x_l. We are, however, particularly interested in the case when the x_l are large numbers, corresponding to actual conditions in a gas mixture. The calculation of the average values may then be considerably simplified, because expressions like W_{x_l} possess a very sharp maximum for a definite set of values of the x_l. By this reason the partition function $\sum W_{x_l}$ reduces approximately to the largest term of the series, and from (10 a) it follows that the average values of the numbers x_l will coincide with the set of values giving maximal probability (W_{x_l} = maximum). We shall not enter more particularly upon these questions here, but only remark that the proof is closely connected with an application of Cauchy's theorem according to Darwin and Fowler.† In fact, the gist of the method introduced by these latter authors may be considered to consist essentially of the proof of this equivalence of average and maximal values.

We have now got the following expression for W_{x_l}:

$$W_{x_l} = \prod_{il} \frac{X_i!}{x_l!} \left(\frac{f_l}{m_i^l!}\right)^{x_l} F,$$

and proceed to calculate its maximal value.

We consider the change in W_{x_l} when each x_l is replaced by $x_l + \delta x_l$, where $\delta x_l / x_l$ are very small quantities in comparison to unity. We may use Stirling's formula

$$\log x! = x \log x - x \ldots,$$

since the x_l in this case are assumed to be large quantities. Taking the logarithm we obtain

$$\log W_{x_l} = \sum_i \log X_i! + \sum_l \left\{ x_l \log \frac{f_l}{\prod_i m_i^l!} - x_l \log x_l + x_l \right\} + \log F.$$

† Darwin and Fowler, *Phil. Mag.* **44** (1927), 450, 823, or Fowler, *Statistical Mechanics*, pp. 24 and 25.

Hence $$\delta \log W_{x_l} = \sum_l \left\{ \log \frac{f_l}{\prod_i m_i^l!} - \log x_l + \frac{\partial \log F}{\partial x_l} \right\} \delta x_l.$$

We write this in the form

$$\delta \log W_{x_l} = \sum_l \left\{ \log \frac{f_l^*}{x_l} + u_l \right\} \delta x_l,$$

where $$f_l^* = \frac{f_l}{\prod_i m_i^l!} \quad \text{and} \quad u_l = \frac{1}{F} \frac{\partial F}{\partial x_l}.$$

The variations δx_l must satisfy the conditions that the number of atoms is constant for each element, that is

$$\sum_l m_i^l \delta x_l = 0 \qquad (i = 1, 2, 3, \ldots).$$

These relations must now be multiplied by Lagrangian factors λ_i and added to the expression of $\delta(\log W_{x_l})$. The variations δx_l may then be considered as independent. We find

$$\delta(\log W_{x_l}) = \sum_l \left(\log \frac{f_l^*}{x_l} + u_l + \sum_{i=1}^{r} m_i^l \lambda_i \right) \delta x_l.$$

The condition that δW_{x_l} shall vanish for all imaginable variations δx_l is satisfied only by reducing to zero the coefficient of each δx_l, which gives the following system of equations

$$x_l = f_l^* e^{\sum_{i=1}^{r} m_i^l \lambda_i + u_l}, \tag{11}$$

which, in connexion with equations (10), suffices for the determination of the quantities x_l, which from now on will be identified with the average values $\bar{x}_l$.

III

QUANTUM MECHANICS

6. The Wave Equation

THE point of view of statistical mechanics in the above sense represents something of an intermediate stage between ordinary mechanics and quantum mechanics. When the thermodynamical statistician disregards the motion of the individual atoms or molecules and focuses his attention on the recurrent characteristics of the motion, it is only because of a sort of physical and mental laziness, and not because he entertains any doubt about the abstract applicability of mechanical laws to the motion of the ultimate particles of matter. On the other hand, the point of view of quantum theory, which has emerged from a mass of experiments and theoretical discussions, is that the notion of the orbit of a molecule or an electron is not to be understood literally as an observable entity. It is not more of an ultimate character than the average values of thermodynamic statistics. To bring this point of view forcefully to the mind we therefore in modern quantum theory build up a theoretical structure in which the statistical point of view is introduced explicitly, even though we talk about a single system only. The quantum theory does not attempt to construct a micro-dynamics behind the motion of the ultimate particles. It limits itself to the construction of a statistical machinery, suitable for the description of all observable properties of the atomic processes, which merges into ordinary mechanics on passing to macroscopic conditions.

To show the connexion between the methods of classical dynamics, statistical mechanics, and quantum mechanics we may proceed as follows. Imagine equation (9) to be integrated over all momenta. The sum involving the p's explicitly in this equation may then be neglected, since it is reasonable to exclude the occurrence of infinite momenta. Further we denote by $\bar{\dot{q}}_l$ the average value of $\dot{q}_l$, averaged with respect to the momenta, and write

$$Q = \int P \, dp_1 \ldots dp_s.$$

The quantities Q and $\bar{\dot{q}}_1, \bar{\dot{q}}_2 \ldots$ will then satisfy the simplified equation of continuity

$$\frac{\partial}{\partial t} Q + \sum_l \frac{\partial}{\partial q_l} (\bar{\dot{q}}_l Q) = 0. \tag{12}$$

In the limit of classical mechanics this equation must become a direct consequence of the corresponding Hamiltonian equations of motion. The quantity Q in the above equation, the probability coefficient in coordinate space, must therefore be intimately correlated with the Hamiltonian function S. A glance at (12) shows that the quantum-mechanical equation to be satisfied by S must be of the second order, in contradistinction to the Hamilton-Jacobi equation which is of the first. This is because the Hamiltonian velocities are proportional to the gradient of S, which, when introduced into (12) gives an equation of the second order.

There are many ways leading from the equations of classical mechanics to quantum mechanics, but all of them have the common feature that they involve arbitrary assumptions, and cannot be carried out logically without an ample appeal to experimental data. Skipping this part we may be content to say that the upshot of the discussion† is that the transition may be performed by considering the Hamiltonian H as an *operator*, in the sense that we consider the momenta to represent differential operators, as follows:

$$p_l = \frac{h}{2\pi i}\frac{\partial}{\partial q_l},$$

and consider the partial and linear differential equation

$$\left(H + \frac{h}{2\pi i}\frac{\partial}{\partial t}\right)\psi = 0, \tag{13}$$

where ψ is the so-called wave function, or probability amplitude, the reason for these names being apparent in the following.

In the limit $h = 0$ the solutions of this equation will coincide asymptotically with those of the corresponding Hamilton-Jacobi equation for S, provided we set up the relation

$$S = \frac{h}{2\pi i}\log\psi \tag{14}$$

between S and ψ. By this relation it will be possible to find out the mechanical equivalents of the quantum-mechanical quantities, as the left side of the equation represents mechanics, and the right hand quantum theory.

† A very complete discussion of the theoretical background for the symbolic interpretation of the canonical momenta as gradient operators is given by Dirac in his book *Quantum Mechanics*, Oxford, 1935.

Denote the complex conjugates of ψ and H by ψ^* and H^* respectively. Then ψ^* will satisfy the equation

$$\left(H^* - \frac{h}{2\pi i}\frac{\partial}{\partial t}\right)\psi^* = 0. \tag{15}$$

Multiplying (13) on the left-hand side by ψ^* and (15) by ψ and subtracting, we find

$$\frac{\partial}{\partial t}(\psi\psi^*) + \frac{2\pi i}{h}(\psi^* H\psi - \psi H^*\psi^*) = 0.$$

Comparing this equation with (12) it appears that a one-one correspondence between the two equations is possible, provided the second term in the above equation reduces to the divergence of a certain vector. In that case the operator H is said to be of the self-adjoint form. In case of ordinary mechanics the Hamiltonian operator is self-adjoint by itself, as may be illustrated by the case of a single electron. We have then, in rectangular coordinates, μ and e denoting mass and charge of the electron, V the potential energy, and $\mathbf{A}$ the vector potential:

$$H = \frac{1}{2\mu}\left(\mathbf{p} - \frac{e}{c}\mathbf{A}\right)^2 + V.$$

Introducing here

$$\mathbf{p} = \frac{h}{2\pi i}\nabla,$$

we find

$$H = \frac{1}{2\mu}\left(\frac{h}{2\pi i}\nabla - \frac{e}{c}\mathbf{A}\right)^2 + V = -\frac{h^2}{8\pi^2\mu}\nabla^2 + \frac{ihe}{2\pi\mu c}(\mathbf{A}\nabla) + \frac{e^2}{2\mu c^2}\mathbf{A}^2 + V.$$

Hence it follows that

$$\begin{aligned} -\frac{\partial}{\partial t}(\psi\psi^*) &= \frac{2\pi i}{h}(\psi^* H\psi - \psi H^*\psi^*) \\ &= \operatorname{div}\left\{\frac{h}{4\pi i\mu}(\psi^*\nabla\psi - \psi\nabla\psi^*) - \frac{e}{\mu c}\psi\psi^*\mathbf{A}\right\}, \end{aligned} \tag{16}$$

it being remembered that $\operatorname{div}\mathbf{A}$ vanishes by the ordinary definition of the vector potential. Identifying now equation (12) with (16), it follows that the coefficient of probability Q is given by

$$Q = \psi\psi^*$$

and the velocity by

$$\mathbf{V} = \frac{h}{4\pi i\mu}\left(\frac{1}{\psi}\nabla\psi - \frac{1}{\psi^*}\nabla\psi^*\right) - \frac{e}{\mu c}\mathbf{A}. \tag{17}$$

A glance at (14) shows immediately that this is just what might have

been anticipated, since it gives the correct asymptotic transition to the case of ordinary mechanics. The definition of energy may, similarly, be taken over directly from ordinary mechanics via equation (13). For in the Hamiltonian theory of conservative systems energy is defined by

$$E = -\frac{\partial S}{\partial t}.$$

Introducing S from (14) we find the wave-mechanical expression of energy

$$E = -\frac{h}{2\pi i}\frac{1}{\psi}\frac{\partial \psi}{\partial t}.$$

Consider the case when the Hamiltonian does not involve the time explicitly. A particular solution of (13) is then

$$\psi = u(q)e^{-2\pi i E t/h}, \tag{18}$$

where E is a constant and $u(q)$ a function of the coordinates only. Comparing these two expressions we see that E in the latter equation is just equal to the energy of the system. The wave function of a conservative system in a definite energy state is thus seen to be simply harmonic in the time.

7. The Quantum Conditions

Equation (13) is the fundamental equation of wave mechanics. The quantum conditions to be imposed upon the solutions of this equation consist in requiring the solutions to be *uniform*, *single-valued*, and *quadratically integrable*. The first of these conditions is common to all functions applicable to real physical problems. The second is practically identical with the earlier quantum conditions of Sommerfeld, Schwarzschild, etc., as will be realized by an inspection of (14). For these earlier conditions demanded that the increase suffered by S when a particular coordinate performed a complete cycle of its possible variations shall be an *integral* multiple of Planck's constant. But this is just the necessary condition that ψ shall be single-valued, provided S and ψ are connected by equation (14). The third condition is necessary for $\psi\psi^*$ to play the part of a true probability coefficient, because the integral over $\psi\psi^*$ extended to the whole of configuration space must give the total number of systems in the assembly under consideration, and it is convenient for the theory to keep this number finite. It is easy to see that this number remains constant during the motion. If, namely, equation (16) is

integrated over all space, the right-hand expression may be transformed into an integral over a surface lying entirely at infinity. Considering material systems of finite dimensions, the flux of probability may be assumed to vanish so rapidly on approaching infinity that this integral may also be neglected. Hence it follows that

$$\frac{\partial}{\partial t}\int \psi\psi^* \, d\tau = 0,$$

where the integral is extended to infinite configuration space.

In this equation is hidden a relation of great importance, which may be brought out at once. Assume that ψ_k and ψ_l correspond to two *different* energy values, say E_k and E_l respectively. Repeating the steps we find in the same way that

$$\frac{\partial}{\partial t}\int \psi_k \psi_l^* \, d\tau = 0.$$

For a conservative system the wave function is given by (18), so that the above equation becomes

$$(E_k - E_l)\int \psi_k \psi_l^* \, d\tau = 0.$$

In so far as E_k differs from E_l, ψ_k and ψ_l^* will thus be *orthogonal* to each other, as expressed by the equation

$$\int \psi_k \psi_l^* \, d\tau = 0 \qquad (E_k \neq E_l).$$

If, however, E_k is equal to E_l, the orthogonality no longer holds. This may mean that ψ_k is equal to ψ_l, but they may also be different. In the latter case the system is said to be *degenerate* and the *degree of degeneracy* of the state in question is equal to the number of different wave functions which correspond to the same energy value E_k.

The orthogonal properties of the wave functions disclosed above lie at the base of the perturbation method of solving wave-mechanical problems, as we shall meet it at different places.

8. The Einstein Probability Coefficients and the Theory of Thermal Radiation

The evolution of the system in time is, in the present theory, in most cases satisfactorily described by giving certain coefficients, which give the probability that a system initially in a definite quantum state i shall in unit time suffer a transition to a state k. In

astrophysical applications the case is of particular importance when collisions between the atoms may be neglected, and the only external influence is represented by a radiation field. If the transition in question transfers the atom from a lower to a higher energy level, the transition coefficient will have the form

$$a_{ik} = B_{ik}\rho(\nu) \qquad (E_i < E_k),$$

where $\rho(\nu)$ is the energy density of radiation of the external field, calculated for the frequency ν and per unit frequency interval. The frequency ν is obtained from the energy difference between the two states by the frequency relation

$$h\nu = E_k - E_i.$$

It is assumed that the radiation field is uniform over the spectral line in question, so that the natural width of the line is irrelevant to the problem. The coefficient B_{ik} will depend on the structure of the atom in the states i and k, but is independent of other states of the atom.

For the converse transition we have

$$a_{ki} = A_{ki} + B_{ki}\rho(\nu).$$

Here A_{ki} is called the coefficient of spontaneous transitions, because it gives the rate of transitions when external influences are absent. The coefficient B_{ki} is analogous to B_{ik}, and is called the coefficient of stimulated emission.

The first attempt to describe the absorption and emission process by such coefficients was made by Einstein,† who also derived some important relations between the A- and the B-coefficients. Einstein showed further how it is possible by simple considerations of this kind to arrive at a theory of the intensity of thermal radiation.

For this purpose it is necessary to know the number n_k of atoms in a quantum state k in a state of thermal equilibrium. This is obtained by a direct application of Gibbs's canonical distribution, by partitioning phase space into cells, and coordinating one cell to each quantum state. Denoting the volume of the cell which is coordinated to the kth quantum state by σ_k, and the energy of this state by E_k, the probability of finding an atom in this state is, according to Gibbs,

$$\sigma_k e^{(\psi - E_k)/kT},$$

where ψ is the free energy of the assembly. If there are n atoms in all,

† *Physikal. Zeits.* **18** (1917), 122.

the number in the kth state is accordingly

$$n_k = n\sigma_k e^{(\psi - E_k)/kT}.$$

In a state of thermal equilibrium it is reasonable to expect that the number of transitions from a state i to a state k is equal to the number of transitions the other way, or that any cycle of the two directions of transitions has the same probability.

This is the principle of detailed balancing. Although it cannot be said that this principle is proved to be a mathematical necessity for thermodynamic equilibrium to be possible, it is natural to narrow the conditions down to this statement, and thus far no contradictions have been encountered on this account. Then for two states i and k,

$$n_i a_{ik} = n_k a_{ki},$$

or, introducing the above expressions of the transition probabilities

$$n\sigma_k e^{-E_k/kT}(A_{ki} + B_{ki}\rho(\nu)) = n\sigma_i e^{-E_i/kT} B_{ik}\rho(\nu).$$

Writing $E_k - E_i = h\nu$ and solving for $\rho(\nu)$ we find

$$\rho(\nu) = \frac{A_{ki}/B_{ki}}{(\sigma_i B_{ik}/\sigma_k B_{ki}) e^{h\nu/kT} - 1}.$$

It is clear that $\rho(\nu)$ must approach infinity when T approaches infinity, so that we must have

$$\sigma_i B_{ik} = \sigma_k B_{ki}.$$

It will be proved later (cf. § 30) that the ratio A_{ki}/B_{ki} is given by

$$\frac{A_{ki}}{B_{ki}} = \frac{8\pi h\nu^3}{c^3},$$

in consequence of which the above formula for the energy density assumes the form

$$\rho(\nu) = \frac{8\pi h\nu^3}{c^3}(e^{h\nu/kT} - 1)^{-1},$$

which was first given by Planck.†

9. Wave-mechanical Calculation of B_{ik}

We shall now pass on to the actual calculation of the transition coefficients. It is then sufficient to calculate one of the three coefficients A_{ki}, B_{ki}, and B_{ik} as the remaining two may then be calculated from the above relations. But it is also of interest to calculate all three coefficients independently, and thus be able to confirm the above relations. Provisionally, however, we shall adopt the first

† *Ann. d. Physik*, **4** (1901), 553.

procedure, and begin with the calculation of B_{ik}. We shall postpone the refined theory for later consideration, and first consider those features of the problem which are inherently connected with the transitions between the stationary states of an atom in a field of radiation. The forces exerted by the radiation on the atom will in all practical cases be very small in comparison with the inner forces of the atom, and it is hence sufficient to solve the problem to a first approximation only. In the more consistent form of the theory, to be considered later, the atom *and* the radiation field are considered as a single conservative mechanical system. Hence also the radiation field is quantized, and absorption and emission of radiation is considered as an interaction of two mechanical systems. For the moment, we shall, however, take a less correct, but more convenient, view, in that we consider the radiation field merely as a small perturbing external field of force, which is not subjected to quantization. We consider for simplicity an atom containing one electron only, the generalization to any number of electrons being obvious in all cases. The wave equation of the atom is then

$$\left\{\frac{1}{2\mu}\left(\mathbf{p}-\frac{e}{c}\mathbf{A}\right)^2+V+\frac{h}{2\pi i}\frac{\partial}{\partial t}\right\}\psi = 0,$$

where $\mathbf{p}$ is the symbolic momentum $\frac{h}{2\pi i}\nabla$ of the electron, μ its mass, and e its electric charge. $\mathbf{A}$ is the vector potential of the light wave and V the potential energy of the electron in the atom. The neglect of the scalar potential of the light wave is defensible because the waves carry no electricity.

When the term in $\mathbf{A}$ is included, the system is not conservative, but becomes so when $\mathbf{A}$ is neglected. Let

$$\phi_k \qquad (k = 1, 2, \ldots)$$

be the characteristic functions of the atom for this latter case. The most general solution of the problem is then

$$\psi = \sum_k a_k \phi_k, \tag{19}$$

where the a_k are a series of constants. We assume the characteristic functions to be *normalized with respect to unity* by the equation $\int \phi_k \phi_k^* \, d\tau = 1$, which is convenient in most cases met with in the sequel.

When the vector potential differs from zero the wave equation

still admits a solution of the form (19) with the only difference that the coefficients a_k are now functions of the time. To prove this proposition we assume the perturbing term $\frac{e}{\mu c}(\mathbf{Ap})\phi_n$ (we neglect $\mathbf{A}^2$) to be expanded in an infinite series of the form

$$\frac{e}{\mu c}(\mathbf{Ap})\phi_n = \sum_k A_{nk}\phi_k. \tag{20}$$

The neglect of $\mathbf{A}^2$ is justifiable in the case under consideration at present, but it may become of importance in considering absorption by free electrons.

The coefficients of the expansion (20) are obtained by multiplying by ϕ_m^*, say, and integrating over the whole of configuration space. All terms with $E_k \gtrless E_m$ will then vanish owing to the orthogonality of the ϕ-functions, the only surviving term being that in ϕ_m. Thanks to our requirement that the solutions are normalized with respect to unity we find

$$A_{nm} = \frac{e}{\mu c}\int \phi_m^*(\mathbf{Ap})\phi_n \, d\tau.$$

Introducing this expression into the wave equation with the perturbation term retained, it reduces to

$$\sum_{nm} a_n A_{nm}\phi_m - \frac{h}{2\pi i}\sum_n \phi_n \frac{da_n}{dt} = 0.$$

Multiplying this relation by $\phi_{n'}^*$, and integrating over all configuration space we find

$$\frac{h}{2\pi i}\frac{da_{n'}}{dt} = \sum_n a_n A_{nn'}. \tag{21}$$

This proves the proposition, and gives the equations for determination of the a's.

Consider next the solution of these equations in successive approximations. For the present purpose we only require a solution valid for a short time interval—short in comparison with the mean life period of the stationary state in question. It is therefore natural to adopt as a first approximation

$$a_n = \alpha_n + \frac{2\pi i}{h}\sum_m \alpha_m A_{mn}^t,$$

where

$$A_{mn}^t = \int_0^t A_{mn}\, dt,$$

while α_n stands for a series of constants—the values of the a_n at the time $t = 0$. Introducing the above expressions into the series in (21) we find the second approximation to be

$$a_n = \alpha_n + \frac{2\pi i}{h} \sum_m \alpha_m A^t_{mn} - \frac{4\pi^2}{h^2} \sum_{mk} \alpha_k B^t_{mnk}, \tag{22}$$

where $$B^t_{mnk} = \int_0^t A_{mn} A^s_{km}\, ds.$$

Let us stop for a moment and recall the meaning of the quantities a_n. We have shown earlier that $\psi\psi^* d\tau$ is the number of systems of which the representative points are found in the volume element $d\tau$ of configuration space. Forming $\psi\psi^*$ from (19) and integrating over all configuration space we find, on remembering the conditions of normalization and the orthogonal properties of the wave functions, that

$$N = \int \psi\psi^*\, d\tau = \sum_n a_n a^*_n.$$

Since N is the total number of systems in the assembly, $a_n a^*_n$ must be the number of systems in the nth quantum state, which we shall denote by N_n. Denoting by a'_n and a''_n the first and second order terms in (22) we therefore find for the net number of systems which in the time interval t have entered the nth state

$$\Delta N_n = a_n a^*_n - \alpha_n \alpha^*_n = a'_n \alpha^*_n + \alpha_n a'^*_n + a''_n \alpha^*_n + \alpha_n a''^*_n + a'_n a'^*_n. \tag{23}$$

According to this formula, the induced transition between the stationary states will depend essentially upon the individual values of the constants α_n, and not solely upon the combinations $\alpha_n \alpha^*_n$ which give the initial number of atoms in the state. This is contrary to the original form of the quantum theory as it was formulated by Einstein.

In practical cases, however, the phases of the constants α_n will probably be distributed at random, and in this case the present theory gives the same result as that of Einstein. For any term $\alpha_n \alpha^*_m$ will then be just as likely to be negative as positive, except for the case $n = m$. The only surviving terms in (23) are then

$$\Delta N_n = \frac{4\pi^2}{h^2} \left\{ \sum_m N_m A^t_{mn} A^{t*}_{mn} - N_n \sum_m 2\mathrm{R} \int_0^t A_{mn} A^s_{nm}\, ds \right\},$$

where R means 'real part of'. A term linear in A^t_{nm} has here been omitted because it is clearly unable to produce any secular increase

in the number of atoms in any state, since it is periodic, practically with the frequency $(E_n - E_m)/h$.

For further progress it is now necessary to calculate the coefficient A_{nm}. By its definition

$$A_{nm} = \frac{h}{2\pi i}\frac{e}{\mu c}\int (\mathbf{A}\nabla\phi_n)\phi_m^* \, d\tau.$$

If the wave-length of the incident radiation is long in comparison with atomic dimensions, $\mathbf{A}$ will be sensibly constant within that part of space in which ϕ_n and ϕ_m^* differ appreciably from zero, and may be taken outside the integral sign. From the identity

$$J_x = \operatorname{div}(x\mathbf{J}) - x \operatorname{div} \mathbf{J}$$

we find, on assuming $\mathbf{J}$ to vanish sufficiently fast towards infinity,

$$\int \mathbf{J} \, d\tau = -\int \mathbf{r} \operatorname{div} \mathbf{J} \, d\tau.$$

Let us apply this equation to the flux of probability, which from (16) is

$$\mathbf{J}_{nm} = \frac{h}{4\pi\mu i}\{\phi_m^* \nabla\phi_n - \phi_n \nabla\phi_m^*\}$$

since $\mathbf{A}$ was neglected in the first approximation. The divergence of this $\mathbf{J}_{nm}$ is

$$\operatorname{div} \mathbf{J}_{nm} = -\frac{\partial}{\partial t}(\phi_n \phi_m^*).$$

We then find $\int (\phi_m^* \nabla\phi_n - \phi_n \nabla\phi_m^*) \, d\tau = \frac{4\pi\mu i}{h} \dot{\mathbf{P}}_{nm}$,

where we have written

$$\mathbf{P}_{nm} = \int \mathbf{r}\phi_n \phi_m^* \, d\tau. \tag{24}$$

By partial integration we find further

$$\int (\phi_m^* \nabla\phi_n - \phi_n \nabla\phi_m^*) \, d\tau = 2\int \phi_m^* \nabla\phi_n \, d\tau,$$

and hence finally

$$A_{nm} = -\frac{e}{c}\mathbf{A}\dot{\mathbf{P}}_{nm} = -\frac{2\pi i e}{c}\nu_{nm}\mathbf{A}\mathbf{P}_{nm}; \qquad \nu_{nm} = \frac{E_n - E_m}{h}.$$

The next task is the calculation of A_{nm}^t. We then write first

$$\mathbf{P}_{nm} = \mathbf{P}_{nm}^0 e^{-2\pi i \nu_{nm} t},$$

so that $$A_{nm}^t = -2\pi i \nu_{nm}\frac{e}{c}\mathbf{P}_{nm}^0 \int_0^t \mathbf{A} e^{-2\pi i \nu_{nm} t} \, dt.$$

The integral on the right-hand side is just the harmonic Fourier

amplitude of the vector potential, calculated for the time interval t and for the frequency ν_{nm}. From this harmonic amplitude of the vector potential we may pass over to the harmonic components of the electric field intensity $\mathbf{E}$. From the formula connecting the energy density U of a plane wave and the field intensity $\mathbf{E}$,

$$U = \mathbf{E}^2/4\pi,$$

we find the intensity of radiation per unit frequency range expressed by the vector potential. The relation in question is

$$I_{\nu_{nm}} = \frac{\nu_{nm}^2}{2c} \int_0^t \mathbf{A}(s) e^{-2\pi i \nu_{nm} s}\, ds \int_0^t \mathbf{A}^*(s) e^{2\pi i \nu_{nm} s}\, ds,$$

which gives $$A^t_{nm} A^{t*}_{nm} = \frac{8\pi^2 e^2}{c} (\mathbf{P}_{nm} \mathbf{P}^*_{nm})_A I_{\nu_{nm}},$$

the suffix A being added to mark that we are concerned with the component of $\mathbf{P}_{nm}$ parallel to the vector potential. A little consideration shows further that

$$2R \int_0^t A_{nm} A^s_{mn}\, ds = A^t_{nm} A^{t*}_{nm}.$$

In fact, the integral on the left is taken over just half the area of that on the right-hand side. In as much as the real part of the integrand is perfectly symmetric with respect to the two halves, the right-hand integral will be just twice as large as that on the left.

We may further average over all directions in space of the vector potential. We then find

$$(\mathbf{P}_{nm} \mathbf{P}^*_{nm})_A = \tfrac{1}{3} |\mathbf{P}_{nm}|^2$$

since all directions are equally probable. The rate of change of N_n then becomes finally

$$\Delta N_n = \sum_m (N_m - N_n) \frac{32\pi^4 e^2}{3h^2 c} |\mathbf{P}_{nm}|^2 I_{\nu_{nm}}.$$

The probability that a system in unit time will pass from the state m to the state n, or vice versa, under the influence of a radiation field is thus determined by the Einstein B-coefficients

$$B_{nm} = B_{mn} = \frac{32\pi^4 e^2}{3h^2 c} |\mathbf{P}_{nm}|^2. \tag{25}$$

If we use the energy density of radiation ρ_ν instead of the specific

intensity I_ν, the B-coefficients given above must be multiplied by $c/4\pi$. Having thus found the B-coefficients, the corresponding Einstein A-coefficients follow from the relation

$$A_{nm} = \frac{8\pi h\nu^3}{c^3} B_{nm},$$

which was mentioned earlier in this chapter. Later we shall, moreover, develop a more refined theory which gives the A-coefficients directly.

The vector $e\mathbf{P}_{nm}$ introduced in the above calculations is the quantum analogue of the electric moment, which in classical theory determines the radiation from an electric dipole. In fact, a classical dipole of electrical moment $e\mathbf{P}_\nu$ oscillating with a frequency ν radiates energy spontaneously at an average rate

$$\frac{2e^2}{3c^3}|\ddot{\mathbf{P}}_\nu|^2 = \frac{4e^2}{3c^3}(2\pi\nu)^4|\mathbf{P}_\nu^0|^2,$$

where $\mathbf{P}_\nu^0$ is the constant harmonic amplitude of $\mathbf{P}_\nu$. The quantum theory expression of the rate of spontaneous radiation from a system is, on the other hand, $h\nu_{nm}A_{nm}$. Expressing A_{nm} by B_{nm} by the Einstein relation given above, and B_{nm} by (25), we get

$$h\nu_{nm}A_{nm} = \frac{4e^2}{3c^3}(2\pi\nu_{nm})^4|\mathbf{P}_{nm}^0|^2.$$

Comparing these two expressions it will be realized that $\mathbf{P}_\nu^0$ corresponds directly to $\mathbf{P}_{nm}^0$. For this reason $e\mathbf{P}_{nm}$ will frequently be referred to as the electric moment associated with the transition $n \to m$ of the system.

IV

SOLUTIONS OF SPECIAL PROBLEMS

10. The Free Particle

THE simplest case of all is the motion of free particles in unlimited space. The wave equation then reduces to

$$-\frac{h^2}{8\pi^2\mu}\nabla^2\psi+\frac{h}{2\pi i}\frac{\partial\psi}{\partial t}=0.$$

Assume a solution of the form

$$\psi=\xi(x)\eta(y)\zeta(z)e^{-2\pi iEt/h}.$$

The three functions ξ, η, and ζ must each satisfy an equation of the form

$$\frac{d^2\xi}{dx^2}+\frac{8\pi^2\mu}{h^2}E_x\xi=0,$$

where E_x is the energy associated with the motion in the x-direction, so that the total energy E is given by

$$E=E_x+E_y+E_z.$$

Writing $\quad k_x=\sqrt{(2\mu E_x)}; \quad k_y=\sqrt{(2\mu E_y)}; \quad k_z=\sqrt{(2\mu E_z)},$

we have the particular solution

$$\psi=\text{const.}\,e^{-2\pi i(Et\pm\mathbf{kr})/h}. \tag{26}$$

The corresponding wave function in case of a shower of particles is equal to the product of such wave functions, one for each particle.

The constant vector $\mathbf{k}$ is equal to the momentum of the particle, as is easily realized by application of (17). The wave function of a free particle is thus a simple three-dimensional sine wave, the wavelength of which is given by

$$\lambda=h/|\mathbf{k}|=h/\mu v,$$

v being the velocity of the particle in question.

The above solution may also be adapted to the case of free particles enclosed in a rectangular box with perfectly reflecting walls. Since the particle is supposed to be completely imprisoned in the box, ψ must vanish at the walls of the box, and be zero outside. If a, b, and c are the lengths of the sides of the box in the x-, y-, and z-directions respectively, the momentum vector must be quantized in the following manner

$$k_x=l\frac{h}{2a}, \qquad k_y=m\frac{h}{2b}, \qquad k_z=n\frac{h}{2c},$$

where l, m, n are integers. The exponentials involving the space coordinates in (26) reduce to simple sines, and the resulting wave-function becomes

$$\psi = \text{const.}\, e^{-2\pi i Et/h} \sin l\pi \frac{x}{a} \sin m\pi \frac{y}{b} \sin n\pi \frac{z}{c},$$

with

$$E = \frac{h^2}{8\mu}\left(\frac{l^2}{a^2} + \frac{m^2}{b^2} + \frac{n^2}{c^2}\right).$$

It is a rather paradoxical result to find a moving particle represented by a harmonic wave train in the above way. A plane sine wave, extending to infinity on all sides, is a conception very different indeed from that of a material particle, which at each instance occupies a definite point of space.

Since a sine wave, as shown above, is associated with a definite momentum k, it is clear that a progressive wave which is limited in space will represent a particle the momentum of which is allowed a certain range of variation. If this range is Δk and the maximum amplitude of the wave is localized within a range Δx in the direction of propagation, it follows from rather simple considerations, which we do not care to reproduce here, that

$$\Delta k \Delta x \geqslant h.$$

When the locus of the particle is determined with certainty, the momentum is arbitrary, and vice versa. This is a particular case of the Heisenberg uncertainty principle, which states that the accuracy with which a given component momentum p and its associated canonical coordinate q of a system may be determined simultaneously is limited by the condition

$$\Delta p \Delta q \geqslant h.$$

Similarly, the determination of energy is associated with the accurate determination of time by the analogous relation

$$\Delta E \Delta t \geqslant h.$$

The thorough discussion of these relations is of fundamental importance in any attempt at a philosophical valuation of wave mechanics.†

† Cf. for instance W. Pauli, 'Die allgemeinen Prinzipien der Wellenmechanik', *Handb. d. Physik*, **24**, 1; or W. Heisenberg, *The Physical Principles of Quantum Mechanics*.

11. The Harmonic Oscillator

Assume the oscillator to have one degree of freedom. Let ν be the frequency of the oscillator and μ its mass. The potential energy is then

$$V = \tfrac{1}{2}\mu(2\pi\nu)^2x^2,$$

x being the distance of the oscillating particle from its position of equilibrium. The wave equation of the problem becomes

$$-\frac{h^2}{8\pi^2\mu}\frac{d^2\psi}{dx^2}+\tfrac{1}{2}\mu(2\pi\nu)^2x^2\psi+\frac{h}{2\pi i}\frac{\partial\psi}{\partial t} = 0.$$

Writing the solution in the form

$$\psi = u(x)e^{-2\pi iEt/h}$$

this equation assumes the form

$$\frac{h^2}{8\pi^2\mu}\frac{d^2u}{dx^2}+\{E-\tfrac{1}{2}\mu(2\pi\nu)^2x^2\}u = 0.$$

Transform to new variables y and w by the equation

$$y = kx; \qquad u = w(y)e^{-\frac{1}{2}y^2}; \qquad k = 2\pi\sqrt{(\mu\nu/h)}.$$

We then find immediately

$$w''-2yw'+(\lambda-1)w = 0; \quad \lambda = \frac{8\pi^2\mu E}{h^2k^2} = \frac{2E}{h\nu}; \quad w' = \frac{dw}{dy}. \tag{27}$$

As a tentative solution we assume an expansion of the form

$$w = \sum_n a_n y^n,$$

which, when introduced in (27), yields the following relation between the expansion coefficients

$$a_{n+2} = a_n\frac{2n+1-\lambda}{(n+1)(n+2)}.$$

For very large values of the index n the function defined in this way behaves approximately like e^{2y^2}, so that $u(x)$ will tend to infinity for very large values of x, *unless* the series expansion breaks off after a finite number of steps. For this to take place it is necessary that λ conforms to the quantum condition

$$\lambda = 2s+1 \qquad (s = 0, 1, 2,\ldots).$$

The corresponding quantum condition for the energy is found from (27) to be

$$E_s = (s+\tfrac{1}{2})h\nu. \tag{28}$$

The series $w(y)$ is a Hermite polynomial, the standard form of which is

$$H_s = (-1)^s e^{y^2}\frac{d^s}{dy^s}(e^{-y^2}).$$

Write

$$H_s(y) = (-1)^s\left[\frac{d^s}{dx^s}(e^{y^2-x^2})\right]_{x=y} = \left[\frac{d^s}{dx^s}(e^{y^2-(x-y)^2})\right]_{x=0},$$

so that the $H_s(y)$ is also defined by the Taylor expansion of its 'generating function'

$$\theta(y,x) = e^{y^2-(x-y)^2} = \sum_{s=0}^{\infty} H_s(y)\frac{x^s}{s!}.$$

Since $e^{y^2-(x-y)^2}e^{y^2-(z-y)^2} = e^{2y(x+z)-x^2-z^2} = \sum_{s,t} H_s(y)H_t(y)\frac{x^s z^t}{s!\,t!}$,

we have

$$\sum_{s,t}\frac{x^s z^t}{s!\,t!}\int_{-\infty}^{\infty} H_s(y)H_t(y)e^{-y^2}\,dy = e^{2xz}\int_{-\infty}^{\infty} e^{-(y-x-z)^2}\,dy = \sqrt{\pi}\sum\frac{(2xz)^t}{t!}.$$

By equating coefficients of the same powers of x and z we find

$$\int_{-\infty}^{\infty} H_s(y)H_t(y)e^{-y^2}\,dy = \sqrt{\pi}\,s!\,2^t\delta_{st}; \qquad \delta_{st} = \begin{cases}0, & s \neq t\\ 1, & s = t.\end{cases}$$

But in as much as $H_s e^{-\frac{1}{2}y^2}$ is the wave function of our problem (disregarding the complex phase factor), the normalized wave function is

$$u_s(y) = \frac{H_s(y)e^{-\frac{1}{2}y^2}}{\sqrt{(2^s s!\sqrt{\pi}\,k^{-1})}}.$$

Selection Principle. The harmonic amplitude of the electric moment, which according to (25) determines the probability of a transition $s \to s'$, is for the present case of the form

$$P_{ss'} = \frac{1}{k\sqrt{(2^{s+s'}s!\,s'!\,\pi)}}\int_{-\infty}^{\infty} H_s H_{s'} e^{-y^2} y\,dy.$$

For the further reduction of this integral we need the following recurrence formula of the Hermite polynomials

$$yH_s = \tfrac{1}{2}H_{s+1} + sH_{s-1}. \tag{29}$$

The proof of this formula is obtained most easily by using the generating function which was given above:

$$\theta(y,x) = e^{y^2-(x-y)^2} = \sum_{0}^{\infty} H_s(y)\frac{x^s}{s!}.$$

Differentiating this formula with respect to x we find

$$2(y-x)\sum_{0}^{\infty} H_s(y)\frac{x^s}{s!} = \sum_{1}^{\infty}\frac{H_s(y)}{(s-1)!}x^{s-1}.$$

Or arranging the equation into a single power series in x

$$\sum_0^\infty \{2yH_s(y)-2sH_{s-1}(y)-H_{s+1}(y)\}x^s = 0.$$

Here the coefficient of each power of x must vanish, which is just the condition expressed by the recurrence formula (29).

Introduce now yH_s from (29) into the expression of $P_{ss'}$

$$P_{ss'} = \frac{1}{k\sqrt{(2^{s+s'}s!\,s'!\,\pi)}} \int_{-\infty}^{\infty} \{\tfrac{1}{2}H_{s'}H_{s+1}+sH_{s'}H_{s-1}\}e^{-v^2}\,dy.$$

Remembering the orthogonal properties of the wave functions it will be realized that the amplitude will vanish except for the case when s' is equal to $s\pm 1$. The corresponding values of the amplitude $P_{ss'}$ are then

$$P_{s,s+1} = \frac{1}{k}\sqrt{\{\tfrac{1}{2}(s+1)\}} = \sqrt{\left(\frac{h(s+1)}{8\pi^2\mu\nu}\right)}$$

and

$$P_{s,s-1} = \frac{1}{k}\sqrt{(\tfrac{1}{2}s)} = \sqrt{\left(\frac{hs}{8\pi^2\mu\nu}\right)},$$

the latter being obtained from the former by changing s into $s-1$, as was to be expected. The harmonic oscillator will thus only suffer transitions between adjacent states, and equation (28) shows that the frequency of the radiation absorbed or emitted in the process is just equal to the frequency of the oscillator itself.

12. Central Fields

The special case of a central field of force is of such importance in atomic theory as to merit separate consideration. First of all the hydrogen atom belongs to this class, and extensive calculations on the state of other atoms have been carried out by approximating to the actual field by an intermediate central field (Hartree's method of self-consistent fields).†

Let V be the potential energy of an electron in the central field under consideration. The wave equation of the problem is then

$$\nabla^2\psi + \frac{8\pi^2\mu}{h^2}(E-V)\psi = 0,$$

where E is the energy of the electron. Since V depends only upon the distance from some fixed centre, r say, it is natural to introduce polar coordinates r, θ, ϕ.

† *Proc. Cambridge Phil. Soc.* **26** (1928), 89.

The Laplacean operator in polar coordinates is

$$\nabla^2 = \frac{1}{r^2}\left\{\frac{\partial}{\partial r}\left(r^2\frac{\partial}{\partial r}\right)+\frac{1}{\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial}{\partial\theta}\right)+\frac{1}{\sin^2\theta}\frac{\partial^2}{\partial\phi^2}\right\}.$$

For those who know the tensor theory the proof of this formula is verified most simply by observing that the square of the line element in these coordinates is $ds^2 = dr^2+r^2\,d\theta^2+r^2\sin^2\theta\,d\phi^2$, so that the co- and contravariant components of the metric tensor are given by

$$g_{rr} = 1; \qquad g_{\theta\theta} = r^2; \qquad g_{\phi\phi} = r^2\sin^2\theta;$$
$$g^{rr} = 1; \qquad g^{\theta\theta} = r^{-2}; \qquad g^{\phi\phi} = r^{-2}\sin^{-2}\theta, \quad \text{and } g = r^4\sin^2\theta.$$

Introducing these values in the general formula

$$\nabla^2 = \frac{1}{\sqrt{g}}\frac{\partial}{\partial x^\mu}\left(\sqrt{g}\,g^{\mu\nu}\frac{\partial}{\partial x^\nu}\right),$$

the above formula for ∇^2 is obtained directly.

Since V depends on r only, it is clearly possible to split up the wave equation into several equations each depending on one coordinate only. Write

$$\psi = \xi(r)\eta(\theta)\zeta(\phi)$$

and introduce it into the wave equation. Denote by $l(l+1)$ and m^2 two constants of integration. The reason why these constants are given this particular form will become evident in the sequel. The wave equation may then be split up into three separate equations, as follows:

$$\frac{\partial}{\partial r}\left(r^2\frac{\partial\xi}{\partial r}\right)-\left\{l(l+1)-\frac{8\pi^2\mu}{h^2}r^2(E-V)\right\}\xi = 0, \tag{30a}$$

$$\frac{1}{\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial\eta}{\partial\theta}\right)-\left\{\frac{m^2}{\sin^2\theta}-l(l+1)\right\}\eta = 0, \tag{30b}$$

$$\frac{\partial^2\zeta}{\partial\phi^2}+m^2\zeta = 0. \tag{30c}$$

The first of these equations depends upon V, and can thus only be solved explicitly when the central field is specified. The functions η and ζ, on the other hand, are common to all central fields. We have, firstly,

$$\zeta = Ae^{-im\phi}+Be^{im\phi},$$

where A and B are constants. In order that ζ may be single-valued, m must be a positive or negative integer

$$m = \pm 0, 1, 2, \ldots.$$

We call *m the magnetic quantum number*, for reasons which will become obvious later on, while the number l is called the subsidiary, or the azimuthal, quantum number.

The equation satisfied by η is the equation for *Ferrers's associated Legendre function of order m and degree l*, which we shall denote by the symbol P_l^m in conformity to general usage. Let us recall briefly the connexion of this function with the ordinary Legendre function and its principal properties.

The differential equation of an ordinary Legendre function is obtained from (30 b) by dropping the term in m

$$\frac{1}{\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial\eta}{\partial\theta}\right)+l(l+1)\eta = 0.$$

Writing $$x = \cos\theta$$

this equation becomes

$$(1-x^2)\frac{\partial^2\eta}{\partial x^2}-2x\frac{\partial\eta}{\partial x}+l(l+1)\eta = 0. \tag{31}$$

Assuming an expansion in power series

$$\eta = \sum_n a_n x^n$$

and introducing it in (31), we find the following relation between the expansion coefficients

$$a_{n+2} = -\frac{(l-n)(l+n+1)}{(n+1)(n+2)}a_n.$$

The series has therefore the form $AP(x)+BQ(x)$, where A and B are arbitrary constants and

$$P = 1-l(l+1)\frac{x^2}{2!}+(l-2)l(l+1)(l+3)\frac{x^4}{4!}-\ldots$$

and $$Q = x-(l-1)(l+2)\frac{x^3}{3!}+(l-3)(l-1)(l+2)(l+4)\frac{x^5}{5!}+\ldots.$$

If l is not a positive integer, both series will diverge at $x = \pm 1$. If l is even, P will break off after a finite number of steps while Q will diverge. If l is odd, the conditions are reversed. In both cases the polynomial which remains finite may be written in the form

$$P_l = \frac{1}{2^l l!}\frac{d^l}{dx^l}(x^2-1)^l,$$

as was first shown by Rodrigues (1816). It is easily verified that (31) is of the self-adjoint form, and that hence the polynomials P_l, the Legendre polynomials, form an orthogonal sequence of functions.

Recurrence Formulae. Rodrigues' formula forms a convenient starting-point for the derivation of some recurrence formulae, which

are frequently useful in the theory of Legendre polynomials. Denoting differentiation with respect to x by a dash we thus find

$$P'_{l+1} = \frac{1}{2^l l!}\frac{d^{l+1}}{dx^{l+1}}\{(x^2-1)^l x\} = \frac{x}{2^l l!}\frac{d^{l+1}}{dx^{l+1}}(x^2-1)^l + \frac{l+1}{2^l l!}\frac{d^l}{dx^l}(x^2-1)^l,$$

which may be written in the form

$$P'_{l+1} - xP'_l = (l+1)P_l, \tag{32a}$$

which is the first recurrence formula.

Write next in Rodrigues' formula

$$(x^2-1)^l = (x^2-1)^{l-1}x^2 - (x^2-1)^{l-1}.$$

Then

$$2lP_l = \frac{1}{2^{l-1}(l-1)!}\left\{\frac{d^l}{dx^l}\{(x^2-1)^{l-1}x^2\} - \frac{d^l}{dx^l}(x^2-1)^{l-1}\right\}$$

$$= \frac{1}{2^{l-1}(l-1)!}\left\{x\frac{d^l}{dx^l}\{(x^2-1)^{l-1}x\} + l\frac{d^{l-1}}{dx^{l-1}}\{(x^2-1)^{l-1}x\}\right\} - P'_{l-1},$$

or

$$2lP_l = xP'_l + lP_l - P'_{l-1},$$

from which follows the second recurrence formula

$$lP_l = xP'_l - P'_{l-1}. \tag{32b}$$

Further, multiplying (32a) by x, replacing l by $l+1$ in (32b), and adding we get

$$(1-x^2)P'_l = (l+1)(xP_l - P_{l+1}). \tag{33}$$

Eliminating the derivatives between these equations we find the last recurrence formula

$$(l+1)P_{l+1} - (2l+1)xP_l + lP_{l-1} = 0. \tag{34}$$

Associated Legendre Polynomials. Differentiate the Legendre equation (31) m times and write

$$P_l^m = (1-x^2)^{\frac{1}{2}m}\frac{d^m P_l}{dx^m} = \frac{1}{2^l l!}(1-x^2)^{\frac{1}{2}m}\frac{d^{m+l}}{dx^{m+l}}(x^2-1)^l.$$

This expression is then found to satisfy equation (30b). It is usually called Ferrers's associated Legendre polynomial of degree l and order m. Just like the Legendre polynomials, the P_l^m form an orthogonal system for different values of the index l. For the sake of normalization it is important to know the expression of the integrated square of P_l^m, which then will also give the normalization constants of the Legendre polynomials. The formula in question is

$$N_{lm}^2 = \int_{-1}^{1}(P_l^m)^2\,dx = \frac{2}{2l+1}\frac{(l+m)!}{(l-m)!}.$$

In order to prove this formula we proceed by steps of partial integrations. From one of the P_l^m under the integral sign we select the differential quotient

$$\frac{d^m P_l}{dx^m} = \frac{1}{2^l l!} \frac{d^{m+l}}{dx^{m+l}}(x^2-1)^l,$$

and denote the remaining factor by F. As this factor will be a polynomial of the $(l+m)$th degree, differential coefficients higher than the $(l+m)$th will vanish. The integrated parts will vanish due to the factor $(1-x^2)$, and we arrive at an expression

$$N_{lm}^2 = \frac{(-1)^{l+m}}{(2^l l!)^2} \int_{-1}^{1} \left(\frac{d^{l+m}F}{dx^{l+m}}\right)(x^2-1)^l\, dx.$$

The differential coefficient under the integral sign is now a constant, and may be taken out. Writing

$$q_l = \int_{-1}^{+1} (x^2-1)^l\, dx = \int_{-1}^{1} x^2(x^2-1)^{l-1}\, dx - \int_{-1}^{1} (x^2-1)^{l-1}\, dx,$$

and integrating the first integral by parts, it appears that q_l satisfies the recurrence formula

$$q_l = -\frac{2l}{2l+1} q_{l-1}.$$

Since $q_0 = 2$ we find

$$q_l = (-1)^l \frac{2(2^l l!)^2}{(2l+1)!}.$$

We find further easily

$$\frac{d^{l+m}F}{dx^{l+m}} = (-1)^m \frac{(l+m)!(2l)!}{(l-m)!},$$

and hence finally

$$N_{lm}^2 = \frac{2}{2l+1} \frac{(l+m)!}{(l-m)!}. \tag{35}$$

By differentiation of (32 a) and (34) respectively, and multiplication by $(1-x^2)^{\frac{1}{2}m}$, the following two formulae are obtained after some calculation:

$$P_l^{m+1} - \frac{2m}{\sqrt{(1-x^2)}} xP_l^m + (l+1-m)(l+m)P_l^{m-1} = 0, \tag{36}$$

and

$$(l+1-m)P_{l+1}^m - (2l+1)xP_l^m + (l+m)P_{l-1}^m = 0. \tag{37}$$

Selection Rules. We have thus far found the characteristic function so far as it depends upon the polar angles, and this is sufficient for

the proof of the selection principles for the magnetic and azimuthal quantum numbers, and the calculation of the intensities of the transitions involving different jumps in l and m, but one and the same jump in the principal quantum number, which depends also on the radial characteristic function. These selection principles are as follows: The magnetic quantum number changes by ± 1 or zero in a jump, and the azimuthal quantum number changes by ± 1 only:

$$m' = \begin{matrix} m \pm 1 \\ m \end{matrix}; \qquad l' = l \pm 1.$$

The first of these follows immediately from the fact that $\psi \propto e^{im\phi}$. This makes the components of the electric moment perpendicular to the polar axis vanish except when $m' = m \pm 1$, while the component in the direction of the axis vanishes in all cases except $m' = m$. These calculations require no further elucidation.

The selection principle for l is more difficult to prove, as it requires knowledge of the properties of associated Legendre functions. Denote the rectangular components of the electric moment by M_x, M_y, and M_z, the orientation of the axes being the usual one. Performing the integration with respect to ϕ and leaving out the factor due to the radial characteristic function and common constant factors, we find

$$M_x + iM_y = \frac{1}{N_{lm} N_{l',m\pm1}} \int_{-1}^{+1} P_l^m \, P_{l'}^{m\pm1} \sqrt{(1-x^2)} \, dx, \tag{38a}$$

$$M_z = \frac{1}{N_{lm} N_{l',m}} \int_{-1}^{+1} P_l^m \, P_{l'}^m \, x \, dx, \tag{38b}$$

where N_{lm} and $N_{l,m\pm1}$ denote the normalization factors of the associated Legendre functions involved:

$$N_{l,m}^2 = \int_{-1}^{+1} (P_l^m)^2 \, dx, \qquad N_{l',m\pm1}^2 = \int_{-1}^{+1} (P_{l'}^{m\pm1})^2 \, dx.$$

As regards M_z, the problem is solved by introducing for $x P_l^m$ from the recurrence formula (37)

$$x P_l^m = \frac{1}{2l+1} \{(l+1-m) P_{l+1}^m + (l+m) P_{l-1}^m\}.$$

In virtue of the orthogonality of the P_l^m-functions it then follows that

l' is either $l+1$ or $l-1$ as stated by the principle. By a simple application of the formula (35) we find

$$(M_z)_{l'=l-1} = \sqrt{\left(\frac{(l+m)(l-m)}{(2l+1)(2l-1)}\right)},$$

$$(M_z)_{l'=l+1} = \sqrt{\left(\frac{(l+1+m)(l+1-m)}{(2l+1)(2l+3)}\right)}.$$

To prove that $M_x \pm iM_y$ obeys the same selection principle we reduce the order of the integrals as regards m by successive applications of the recurrence formula (36), eliminating P_l^{m+1} in case of a positive sign, and P_l^m in case of a negative sign. We then get one integral of the form M_z and one of the original form, but with m lowered by one unit. The same operation is then applied to this integral, and so on, until we arrive at an integral of the form

$$\int_{-1}^{1} P_l P_{l'}^1 \sqrt{(1-x^2)}\, dx = \int_{-1}^{1} (1-x^2) P_l \frac{dP_{l'}}{dx}\, dx.$$

But eliminating $(1-x^2)\dfrac{dP_{l'}}{dx}$ from this expression by the recurrence formula (33) we are led to integrals over $P_l P_{l'\pm 1}$, which obey the selection principle.

To find the analytical expression of $M_x \pm iM_y$ another procedure is preferable. We consider first the case $m' = m+1$, $l' = l+1$, and assume an expansion of the form

$$\sqrt{(1-x^2)} P_l^m = \sum_{s=m+1}^{l+1} a_s P_s^{m+1}, \tag{39}$$

where the a_s's are a set of constants. This means merely a rearrangement of the powers of x in a polynomial, which can always be performed. Due to the orthogonality properties of the P_l^m we find, on introducing this expansion in (38 a), that the integral reduces to

$$a_{l+1} N_{l+1,m+1}^2.$$

The coefficient a_{l+1} is simply found by comparing the coefficient of the highest power of x on both sides of (39), which gives the relation

$$a_{l+1} = \frac{1}{2l+1}.$$

Hence, finally, $M_x + iM_y = \sqrt{\left(\dfrac{(l+m+2)(l+m+1)}{(2l+1)(2l+3)}\right)}.$

the preceding theory of central fields suffices for the rotator in all respects.

Let us compare this formula with the ordinary mechanical expression of the energy of a rotator, $M^2/2A$, where M is the angular momentum. We find at once

$$\sqrt{\{j(j+1)\}} = \frac{2\pi}{h} M.$$

For large values of j this quantity thus signifies the angular momentum in units of $h/2\pi$, and it is easy to see that the subsidiary quantum number l has the same significance in the general case of a central field.

14. The Balmer Formula

Consider next the problem of an electron moving in a Coulomb field of force so that the potential energy is given by

$$V = -\frac{Ze^2}{r},$$

where Z is the atomic number of the nucleus. Equation (30a) for the radial wave function then assumes the form

$$(r^2\xi')' - \left\{l(l+1) - \frac{8\pi^2\mu}{h^2} r^2 \left(E + \frac{Ze^2}{r}\right)\right\}\xi = 0. \tag{41}$$

When r approaches infinity, ξ will satisfy approximately the reduced equation

$$(r\xi)'' + \frac{8\pi^2\mu}{h^2} E(r\xi) = 0.$$

If E is positive, this equation defines ξr as a trigonometric function, which remains finite for all values of r, so that ξ will approach zero when r approaches infinity. For small values of r the functional form of ξ will depend essentially on the term $l(l+1)/r^2$, so that the question whether ξ is finite and uniform or not depends on l but not on E. This means that all positive values of E are admissible. A positive E means mechanically that the electron is moving in a hyperbolic orbit, and we have the expected result that free electrons may take on a continuous range of energy values.

When E is negative, conditions are different, since $r\xi$ will approach towards an exponential function as r approaches infinity:

$$r\xi = Ae^{\kappa r} + Be^{-\kappa r}, \qquad \kappa = \sqrt{\left(-\frac{8\pi^2\mu}{h^2} E\right)},$$

A and B being constants. It is clear that A must be zero in the true solution and that B will then no longer be a constant, but a slowly varying function of r. To find the actual form of this function we write

$$\xi = y(z)e^{-z}; \qquad z = \kappa r,$$

in consequence of which equation (41) assumes the form

$$y''+2(z^{-1}-1)y'-y\left(\frac{l(l+1)}{z^2}+\frac{2-\lambda}{z}\right) = 0,$$

where

$$\lambda = \frac{Ze^2}{\kappa}\frac{8\pi^2\mu}{h^2}.$$

To solve this equation we assume a series expansion of the form

$$y = \sum_s a_s z^s,$$

which, when introduced in the differential equation for y, demands the following relation between two consecutive expansion coefficients:

$$a_s = a_{s-1}\frac{2s-\lambda}{(s-l)(s+l+1)}. \tag{41 a}$$

This series must break off for some upper index in order to prevent y from behaving ultimately like the exponential function e^{2z}, which again means an infinite value of the wave function at infinity. But the series can only break off if the numerator happens to become zero for a particular value of s. This means that $\frac{1}{2}\lambda$ must be an integer, usually called the principal quantum number and denoted by n.

The highest power in the series expression of $y(z)$ is thus z^{n-1}. From the form of the recurrence formula for the coefficients it follows further that the series will start with a term z^l, because the coefficient of a_{l-1} in (41 a) is infinite, so that a_{l-1} must itself vanish, in order that a_l shall be a finite quantity.

Introducing the proper expression for κ in the equation for λ and solving for the energy we find

$$E_n = -\frac{2\pi^2\mu Z^2e^4}{h^2n^2}.$$

This will give the correct theory for the Balmer terms of hydrogen, provided the Rydberg constant is given by

$$R = \frac{2\pi^2\mu e^4}{h^3}.$$

This relation was discovered by Bohr in 1913 in his first attempt to develop the quantum theory of spectra.

In the above calculations we have assumed the nucleus to be at rest. This means, as was proved earlier, that the electronic mass occurring in the above formulae is not the true but the reduced mass

$$\mu = \frac{mM}{m+M},$$

where m is the mass of the electron and M that of the nucleus. Since a proton is about 1,800 times heavier than an electron, there will be very little difference between μ and m. But the difference is large enough to be measured spectroscopically. This may be done by comparing the ordinary Balmer lines of hydrogen with the corresponding spark lines of helium. Since the mass of the alpha particle is about four times larger than that of a proton, the helium lines will appear displaced a little to the violet with respect to the hydrogen lines. This displacement is even large enough to be measurable in certain stellar spectra, where it was found in quantitative agreement with theory.†

15. Atomic Dimensions

It is of interest to consider how the atomic dimensions come out of the present theory. Being statistical in its nature, we can only expect to be able to define mean values for the dimensions.

We have already seen how the probability distribution of the electron in the hydrogen atom is determined as regards dependence on direction from the nucleus. In order to find its dependence on distance from the nucleus we must turn to the radial wave function. Let us first of all consider the unit of length in which our radial coordinate z is expressed. In our earlier notation this is equal to $1/\kappa$. Since we now have determined the characteristic values of E, we may express κ as a function of the principal quantum number n. This gives for the unit length

$$\frac{1}{\kappa} = \frac{h^2}{4\pi^2\mu e^2}\frac{n}{Z} \sim 5.10^{-9}\frac{n}{Z}\text{ cm}.$$

For the normal state of hydrogen $n = Z = 1$. This unit length is thus of the order of magnitude of ordinary atomic dimensions, and we must demand that for z-values appreciably larger than unity the radial wave function must fall off very rapidly for atoms in the normal state.

† H. H. Plaskett, *Pub. Dom. Astrophys. Obs. Victoria*, **1** (1922), 325.

Let us now turn to the actual expression of the radial wave function for hydrogen. The relation between consecutive expansion coefficients of y is

$$a_s = \frac{2(s-n)}{(s-l)(s+l+1)} a_{s-1}.$$

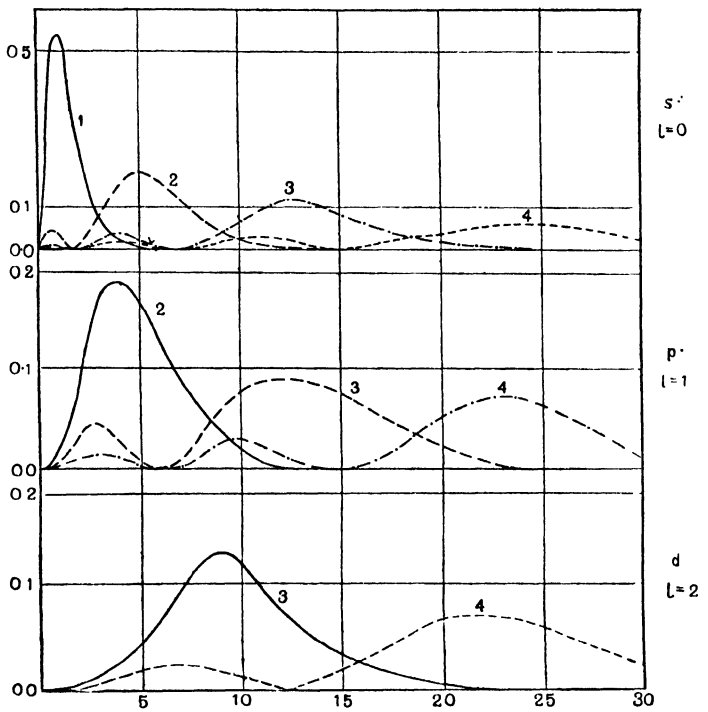


FIG. 1. Graphs of the probability distribution of the electron in a hydrogen atom as a function of the distance from the nucleus. Ordinates are $\xi^2 r^2$ and abscissae the radius vector r in units of $a_H = \frac{h^2}{4\pi\mu e^2} = 5.10^{-9}$ cm. The numbers written on the curves signify the corresponding principal quantum number n. The subsidiary quantum number l is written to the right.

Writing $\alpha = l-n+1$, and $\beta = 2l+2$,

the radial wave function $\xi = ye^{-z}$ assumes the form

$$\xi = \text{const.}\,(2z)^l e^{-z}\left\{1+\frac{\alpha}{\beta}\frac{2z}{1!}+\frac{\alpha(\alpha+1)}{\beta(\beta+1)}\frac{(2z)^2}{2!}+\ldots+\right.$$
$$\left.+\frac{\alpha(\alpha+1)\ldots(\alpha+n-l-2)(2z)^{n-l-1}}{\beta(\beta+1)\ldots(\beta+n-l-2)(n-l-1)!}\right\}.$$

For the normal state ($n = 1$, $l = 0$) the wave function reduces to a mere exponential e^{-z}.

For atoms in excited states (i.e. with $n > 1$) the wave function will always possess maxima outside the centre. In particular, when $n = l+1$ the wave function reduces to the form

$$\xi = \text{const.}\, z^{n-1} e^{-z},$$

which has a single maximum at

$$z = n-1.$$

This maximum is likely to persist as the principal maximum also in cases when l is smaller than $n-1$, but then there will, besides, occur a series of subsidiary maxima for smaller z-values. When $l = 0$ there is always a pronounced maximum at the centre. Some interesting cases are illustrated in Fig. 1.

The marked variation in the distribution function shown by these figures has a simple mechanical meaning. It will be noted, for instance, that when $l = n-1$ there exists only a single maximum of $\xi^2 r^2$. The electron is then running preferentially in a spherical shell, and it is interesting to note that in the earlier form of the quantum theory this case corresponded to a circular orbit. As l approaches zero, the maxima of $\xi^2 r^2$ approach towards the origin. This means, mechanically, that the orbits become more eccentric, so the electron spends more and more time in the vicinity of the nucleus. This is clear from the fact that the eccentricity of a quantized orbit in a hydrogen atom was earlier given by $\sqrt{\{1-(l+1)^2/n^2\}}$.

For the purpose of qualitative considerations it is frequently of advantage to use this mechanical picture of the electron orbits rather than the wave-mechanical language. This will be largely done in the account of the periodic system given in Chapter VI.

V

ATOMS WITH ONE OUTER ELECTRON AND AN ELECTRONIC CORE

16. The Rydberg-Ritz Formula

THE theory of the hydrogen atom may be considered as a first approximation for a much wider class of atoms, those, namely, where a single electron is moving at fairly large distances from the rest of the atom, consisting of the nucleus and an attendant core of inner electrons. Such cases are frequently met with in atomic theory, first of all for ordinary electrically neutral alkali atoms, or singly ionized alkaline earth atoms. The classification of the energy levels by quantum numbers n and l is equally useful in both cases. But while in hydrogen (to our order of approximation) all sub-states with the same n have the same energy, there will in other atoms also appear a dependence on l. The very large increase in distinguishable energy levels produced in this way does not show up to a corresponding degree in the form of spectral lines because of the selection principle for the subsidiary quantum number. Since this number can only change by $+1$ or -1 in a jump, the number of possible spectral lines corresponding to a jump n to n' will not be nn' but only $2n'-1$, where n' is supposed to be smaller than n.

The spectral terms are usually designated by first giving the principal quantum number, and then the subsidiary quantum number symbolized by letters in the order $s, p, d, f, g, \ldots$. This designation is taken over from early spectroscopic literature, meaning sharp, principal, diffuse, fundamental, etc., which names were used for certain spectral line series. Thus $2s$ means $n = 2$, $l = 0$. The extension of this notation to atoms with several outer electrons will be given later on.

From the work of Rydberg and Ritz it follows that the energy levels of an atom of the kind under consideration may be given by an expression of the form

$$E_{nl} = -\frac{RhZ^2}{(n-\delta)^2},$$

where δ, according to Ritz, has the form

$$\delta = \delta_1(l) + \delta_2(l)n^{-2}.$$

It is customary to call $n-\delta$ the effective quantum number, and denote it by n^*.

We shall now show how it is possible to give a theoretical interpretation of this formula on the basis of the wave-mechanical theory, by assuming the atoms in question to consist of an inner electronic core and a single outer electron.

For this purpose we recall the theory of small perturbations given earlier which leads up to the equation

$$\frac{da_n}{dt} - \frac{2\pi i}{h} \sum_m a_m A_{mn} = 0. \tag{42}$$

It may be recalled that the wave function of a certain distribution of atoms over various quantum states is supposed to be given by

$$\psi = \sum a_n(t)\phi_n,$$

where the a_n's are constants or functions of the time, and ϕ_n is the wave function of the nth state in the unperturbed problem. Finally, A_{mn} is given by

$$A_{mn} = \int \phi_n^* \mathrm{K} \phi_m \, d\tau,$$

where K is the perturbation term in the Hamiltonian of the wave equation, while the integration is extended to the whole of the configuration space of the atom.

In the present simple case K stands simply for the change in potential energy due to the deformation of the electronic core, and is independent of the time. This has the obvious consequence that A_{mn} must be proportional to $e^{-2\pi i(E_m - E_n)t/h}$ and will hence oscillate very rapidly about zero as a mean value, except in the case $E_m = E_n$. All terms in the sum occurring in the variational equation (42) will therefore vanish in the mean, except the nth term, when $E_m = E_n$. The equation is then easily integrated, giving

$$a_n = \text{const.}\, e^{-2\pi i A_{nn} t/h}.$$

The perturbation will thus result in an increase of the total energy by an amount

$$A_{nn} = \int \phi_n^* \mathrm{K} \phi_n \, d\tau.$$

The next step is to find the proper expression of K for the case in question. Imagine the outer electron to be so far off from the core that the force on the electron will correspond to the Coulomb force due to the net electric charge on the core. Then each electron of the core screens effectively one unit positive charge on the nucleus.

Assume next the distance to diminish. The core will then by and by become electrically polarized by the electron, the axis of the induced electric doublet passing through the electron. Let M be the strength of the induced doublet and r the distance of the electron from the nucleus. The doublet will then exert an attractive force $2Me/r^3$ on the electron—attractive because the outer electron will push the inner electrons farther off and attract the nucleus, so as to give this latter a better chance to act. But the doublet strength itself is proportional to the external force, that is, $M = \alpha e/r^2$,

where α is a constant. The reactive force of the doublet is therefore $-2\alpha e^2/r^5$ and derives from a potential $-\alpha e^2/2r^4$, which then is the quantity denoted by K above:

$$\mathrm{K} = -\frac{\alpha e^2}{2r^4}.$$

The calculation of the average value of K is much simplified by a reference to the theory of Laguerre polynomials,† which are defined by

$$L_n(x) = (n!)^2 \sum_{s=0}^{n} \frac{(-x)^s}{(s!)^2(n-s)!} = e^x \frac{d^n}{dx^n}(x^n e^{-x}).$$

Denote by L_n^r the rth derivative of the nth Laguerre polynomial. Then the radial wave function of the hydrogen atom may be written in the form

$$\xi(x) = \text{const.}\, x^l L_{n+l}^{2l+1}(x) e^{-\frac{1}{2}x} \qquad (x = 2\kappa r),$$

or $x = 2z$ in our earlier notation.

There is a generating function $G(x,t)$ associated with these polynomials, which is such that

$$G(x,t) = e^{-xt/(1-t)}(1-t)^{-1} = \sum_{p=0}^{\infty} L_p(x)\frac{t^p}{p!}.$$

Differentiating this relation r times with respect to x we find

$$\sum_{p=0}^{\infty} L_p^r(x)\frac{t^p}{p!} = (-1)^r t^r e^{-xt/(1-t)}(1-t)^{-(1+r)}.$$

Hence

$$\sum_{p,q} \frac{t^p u^q}{p!\,q!} \int_0^{\infty} x^{s+2l+2} e^{-x} L_p^r L_q^r \, dx$$

$$= \frac{(ut)^r}{\{(1-t)(1-u)\}^{r+1}} \int_0^{\infty} x^{s+2l+2} \exp\frac{-x(1-ut)}{(1-t)(1-u)}\, dx,$$

† I. Waller, *Zeits. f. Physik*, **38** (1926), 635.

where s is a constant. Remembering that $\int_0^\infty x^n e^{-x}\,dx = n!$ we find easily

$$\int_0^\infty x^{s+2l+2}e^{-x}(L_{n+l}^{2l+1})^2\,dx = \{(n+l)!\}^2(s+2l+2)!P_{nl}^s,$$

where P_{nl}^s is the coefficient of $(ut)^{n-l-1}$ in the power series expansion of

$$(1-t)^{s+1}(1-u)^{s+1}(1-ut)^{-s-2l-3}.$$

The above formula is just of the type we need, since it gives the integral over the square of a wave function multiplied by an arbitrary power of the argument.

In the application in view we need the above formula for the calculation of the integral over K, in which case we must put $s = -4$ in the above formulae. However, it is also necessary to evaluate the integrals for $s = 0$, in order to obtain the factor of normalization. For this latter case we find

$$P_{nl}^0 = \frac{2n(n+l)!}{(2l+2)!(n-l-1)!}$$

and

$$\int_0^\infty e^{-x}(x^l L_{n+l}^{2l+1})^2 x^2\,dx = \frac{2n(n+l)!^3}{(n-l-1)!}.$$

For $s = -4$ we find after some lengthy calculations

$$P_{nl}^{-4} = \frac{2(n+l)!\{3n^2-l(l+1)\}}{(2l+3)!(n-l-1)!}$$

and

$$\int_0^\infty x^{2l-2}e^{-x}(L_{n+l}^{2l+1})^2\,dx = \frac{2(n+l)!^3\{3n^2-l(l+1)\}}{(n-l-1)!(2l-1)l(2l+1)(2l+2)(2l+3)}.$$

Hence it follows easily for the additional energy that

$$\epsilon_{nl} = \int_0^\infty \mathrm{K}\xi^2 x^2\,dx = -\frac{4\alpha e^2\kappa^4\{3n^2-l(l+1)\}}{n(2l-1)l(2l+1)(2l+2)(2l+3)}.$$

Writing the total energy of the state in question in the form

$$E_{nl} = -\frac{RhZ^2}{(n-\delta)^2} \sim -\frac{RhZ^2}{n^2} - \frac{2RhZ^2}{n^3}\delta + \ldots$$

and comparing the last term with the above expression of ϵ_{nl} we find the Rydberg-Ritz correction to be given by

$$\delta = \frac{4\alpha Z^2}{a^3}\,\frac{3-l(l+1)/n^2}{(2l-1)l(2l+1)(2l+2)(2l+3)}; \qquad a = \frac{h^2}{4\pi^2 e^2\mu}.$$

It will be noticed that this expression is just of the form adopted

by Ritz for the empirical correction. Also the absolute value is in fair agreement with observations; but some outstanding exceptions may be noted. First it breaks down in case of sharp terms ($l = 0$). This is due to the fact that our expression for the perturbing potential is then erroneous, and better results are obtained by using an ordinary Coulomb potential inside the core and neglecting the correction outside the core.

In other cases the approximation breaks down although l is not nearly zero. But then it has been shown to be due to a kind of resonance, that is, the lines in the spectrum of the radiation absorbed by the core come close to the lines emitted or absorbed by the outer electron, or, otherwise expressed, certain arc lines tend to coincide with certain spark lines of the same element.†

17. The Spinning Electron

Under sufficiently high resolving power the hydrogen lines are shown to consist of a number of components. This feature was explained by Sommerfeld in the early days of the quantum theory as a consequence of the relativity formula for the variation of mass with velocity, and was for some time regarded as the most convincing proof of this formula. The expression for the energy in the quantized states found by Sommerfeld was the following:

$$1+E/\mu c^2 = \left\{1+\frac{\gamma^2 Z^2}{\{n-l-1+\sqrt{([l+1]^2-\gamma^2 Z^2)}\}^2}\right\}^{-\frac{1}{2}}.$$

Here γ is the so-called fine structure constant

$$\gamma = \frac{2\pi e^2}{hc},$$

while the other symbols have their usual significance. The value of γ is nearly 1/137, and in Eddington's theory of the ratio between the mass of an electron and that of a proton‡ γ is exactly 1/137. But this theory is at present regarded with reserve by most physicists.

The Sommerfeld formula agreed well with the observed fine structure of the hydrogen lines, and also with the fine structure of the corresponding lines of ionized helium. It was further strikingly confirmed by the fine structure of X-rays, which present an analogous problem.

† Cf. E. Schrödinger, *Ann. d. Physik*, **77** (1925), 43.
‡ *Proc. Roy. Soc.* A, **134** (1931), 524.

Later on, however, unexplained features appeared, which threw doubt on the exact validity of the Sommerfeld formula. There were difficulties in the classification of X-ray levels, and in the complex structures of optical levels as well. Another serious difficulty which embarrassed atomic theory at that time was the Zeeman effect, which also seemingly was connected with the theory of the complex structure of the lines, although the reason for this was obscure.

All these difficulties were removed by the hypothesis of Goudsmit and Uhlenbeck† that the electron carries an elementary magnet and angular momentum, in addition to its mass and charge. The angular momentum, 'the spin', of an electron is assumed to be

$$P = \frac{1}{2}\frac{h}{2\pi},$$

and the elementary magnetic moment

$$\mathbf{M} = \frac{e}{\mu c}\mathbf{P}.$$

That the magnetic moment will influence the conditions when the atom is in a magnetic field, and also when several electrons are present, is fairly clear. It is perhaps less clear how the structure of the hydrogen lines can be influenced by the magnetic moment, since the electron is moving in an electrostatic field, which should not influence a magnet at all. The solution of this apparent difficulty lies in the fact that relative to the electron the proton is moving very fast, and hence constitutes an electric current, with an attendant magnetic field. The relation between the magnetic field strength $\mathbf{H}$, the electric field strength $\mathbf{E}$, and the velocity of the electron $\mathbf{V}$ is found by observing that the local rate of change of the electric field intensity with time experienced by the electron moving with velocity $\mathbf{V}$ is

$$\frac{\partial \mathbf{E}}{\partial t} = (\mathbf{V}\nabla)\mathbf{E} = \operatorname{curl}(\mathbf{E}\times\mathbf{V}),$$

on assuming the local electric density zero, so that $\operatorname{div}\mathbf{E} = 0$. Maxwell's first equation of the electromagnetic field then assumes the form

$$\operatorname{curl}\mathbf{H} = \frac{1}{c}\frac{\partial \mathbf{E}}{\partial t} = \operatorname{curl}\frac{1}{c}(\mathbf{E}\times\mathbf{V}).$$

† *Naturwiss.* **13** (1925), 953; *Nature*, **117** (1926), 264.

Hence $$\mathbf{H} = \frac{1}{c}(\mathbf{E}\times\mathbf{V}).$$

Formally a vector whose curl and divergence vanish might be added to the right-hand side of this relation. But since this vector must vanish at infinity, it must vanish everywhere. In the further treatment of the problem it is necessary to take account of the relativity modifications of kinematics. This is evident since the effect of the spin is supposed to serve as correction to the Sommerfeld relativity formula. The point at issue is that the effect of a magnetic field on a magnetized spinning top is to induce it to perform a slow precession of its axis of rotation. The accelerated motion of the electron round the proton will, on the other hand, when the relativity modifications are taken into account, be equivalent to another precession of the frame of reference of the spinning electron, which is of the same order of magnitude. This effect was discovered by Thomas,† and is known as the *Thomas Precession*. It works in the opposite direction to the induced magnetic field, and reduces it by a factor $\frac{1}{2}$. If we only had to deal with the fine structure of the hydrogen lines, this factor would not be of so much consequence, as it cannot then be distinguished from the arbitrarily adopted values of the elementary magnetic moment and spin of the electron. But the Thomas precession becomes of essential importance when it comes to the explanation of the anomalous Zeeman effect, as it is only in this way possible to explain the two effects simultaneously with the same values of spin and magnetic moment.

An elementary magnet of moment M subjected to the influence of a magnetic field of strength H has a potential energy

$$U = \mathbf{MH}.$$

The proof of this formula is very simple and need not be repeated here. The next step is to introduce the proper expressions of **M** and **H** in this formula and to calculate the average value of the interaction energy. This was carried out by Darwin‡ on one hand and Heisenberg and Jordan‖ on the other. The result is, for a hydrogenic atom, that the average additional energy in a stationary state is

$$\bar{U} = -E_n \frac{Z^2\gamma^2}{n} \frac{j(j+1)-l(l+1)-s(s+1)}{l(l+1)(2l+1)}.$$

† L. H. Thomas, *Phil. Mag.*, ser. 3, **7** (1927), 1.
‡ *Proc. Roy. Soc.* A, **115** (1927), 1. ‖ *Zeits. f. Physik*, **37** (1926), 263.

We have written down the formula in the above rather complicated form because this is the way in which it was first stated. Here s represents the angular momentum of spin in units of $h/2\pi$, and is thus to be taken equal to $\frac{1}{2}$. Further, j is a new quantum number which appears in consequence of the degeneracy of the Balmer terms. It signifies the total angular momentum of the atom, orbital plus spin, and is thus either equal to $l+\frac{1}{2}$ or $l-\frac{1}{2}$. However, j is always positive, so that the level $l = 0$ is single, having $j = \frac{1}{2}$ only, $j = -\frac{1}{2}$ being rejected. Introducing the proper expressions of j and s as given above, we find that all distinguishable energy levels will be given by the single formula

$$\bar{U} = -E_n \frac{Z^2\gamma^2}{n} \frac{l \text{ or } -(l+1)}{l(2l+1)(l+1)}, \tag{43}$$

the alternatives in the numerator corresponding to $j = l+\frac{1}{2}$ and $j = l-\frac{1}{2}$ respectively.

To complete the solution of the problem it is necessary to calculate the wave-mechanical equivalent of the Sommerfeld relativity formula. This is a straightforward problem, which leads to a resultant expression of the energy

$$1+E/\mu c^2 = \left\{1+\frac{\gamma^2 Z^2}{\{n-(l+\frac{1}{2})+\sqrt{[(l+\frac{1}{2})^2-\gamma^2Z^2]}\}^2}\right\}^{-\frac{1}{2}},$$

which differs from the Sommerfeld formula by the substitution of $l-\frac{1}{2}$ for l. Expanding this expression to the second order in γ^2 we find

$$E = E_n + E_n \frac{\gamma^2 Z^2}{n}\left(\frac{1}{l+\frac{1}{2}} - \frac{3}{4n}\right).$$

Adding (43) to this expression we find that all fine structure levels corresponding to a given n are contained in the formula

$$E = E_n + E_n \frac{\gamma^2 Z^2}{n}\left(\frac{1}{l+1} - \frac{3}{4n}\right) \qquad (l = 0, 1, 2, \ldots),$$

without introducing ambiguous signs. The reason for this is that the levels will coincide in pairs. The nth level will thus be split up into n component levels, which are given by the above formula.

The introduction of an electronic spin also influences the magnetic quantum number. This is now more naturally defined with respect to the total quantum number j, representing the component of j

parallel to a given direction. It may therefore run through the series of values

$$-j, -j+1, -j+2, \ldots, j-1, j.$$

Thus m is no longer an integer. For further details compare the paper of Heisenberg and Jordan referred to above (*Zeits. f. Physik*, **37** (1926), 263). It is sufficient here to give the final formula for the change in the energy levels of an atom caused by a homogeneous magnetic field of field strength H:

$$\bar{U} = \frac{eH}{2\mu c} m \left\{ 1 + \frac{j(j+1) - l(l+1) + s(s+1)}{2j(j+1)} \right\},$$

where the magnetic quantum number m is defined in the above manner. In this formula H is supposed infinitely small, so that the true magnetic energy is essentially smaller than the separation of the fine structure levels.

18. The Alkali Spectra

Consider next the case of an atom in which a single electron moves at relatively large distances from an inner core of electrons, which remains practically unaltered during transitions of the outer electron. This model was previously adopted as a basis for the interpretation of the arrangement of the spectral terms into s, p, d series, etc., the Rydberg-Ritz formula, and the selection principle for the subsidiary quantum number. While these characteristics apply more or less accurately to a large number of elements with rather different chemical properties, which suggest the presence of several outer electrons, we shall now see that when the electronic spin is taken into account we must limit the application of this model to atoms of alkali metals only.

In fact, it will be expected that formula (43) for the magnetic energy of the spinning electron may be taken over also for the more general case, provided the change in the orbit of the electron is compensated by the use of an effective quantum number n^* instead of n. A change of l by one unit will now affect n^* also, and this will have the effect of suppressing the coincidence of levels in pairs. The only fine structure left over is then the one given by the ambiguity of (43) which corresponds to a level separation of

$$\Delta E = \frac{E_n}{n^*} \frac{Z^2\gamma^2}{l(l+1)}. \tag{44}$$

In case of s-terms the ambiguity vanishes, owing to the fact that j is

not permitted to be negative. *Hence the s-terms are always singlets.* That a formula of the form (44) will fit the observational data concerning doublet spectra was shown by Landé, and by Millikan and Bowen.

It should be noted that the relativity term has disappeared from the above expression of the doublet separation. The alkali doublets are thus an exclusive effect of the electronic spin, in contradistinction to the fine structure of hydrogen, which is a combined effect of relativity *and* spin.

In case of elements other than alkali metals, or in case of spectra of higher order, more complicated multiplicities will be found, which cannot be explained on the above lines. These multiplicities are due to the interaction of several valency electrons, both as regards their spins as well as their orbital magnetic moments.

In the above short account of the spinning electron mechanical notions are used considerably beyond the limits of their applicability. It is therefore gratifying that Dirac† was able to give the theory such a turn that the electronic spin appeared as a sole effect of the relativity modifications in mechanics, without introducing the notion of spin explicitly, a state of affairs which is more in keeping with our general feeling about the matter. But for practical purposes it is an established use to talk of the electron 'as if' it carried a spin and a magnetic moment, which is especially convenient in qualitative discussions of spectral structure. For this reason we shall not enter further into details of Dirac's relativistic theory.

† *Proc. Roy. Soc.* A, **117** (1928), 610; **118** (1928), 351.

VI

THE PERIODIC SYSTEM

19. Pauli's Exclusion Principle

BEFORE considering the theory of spectra of higher multiplicities it is necessary to study in more detail how the interaction between several electrons in an atom is to be treated on the basis of the present theory. One soon encounters the difficulty that a large class of solutions, which are perfectly good from a theoretical point of view, are not realized in atomic configurations. How this comes about is best seen from an example like the following.

Consider a hydrogenic atom in its normal state, characterized by the values 1, $\pm\frac{1}{2}$, 0, $\frac{1}{2}$ of the quadruple numbers n, m, l, and j. Assume the nuclear charge to consist of several elementary units, and consider the process of capturing additional electrons in the atom, external agencies being absent. Neglecting the electronic interactions we expect the electron last bound to suffer a series of transitions, and finally also to be bound in a normal state. What difference will the interaction between the electrons make to this result? For the spectrum of the radiation emitted during the binding process the difference may be considerable. The original system happens, namely, to possess an exceptional type of degeneracy, which makes the perturbations in the energy values abnormally large. Otherwise the evolution of the system will proceed very much as if the interaction between the electrons were absent. In particular, the final state of the system is likely to be characterized by the same set of quantum numbers as a normal hydrogenic state containing two electrons. The functional dependence of the energy upon the quantum numbers may, however, be somewhat different from the pure hydrogenic case, but this is an unimportant point here. Adding a third electron to the atom would probably only result in the introduction of a third equivalent electron. In short, we expect all electrons to be bound in equivalent states.

This conclusion, however, is contradicted by a mass of facts in atomic physics, notably the structure of X-ray spectra and optical multiplets, as well as the pronounced periodicity in the chemical properties of the elements. It is therefore necessary to restrict the possible solutions by an exclusion principle, which can be formulated

as follows: an atomic system will never contain equivalent electrons. We then assume equivalence to be defined by equal values of the quantum numbers n, l, m, and j.

Dirac and Heisenberg have brought Pauli's principle into a form which is convenient for discussions on the basis of wave mechanics. Consider an atomic system containing r electrons, and let ψ_1, ψ_2,..., be the characteristic solutions of the wave equation corresponding to this atom. Due to the linearity of the wave equation the solutions in question can be separated into two series $\psi_k^{(s)}$ and $\psi_k^{(a)}$ such that the functions of the first series are unaltered by interchanging the position of two electrons, while the functions of the second series change the sign only by this process. That is, the functions of the first series are *symmetric* and those of the second series *antisymmetric* in the coordinates and the spins of the electrons.

The possibility of this requirement being admitted, it follows that all harmonic amplitudes in the electrical moment of the atom, which involve one symmetric and one antisymmetric wave function are *zero*.

The truth of this statement follows directly from the fact that interchanging two electrons, the integral over the wave functions which gives the harmonic amplitude associated with the transitions in question will change its sign if it involves one symmetric and one antisymmetric wave function. *In as much as the electrons are supposed to possess identical structures* the amplitudes in question cannot possibly alter their value by such an interchange, and the integrals must vanish. The possible stationary states of a system of identical electrons therefore fall in two series between which there can occur no transitions. Each of these series therefore constitutes an independent class of solutions of the quantum problems.

Pauli's principle applies to the antisymmetric functions, while the symmetric functions must be rejected as solutions of atomic problems. However, in the quantum theory of the radiation field the symmetric functions are of importance, while the antisymmetric must be rejected. It should also be realized that symmetric functions are frequently used in atomic problems, when a solution is sought by neglect of the electronic spin. It is then understood, however, that using these solutions there will be no more electrons in the atom than may be labelled in an individual manner by a suitable distribution of the spins.

Pauli's exclusion principle is best illustrated by a discussion of the chemical properties of the elements. For a detailed theory the actual solutions of the wave equation for molecular systems will, of course, play the principal part, but for a qualitative survey Pauli's principle is like a flash-light that suddenly illuminates the night.

It is clear that the pronounced periodicity of the chemical properties of the elements must be primarily determined by the atomic surface electrons, since there is no particular reason why electrons buried deep in the atom should have any influence on the outside, apart from their screening effect on the nuclear charge. And we shall see that the surface electrons must exhibit a periodicity which really goes parallel to the chemical properties in a striking manner, provided Pauli's principle is duly considered.

Consider as an illustration the imaginary case in which the interaction between the electrons in the atom is neglected. In order to fix the otherwise indefinite magnetic quantum number m we assume the atom to be situated in a weak magnetic field. The state of each electron is then characterized by the four numbers n, l, m, and j.

By Pauli's principle no two electrons can occupy the same state. There can thus only be two electrons in the state $n = 1$, since then $l = 0$, $j = \frac{1}{2}$, and $m = \pm\frac{1}{2}$. For $n = 2$ there are eight permissible states:

$$l = 0, \quad j = \tfrac{1}{2}, \quad m = \pm\tfrac{1}{2}$$
$$l = 1, \quad j = \tfrac{1}{2}, \quad m = \pm\tfrac{1}{2}$$
$$l = 1, \quad j = \tfrac{3}{2}, \quad m = \pm\tfrac{3}{2} \text{ and } \pm\tfrac{1}{2}.$$

In general, for a given value of l there are $4l+2$ possible states. We shall adopt the nomenclature that all electrons with the same value of l form a sub-group, and that all electrons with the same value of n form a principal group. In each sub-group there are then $4l+2$ electrons, while a principal group contains $2n^2$ electrons, since

$$\sum_{l=0}^{n-1}(4l+2) = 2n^2.$$

In X-ray spectroscopy a definite nomenclature has been adopted for the principal electronic groups, each group being denoted by a capital letter, starting with K and continuing upwards through the alphabet. Thus the K-group ($n = 1$) contains $2.1^2 = 2$ electrons, the L-group ($n = 2$) contains $2.2^2 = 8$ electrons, the M-group ($n = 3$) contains $2.3^2 = 18$ electrons, and so on. For the sub-groups we shall use the nomenclature of optical spectroscopy, and denote them by

the letters s, p, d, f, etc. Thus the groups with $l = 0, 1, 2, 3, \ldots$ will be denoted by the letters s, p, d, f and will be called the s, p, d, f-groups. In order to indicate the principal group to which a given sub-group belongs we put the value of n in front of it, as is customary in spectroscopy. The d-sub-group of the M-group will thus be denoted by $3d$.

We have already seen that the first two principal groups will contain 2 and 8 electrons respectively, such that these groups will be expected to be just saturated in helium ($Z = 2$) and neon ($Z = 10$). These elements are remarkable for the fact that they form inert gases, which do not enter into chemical combinations of any kind and can only be made to liquefy at extremely low temperatures. This fact immediately suggests that the ability of an atom to enter into chemical combinations depends essentially upon the degree to which its outer electronic group is saturated. Further information is obtained by considering the elements adjacent to an inert gas in the periodic system. In fact, since an inert gas is supposed to contain a saturated electronic group, the succeeding element will presumably have a single electron loosely bound outside a saturated group. This surmise is amply confirmed by the spectra of the alkali metals. The unit valency of the alkali is therefore linked up with the fact that the surface group of the atoms contains a single electron only.

The further fact that these elements carry a positive charge in electrolytic solutions, i.e. are electro-positive, means evidently that the additional electron is only loosely bound and easily lost. This is, of course, a fact which is well known from spectroscopic measurements of the strength of the binding of the electron.

Conversely, the electronic configuration of the elements *preceding* the inert gases, the halogens, must have a surface group which is saturated *save* for the absence of a single electron. The fact that these elements are monovalent and electro-negative means therefore that an unsaturated group has a strong tendency to take up an extra electron, so as to complete the group in question. This happens in spite of the fact that the atom acquires a surplus negative charge in the process.

It is now easy to generalize and to ascribe the divalent and electro-negative properties of the elements preceding the halogens to the fact that the electronic surface groups of these elements will be saturated by the addition of two electrons. Similarly, the fact that

the alkaline earths are divalent and electro-positive means that the surface group contains two electrons only.

This picture gives a suggestive explanation of the chemical affinity between electro-negative and electro-positive elements as a simple recombination between particles of opposite electric polarity. The recombination of a positive sodium atom with a negative chlorine atom to form an electrically neutral NaCl molecule being analogous to the recombination of a free electron with a positive nucleus to form an atom. In this way we obtain the first partial understanding of all heteropolar chemical combinations, i.e. the cases in which the combining entities have opposite electric polarities. Considering electronic groups half-way completed, we shall be somewhat embarrassed, however, to say which way the group is likely to complete itself, either by loosing some electrons and becoming electro-positive or by adding surplus electrons and become electro-negative, or on the whole how many valencies the atom will possess. The rule here seems to be, as has been elaborated in detail by Slater, that the valency is determined by the resultant spin of the electrons in the outer group. That a closed group has no resultant spin is a simple consequence of Pauli's principle.

20. Influence of Electronic Screening

It is now time to improve upon the simple model of an atom, which was adopted as a basis of the discussion. The neglect of the interaction between the electrons in the atom served the purpose of indicating the main origin of the periodicity in the chemical properties, and the parallelism between atomic structure and chemical valency. However, this parallelism raises the pertinent question why atoms with incomplete electronic groups have a tendency to complete themselves by adding or losing electrons. A quick glance over the periodic system shows, moreover, that the chemical properties of successive elements are frequently quite different from the succession predicted by energetic considerations when the interaction between the electrons is neglected.

It so happens that it is fairly simple to find a satisfactory answer to both questions. It is for this purpose largely sufficient to deal with the problem from the approximate point of view which was used in discussing the Rydberg-Ritz formula.

Assume that an atom is constructed in the following manner.

Initially there is a bare nucleus with Z positive charges and Z free electrons. These latter are then one by one carried into the field of the nucleus, so that the atom is being built up in successive steps. The first electron bound by the nucleus will, when left to itself, evidently settle down into a normal hydrogenic state. The binding of the second electron will now be complicated by the fact that the inner electron will exert a screening effect, such that at great distances from the nucleus the conditions will approximate the case of a hydrogenic atom with atomic number $Z-1$. When the second electron has settled down into its normal state we can still assume that the field of force it is running in is, on an average, centrally symmetrical and practically hydrogenic, but the screening by the first electron is not unity any more, but has rather some value between $\frac{1}{2}$ and $\frac{1}{4}$.

Consider next the binding of the third electron, the first electron of the L-group. It runs outside the K-group most of the time, so that the K-electrons will do their full amount of screening. This means, among other things, that the *dimensions* of the L-group will be larger than would have been the case if the K-electrons were absent. Adding more electrons to the atom, the *mutual* screening of electrons in the L-group must also be taken into account.

It is important then to realize that the mutual screening between electrons in the same group may be essentially smaller than the screening by inner groups. The reason for this is essentially embodied in the electrostatic theorem that a spherically symmetric distribution of electricity outside a sphere will not give rise to an electric force inside the sphere. Considering, for instance, the motion of an L-electron we realize that the screening by K-electrons is practically complete, while the L-electrons themselves will occupy an intermediate position.

This incomplete screening between electrons in the same group, particularly when the group approaches completion, is the main cause of the electro-negative properties of the halogens and allied elements. In fact, an incomplete screening must, for electrons in the given group, be equivalent to a surplus positive charge of the nucleus. It is therefore not surprising that in an electronic group nearing completion this virtual charge has become equivalent to one or more real positive charges, such as to warrant the binding of surplus electrons in the group.

21. Relative Importance of Sub-groups

In the case of M-electrons the screening by K- and L-electrons amounts to nearly ten units. But in this case there may be considerable difference between electrons in different sub-groups. The s-electrons, for instance, will frequently penetrate even within the region of the K-electrons, and thus experience the full force of the nucleus. The d-electrons, on the other hand, will run in spherical shells, practically, outside both the K- and L-electrons, and hence be screened by inner groups to the maximum amount, and this differ-

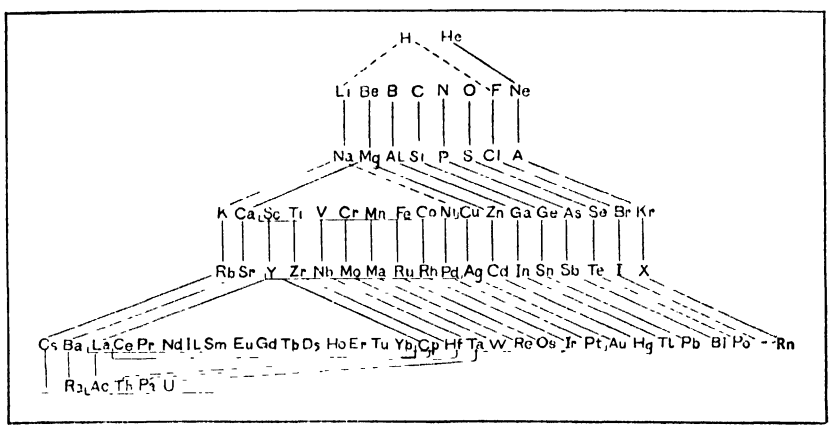


FIG. 2. The periodic system of the elements. The principal periods are represented by the arrangement of the elements in horizontal rows. Elements belonging to the iron group, the zirconium-palladium group, and the rare earths are enclosed by frames. Chemically related elements are connected by full lines.

ence grows more and more important on passing to higher groups. For the s-electrons will penetrate into the K-shell no matter to which shell they belong, and thus break through all intermediate groups.

It is this difference in the screening for electrons in different sub-groups which is the essential cause of the irregular succession in the chemical properties of heavier atoms.

In order to facilitate the survey of the elements we give in Fig. 2 a diagram of the periodic system, which brings out the periodicity in a rather striking way. In this diagram the principal periods, counted from one inert gas to the next, are written in separate rows. Elements in different periods, but with similar chemical properties, are connected by vertical or oblique lines. Adjacent elements with similar chemical properties are underlined.

22. The Problem of Argon

The elements from hydrogen to argon appear to be built in the succession predicted from the simplified scheme with the electronic screening left out of consideration. Thus helium closes the K-group, and neon the L-group. The next inert gas should be expected at atomic number

$$2.1^2+2.2^2+2.3^2 = 28.$$

Element number 28, however, is occupied by the metal nickel, the properties of which differ very much from those of an inert gas. On the other hand, already at atomic number 18 there occurs an inert gas, *argon*. This element marks the completion of the $3p$-group of the M-electrons, but not the M-group itself. The fact that a sub-group here takes over the function played by principal groups in the earlier elements marks the stage when the difference in screening effect of sub-groups has increased to prime importance.

That this is the true explanation is nicely shown by the succeeding element potassium. Since argon closes the $3p$-group, we should expect potassium to possess a closed $3p$-group and a single $3d$-electron running nearly circularly outside. The question here naturally arises, however, whether an electron having the low principal quantum number 3, and being exposed to the full screening of 18 electrons, will have less or more energy than an electron having higher principal quantum number 4, and the advantage of being an s-electron which may penetrate through the intervening shells right into the vicinity of the nucleus. The arc spectrum of potassium shows clearly that the latter case is the one realized in nature, the additional electron being bound in a $4s$-state.

23. The Problem of the Iron Group

The element succeeding potassium is calcium. Here the same question arises as regards the binding of the additional electron, and the answer is the same: it is bound in a $4s$-state. This completes the first sub-group in the N-shell, and the question, what will happen becomes still more pertinent in the succeeding element, scandium. For while a $3d$-orbit may lose in the competition with a $4s$-state, it may win when competing with a $4p$. On the other hand, with increasing nuclear charge the principal quantum number will gain relatively in importance, and may induce the electrons to turn from the $4s$-state to the slighted $3d$-state again. This is exactly what

happens in scandium, where not only the electron last bound but all three outer electrons have settled down in $3d$-states.

This uncertainty about the right choice of sub-group is the characteristic feature of the elements of the iron group. There is a constantly recurring tendency to build up an outer N-shell of electrons round an incomplete M-shell. Although the idea has never been worked out in detail there is no doubt that the characteristic properties of these elements must be closely connected with the simultaneous presence of *two* incomplete electronic shells in the atoms.

The element succeeding nickel is copper ($Z = 29$). According to spectroscopic data this element has a single $4s$-electron running outside a closed electronic group. This must mean that the M-group was effectively closed in nickel, and that the $4s$-electron in copper is the first one which appears in its legitimate order of succession. The next element, zinc, must then have two $4s$-electrons. These conclusions are confirmed by the general similarity of the arc spectra of Cu and Zn on one hand, and the arc spectra of alkali and alkaline earth metals on the other. The next six elements would seem to behave regularly, as would be expected, such that krypton ($Z = 36$) possesses two complete sub-groups of N-electrons, in addition to the completed principal groups K, L, and M. It is true that krypton only marks the completion of the first sub-groups of N. But we learned from the case of argon that this may be sufficient for the development of inert gas characteristics, and this holds for krypton, which is a typical inert gas.

24. The Fifth, Sixth, and Seventh Periods

From this stage on history repeats itself, the development of the succeeding elements exhibiting a striking similarity to the growth of the iron group. The valency electron in rubidium, succeeding krypton, is bound in a $5s$-state, instead of a $4d$-state. This means, in our scheme of interpretation, that now the internal screening has reached the stage when a difference of two units in l outweighs one unit in n. Strontium ($Z = 38$) has, similarly, two electrons in the $5s$-group. But in the next element, yttrium, there occurs the same back-slide which we noticed in scandium; and in the eight elements succeeding strontium, the yttrium-palladium group, an oscillation back and forth between $4d$- and $5s$-states seems to play the same part as an oscillation between $3d$ and $4s$ did in the iron family. The uncertainty

is settled in silver ($Z = 47$), in which case the first three sub-groups of the N-group appear to be saturated with their normal number of 18 electrons, with a single electron in a $5s$-state outside.

The next element, cadmium ($Z = 48$), has two $5s$-electrons, while the next few elements complete the $5p$-states, so that the element xenon, which is eight units above palladium, will have the sub-groups $5s$ and $5p$ completed, and, according to expectations, the properties of an inert gas. It is not easy to explain why the completion of the first two sub-groups of a principal group marks the stage of an inert gas, while the completion of succeeding sub-groups passes unnoticed. But it could probably be explained by a careful study of the screening conditions.

The Sixth Period. Superficially we might now expect that conditions in the next period will be a mere repetition of those in the fourth and fifth periods. But this would be a mistake. The beginning is hopeful, however, as the element succeeding xenon, caesium, has the properties of an alkali metal, and a single electron running in a $6s$-state as was to be expected. Barium ($Z = 56$) falls into its proper place as an alkaline earth metal, and has two $6s$-electrons.

Then, however, something partially new happens. For we must remember that we have been going on adding electrons in new groups at the surface of the atom, without taking care of completing the inner groups in due course. Thus barium still lacks the f-electrons in the N-group, and all d-, f-, and g-electrons of the O-group, while it is busy building up a still higher group, the P-group. This looks like a preparation for a landslide of electrons, as, considering the small screening effect between electrons of the same sub-group, we should expect a group to be filled to the brim as soon as it starts capturing electrons.

This is what happens in the elements succeeding barium. Lanthanum ($Z = 57$) is on the verge, having probably its additional electron bound in a $5p$-state. In the fourteen elements succeeding lanthanum we now witness the completion of the N-group. These elements form *the rare earth family*, and have closely similar chemical properties. This latter circumstance is essentially due to the fact that the spatial dimensions of the N-group will be considerably smaller than the dimensions of the O-group. What happens to the N-group will therefore be only feebly reflected on the atomic surface, which determines the chemical properties of the element.

The rare earths close up with ytterbium, and in the next eight elements, ending with platinum, the first three sub-groups of *O* are completed. This group of elements parallelize, in a certain way, the elements of the iron family, and the yttrium-palladium family in earlier periods. Further on the *O*-group probably remains unaltered, while the first two sub-groups of *P* are built up, terminating in the inert gas *niton*.

The Seventh Period contains only five (or six) known elements, all of which are strongly radioactive. As far as we can judge, this period runs parallel to the previous one in the beginning, as would be expected from analogy. But as we only know a few of its elements it is rather idle to speculate about the conjectural properties of elements which are absent. It is also probable that for such high atomic numbers the problem of constructing elements still heavier than uranium will be complicated by the finite extension of the atomic nucleus, even if it should be possible to find a nucleus which was not subjected to spontaneous disintegration. The dimensions of the *K*-shell in uranium are, in fact, not much larger than the estimated dimensions of the uranium nucleus.

VII

THE THEORY OF MULTIPLETS

25. Nomenclature

In the more detailed theory of the atoms with several electrons it is advantageous to use a generalized nomenclature in which the notation used for atoms with only one electron appears as a special case.

The principal quantum number is not particularly suited for a generalization of this sort, and it is therefore fortunate that the part played by this number becomes more and more ill defined as the multiplicity properties of the atoms become more pronounced. When referring to this number in the theory of complex spectra we shall simply take it to define in the ordinary way a sort of mean value of the energy of all levels belonging to the same multiplet.

As regards the remaining three quantum numbers we introduce capital letters L, J, and S as a generalization of the numbers l, j, and s of single-electron atoms. Here L will appear as a sort of vector sum of all the l's of the individual electrons. Denoting the l of the ith electron by l_i we have therefore

$$L \leqslant \sum_i l_i,$$

the summation being extended to all electrons in the atom. In the same way we have

$$J \leqslant \sum_i j_i, \quad \text{and} \quad S \leqslant \sum |s_i|.$$

The number J signifies the total angular momentum of the atom, measured in units of $h/2\pi$, and it therefore satisfies the inequality

$$J \leqslant |L+S|.$$

In order to obtain a clear view of the rather complicated cases met with in multiplet spectra it is necessary to introduce a generalized sort of nomenclature for the atomic states involved, by which a given state is completely characterized by a single symbol. Several systems of nomenclature have been devised from time to time, but at present it seems that the form originally proposed by Russell and Saunders enjoys universal favour. In this system the terms are denoted by capital letters in the usual order:

$L = 0, 1, 2, 3, 4,\ldots$ by the term symbols: $S, P, D, F, G,\ldots$.

The multiplicity of the term is written as a separate upper index on the left side of the letter. Thus a triplet P term is denoted by ^3P. The value of J, on the other hand, is given as a lower right index. If the above triplet P term corresponds to an inner quantum number $J = 2$, the notation is $^3P_2$. It remains to indicate the principal quantum number. This is frequently not even known or is uninteresting, and may be left out altogether. But if it is important to bring it out it is added in front of the letter. If the above term corresponds to $n = 4$, the complete symbol assumes the form $4\,^3P_2$.

It is frequently of interest to indicate explicitly how a particular configuration of electrons is constructed. We then use the ordinary term symbols, and add the number of electrons in a given state by an upper right-hand index. The normal state of the carbon atom is a triplet and contains two $1s$-electrons, two $2s$-electrons, and two $2p$-electrons, and these electrons interact in such a manner that $L = 1$, $S = 1$, and $J = 0$. The normal state is therefore characterized by the symbol

$$(1s)^2(2s)^2(2p)^2\,^3P_0.$$

But it is often unnecessary to indicate more than the last bound electron or electrons and the above symbol is then abbreviated to $2p\,^3P_0$.

The prototype of complex spectra is set by the so-called *normal* multiplets, which appear as a direct generalization of the alkali doublets, doublets which therefore are included in the group. A normal multiplet spectrum consists of several s, p, d series, etc., the members of each series having a characteristic kind of multiplicity. Thus calcium has one singlet term system and one triplet system. Scandium has doublets and quartets, titanium singlets, triplets, and quintets, etc. The multiplicity of a series is characterized by its S-value. Singlets have $S = 0$, doublets $S = \frac{1}{2}$, triplets $S = 1$, etc. The multiplicity is in any case given by $2S+1$.

For a given value of S and L the number J can only assume the values

$$J = L+S,\ L+S-1,\ldots,\ |L-S+1|,\ |L-S|,$$

in accordance with the presumed vector property of J.

This general nomenclature may be useful in the analysis of any kind of multiplet spectra. The so-called normal multiplets, which correspond to a coupling between the outer electrons of the kind first

distinctly defined by Russell and Saunders,† are singled out among other multiplets by particular selection rules, which result from the assumed coupling scheme. In the Russell-Saunders coupling scheme there is a strong coupling between the individual l_i and the individual s_i, so that they combine into quantized resultants L and S in the manner considered above. But the coupling between L and S is supposed to be slight. From these assumptions the following selection rules may be derived:

$$L \to L \pm 1 \quad \text{or} \quad L \to L,$$
$$J \to J \pm 1 \quad \text{or} \quad J \to J,$$

and

$$S \to S.$$

The last rule states that the multiplicity is conserved during a transition, so that intercombination lines require a departure from the Russell-Saunders coupling scheme in order to occur. From the fact that intercombination lines do occur, occasionally, in almost all multiplet spectra, it follows that the normal multiplet scheme forms a sort of ideal limit, which is not realized rigidly in actual cases.

That L and J must conform to the same selection rules in normal multiplets is a simple consequence of the assumed negligible coupling between L and S. A change in J must then necessarily lead to a corresponding change in L.

Applying these rules we find the following scheme for normal multiplets:

L	J		J					
0	0	singlets $S = 0$	1/2					doublets $S = 1/2$
1	1		1/2	3/2				
2	2			3/2	5/2			
3	3				5/2	7/2		
0	1	triplets $S = 1$		3/2				quartets $S = 3/2$
1	0 1 2		1/2	3/2	5/2			
2	1 2 3		1/2	3/2	5/2	7/2		
3	2 3 4			3/2	5/2	7/2	9/2	
0	2	quintets $S = 2$			5/2			sextets $S = 5/2$
1	1 2 3			3/2	5/2	7/2		
2	0 1 2 3 4		1/2	3/2	5/2	7/2	9/2	
3	1 2 3 4 5		1/2	3/2	5/2	7/2	9/2	11/2
0	3	septets $S = 3$				7/2		octets $S = 7/2$
1	2 3 4				5/2	7/2	9/2	
2	1 2 3 4 5			3/2	5/2	7/2	9/2	11/2
3	0 1 2 3 4 5 6		1/2	3/2	5/2	7/2	9/2	11/2 13/2

† *Astrophys. J.* **61** (1925), 38. A very comprehensive discussion of various possible coupling schemes in multiplets is given by Laporte, *Handb. d. Astrophys.* III. **2**.

One will notice that the sharp series terms ($L = 0$) are always singlets, no matter whether they belong to a singlet *system* or not. Passing from sharp to principal, diffuse terms, etc., the multiplicity increases until it reaches a maximum value, and from thence it remains constant.

The manner in which S is related to the elementary electronic magnets is indicated among other things by the manner in which S takes an upper limit for different elements. The conditions are such that each outer electron contributes at most $\frac{1}{2}$ to the maximum value of S. When the number of electrons in an outer shell approaches its maximum number it seems, however, as if they gradually lose their power of manifesting their presence by affecting the maximum number of S, so that the multiplicity does not reach its theoretically possible upper limit. The following table giving a survey of the observed spectra of the elements from potassium to zinc will make this matter clear, and otherwise serve to illustrate the features of multiplet spectra.

Arc spectra

	K	Ca	Sc	Ti	V	Cr	Mn	Fe	Co	Ni	Cu	Zn
Singlets		×		×						×		×
Doublets	×		×		×			×			×	
Triplets		×		×		×		×		×		×
Quartets			×		×		×		×		×	
Quintets				×		×		×		×		
Sextets					×		×		×			
Septets						×		×				
Octets							×					

Spark spectra

	K+	Ca+	Sc+	Ti+	V+	Cr+	Mn+	Fe+	Co+	Ni+	Cu+	Zn+
Singlets			×									
Doublets		×		×								×
Triplets			×		×							
Quartets				×		×						
Quintets					×							
Sextets						×		×				
Septets							×					
Octets												

It will be seen that the rule of maximum multiplicity holds fairly well up to seven electrons, but then the maximum multiplicity declines. A rule which holds without exception, however, is the rule of alternation, that is, the fact that an even number of outer electrons

gives odd multiplicities while an odd number gives even multiplicities.

26. Displaced Terms

In most cases of observed multiplets all but one of the valency electrons occupy their normal positions. Theoretically any distribution of electrons over the possible quantum states is permissible, but experimental conditions are mostly such that the excitation of one electron is strongly favoured as compared with the simultaneous excitation of several electrons. There are some exceptions to this rule, however, the most conspicuous being found among the terms of the arc spectra of alkaline earths. These exceptions are due to the existence of metastable states of the atom, particularly in the ionized form.

Looking over the spectral terms of singly ionized calcium, for instance, we find that the lowest d-term is intermediate in value between the lowest p- and the lowest s-term, the latter corresponding to the normal state of the atom. Due to the selection principle for l a spontaneous transition from a d- to an s-state is excluded, and once the atom is settled down into the lowest d-state some external influence is required in order to transfer the electron to the normal or some other state. Or, if the metastability is due to the vanishing of the associated dipole moment only, the life period of the state will be very much prolonged, so as to allow the next higher non-vanishing multipole to come into play. Under conditions of sufficiently strong excitation there will in any case result an abnormal aggregation of singly ionized atoms in the lowest d-state, and the case may occur with a rather large frequency that such an atom recombines with a free electron. In the recombination process such atoms will emit radiation, the spectral lines of which differ from the usual ones because the influence of the additional electron proves sufficiently large to destroy the metastability in question and to induce a simultaneous transition of *both* valency electrons to states of lower energy.

The difference between the spectrum emitted in this way and the ordinary spectrum consists essentially in a displacement of all lines towards higher frequencies, and hence the designation 'displaced terms'. The displacement is found to correspond to the energy difference between the lowest d- and the lowest s-state of the ionized

atom, as was to be expected. No exact correspondence is found, since the influence of the outer electron is likely to produce a considerable deformation of the motion of the inner electron. This influence must decrease with increasing distance of the outer electron in its initial state, and must vanish at the series limit. This inference is in complete accordance with experiments. Displaced terms originating in a similar way are found in the spectra of N, B^{+}, C^{++}, Al^{+}, S^{++}, P^{+++}, and S^{++++}.

27. Probabilities of Transitions between Multiplet Levels

The theory of intensities in multiplets has already played some part in astrophysical discussions, and may become of more prominence in the future. We shall therefore give a short account of the formulae most in use, without, however, entering upon their theoretical basis. Most of the work in this line was completed before the discovery of the modern quantum mechanics, but later developments have not led to divergent results.

As emphasized before, each multiplet level of inner quantum number J is itself multiple and splits up into $2J+1$ component levels in a magnetic field. We must therefore first consider the transitions between these sub-levels, which are distinguished by different values of the magnetic quantum number m. These transitions will determine the intensities of the Zeeman components of the lines for the case of an infinitely weak field. This problem has been considered by Darwin† from the point of view of wave mechanics, while Born, Heisenberg, and Jordan,‡ and Heisenberg and Jordan|| have developed the theory from the angle of matrix mechanics. The formulae are quite the same as those derived by earlier authors by assuming the transition probabilities to be quadratic functions of m, and assuming the total intensity and the total polarization of a multiplet to be unaltered by the influence of an infinitely weak field (principle of spectroscopic stability).

The formulae in question are derived on the assumption that it is sufficient to take into account the influence of the associated electric dipole moment of the electric oscillation in the atom. The theory has also been developed so as to take into account the effect of the associated magnetic dipole and electric quadripole. But as that

† *Proc. Roy. Soc.* A, **115** (1927), 1.
‡ *Zeits. f. Physik*, **35** (1926), 557.
|| Ibid. **37** (1926), 263.

question is of special interest to the theory of nebulae, it will be reserved for special consideration later on. We shall therefore here only recall the form of the simpler formulae, without entering on their theoretical foundation.

Let $A^{Jm}_{J'm'}$ denote the probability of a spontaneous jump from a state J, m to a state J', m'. The formulae in question have then the following form:

Transition $J \to J$:

$$A^{Jm}_{J,m\pm1} = \tfrac{1}{2}A^J_J \frac{(J\mp m)(J\pm m+1)}{J(J+1)},$$

$$A^{J,m}_{J,m} = A^J_J \frac{m^2}{J(J+1)}.$$

$J \to J+1$:

$$A^{J,m}_{J+1,m\pm1} = \tfrac{1}{2}A^J_{J+1} \frac{(J\pm m+2)(J\pm m+1)}{(J+1)(2J+3)},$$

$$A^{J,m}_{J+1,m} = A^J_{J+1} \frac{(J+1)^2-m^2}{(J+1)(2J+3)}.$$

$J \to J-1$:

$$A^{J,m}_{J-1,m\pm1} = \tfrac{1}{2}A^J_{J-1} \frac{(J\mp m)(J\mp m-1)}{J(2J-1)},$$

$$A^{J,m}_{J-1,m} = A^J_{J-1} \frac{J^2-m^2}{J(2J-1)}.$$

The quantities A^J_J, etc., are independent of m and J, but may depend in any manner on the other quantum numbers.

In order to use these expressions for the analysis of Zeeman components of multiplet lines it is of course necessary to know the relative number of atoms in the different quantum states involved. If the excitation has a non-selective character, so that states with the same values of n, L, and J have the same probability, practically speaking, the above formulae will be directly applicable to the intensity problem. Thus the empirical rules of Ornstein, Burger, and Landé concerning the intensities of the Zeeman components are all contained in the above formulae.

Consider next the probabilities of transition between the unperturbed multiplet levels. It is now important, especially for astrophysical applications, to remember that the following formulae apply only to transitions dominated by a variable dipole moment, as was emphasized above. The formulae were given by various authors as a suitable generalization from the classical theory. The nomenclature adopted is that of Sommerfeld and Hönl, which is, for instance, also

used in Laporte's article on multiplets in *Handb. d. Astrophys.* III. **2**, p. 673.

Let P, Q, and R be defined by

$$P(J) = (J+L)(J+L+1)-S(S+1),$$
$$Q(J) = -(J-L)(J-L+1)+S(S+1),$$

and $$R(J) = J(J+1)+L(L+1)-S(S+1).$$

We have then for transitions $L \to L-1$:

$$A^{J,L}_{J-1,L-1} = \frac{P(J)P(J-1)}{4JL},$$

$$A^{J,L}_{J,L-1} = P(J)Q(J)\frac{2J+1}{4LJ(J+1)};$$

and for transitions $L \to L$:

$$A^{J,L}_{J,L} = R^2(J)\frac{(2J+1)(2L+1)}{4J(J+1)L(L+1)},$$

$$A^{J,L}_{J-1,L} = P(J)Q(J-1)\frac{2L+1}{4JL(L+1)}.$$

These formulae are due to Kronig,† Sommerfeld, Hönl,‡ and Russell.‖

† *Zeits. f. Physik*, **31** (1925), 885.
‡ *Berl. Ber.* (1925), 141; cf. also *Ann. d. Physik*, **79** (1926), 273.
‖ *Nature*, **115** (1925), 835; *Proc. Nat. Acad. Washington*, **11** (1925), 314, 322.

VIII

THEORY OF ABSORPTION AND EMISSION

28. Energy of the Radiation Field

CONSIDER a radiation field which is constrained to remain inside a finite enclosure with perfectly reflecting walls. For the sake of simplicity we assume the enclosure to be of rectangular shape, parallel to the axes of a rectangular coordinate system, the coordinates of the edges along the axes being 0, α, β, γ respectively.

In such an enclosure the amplitude of a harmonic component of the field must have the form

$$X = A \cos 2\pi\nu t \sin k\pi\frac{x}{\alpha} \sin l\pi\frac{y}{\beta} \sin m\pi\frac{z}{\gamma},$$

where A is an arbitrary constant, and k, l, and m are positive integers. The amplitude will now necessarily vanish on the walls of the box, and may consequently be assumed to be zero on the outside. Since X must satisfy the wave equation

$$\nabla^2 X = \frac{1}{c^2}\ddot{X},$$

it follows that the frequency ν is given by

$$\frac{4\nu^2}{c^2} = \frac{k^2}{\alpha^2} + \frac{l^2}{\beta^2} + \frac{m^2}{\gamma^2}.$$

Each triad of integers k, l, m will correspond to a possible solution, and the total number of possible solutions for which the frequency is less than ν is given by the volume of the positive octant of the ellipsoid defined by the above equation when k, l, and m are considered as rectangular coordinates. This volume is

$$N = \frac{1}{8}\cdot\frac{4\pi}{3}\left(\frac{2\nu\alpha}{c}\right)\left(\frac{2\nu\beta}{c}\right)\left(\frac{2\nu\gamma}{c}\right) = \frac{4\pi}{3}\Omega\frac{\nu^3}{c^3},$$

where Ω is the volume of the enclosure. In case of light waves the number will be twice this because of the transverse nature of the vibrations, so we add a factor 2. Differentiating, we find the number of vibrations in the frequency interval ν to $\nu+d\nu$ to be

$$dN = 8\pi\Omega\frac{\nu^2 d\nu}{c^3}. \tag{45}$$

Assume that the radiation field is defined by its vector potential,

and let us resolve this potential into harmonic components in a slightly different way, a component being given by

$$\mathbf{A} = \mathbf{A}_0 u(t) \sin \frac{2\pi\nu}{c}(\alpha x + \beta y + \gamma z + \epsilon). \tag{46}$$

Here $\mathbf{A}_0$ is a unit vector, α, β, γ are the direction cosines of the direction of propagation of the wave, ϵ is a phase constant, which is to be determined so as to make the amplitude vanish at the boundary, and $u(t)$ is a function of the time, which reduces to a cosine for a wave which is also harmonic in the time, as assumed above, but which otherwise may be slightly different when there is absorption or emission of radiation. The complete radiation field will now be represented by a sum of standing vibrations of the above type, the number of possible vibrations in the frequency interval $d\nu$ being given by (45), while the directions of propagation and polarization are distributed at random. It is evident that knowing all the u-functions of the field the vector potential is determined, and then also, by a simple differentiation process, the electric and magnetic intensities. Hence we may consider the u-functions as the *coordinates* of the radiation field.

The electric and magnetic field vectors $\mathbf{E}$ and $\mathbf{H}$ are given by

$$\mathbf{E} = -\frac{1}{c}\frac{\partial \mathbf{A}}{\partial t}, \qquad \mathbf{H} = \operatorname{curl} \mathbf{A},$$

or expressed in harmonic components by (46):

$$\mathbf{E} = -\sum_s \frac{1}{c}\mathbf{A}_{0s}\dot{u}_s \sin \Gamma_s; \qquad \mathbf{H} = \sum_s \frac{2\pi\nu_s}{c}\mathbf{H}_s \times \mathbf{A}_{0s} u_s \cos \Gamma_s,$$

where $\mathbf{H}_s$ is a unit vector with components α_s, β_s, γ_s, and Γ_s is given by

$$\Gamma_s = \frac{2\pi\nu_s}{c}(\mathbf{H}_s \mathbf{r}) = \frac{2\pi\nu_s}{c}(\alpha_s x + \beta_s y + \gamma_s z).$$

The electromagnetic energy in the space Ω is

$$W = \frac{1}{8\pi}\Omega(\mathbf{E}^2 + \mathbf{H}^2).$$

Leaving out cross-terms, which only entail rapid fluctuations in the field of the period of the light waves, we may write

$$\mathbf{E}^2 = \frac{1}{2c^2}\sum_s \dot{u}_s^2; \qquad \mathbf{H}^2 = \sum_s \frac{2\pi^2\nu_s^2}{c^2} u_s^2,$$

which give the following expression for the energy W

$$W = \frac{\Omega}{8\pi c^2} \sum_s (\tfrac{1}{2}\dot{u}_s^2 + 2\pi^2 \nu_s^2 u_s^2).$$

Introducing the Hamiltonian momenta of the field,

$$v_s = \frac{\partial W}{\partial \dot{u}_s} = \frac{\Omega}{8\pi c^2} \dot{u}_s,$$

the energy assumes the following form

$$W = \sum_s \left(\frac{4\pi c^2}{\Omega} v_s^2 + \frac{\Omega}{8\pi c^2} 2\pi^2 \nu_s^2 u_s^2 \right).$$

The consequent canonical equations are

$$\dot{u}_s = \frac{\partial W}{\partial v_s} = \frac{8\pi c^2}{\Omega} v_s; \qquad \dot{v}_s = -\frac{\partial W}{\partial u_s} = -\frac{\Omega}{8\pi c^2} 4\pi^2 \nu_s^2 u_s.$$

Eliminating v_s we find

$$\ddot{u}_s + (2\pi\nu_s)^2 u_s = 0,$$

which shows, as was to be expected when there is no interaction energy with matter, that the u-functions are harmonic in the time with frequency ν. It is convenient for the following to make a slight change in the definition of the canonical variables by writing

$$u_s = \sqrt{\left(\frac{8\pi c^2}{\Omega}\right)} q_s; \qquad v_s = \sqrt{\left(\frac{\Omega}{8\pi c^2}\right)} p_s.$$

In the new variables p_s and q_s the energy assumes the form

$$W = \sum_s \{\tfrac{1}{2} p_s^2 + 2\pi^2 \nu_s^2 q_s^2\},$$

which is the Hamiltonian for a system of independent harmonic oscillators of unit mass and frequencies $\nu_1, \nu_2, \ldots$. The vector potential of the field assumes the consequent form

$$\mathbf{A} = \sqrt{\left(\frac{8\pi c^2}{\Omega}\right)} \sum_s \mathbf{A}_{0s} q_s \sin \Gamma_s.$$

29. Quantization of the Radiation Field

Since the radiation field is resolvable into a system of independent oscillators, the characteristic functions of the field are

$$\psi_{s'} = H_{s'}(kq) e^{-\frac{1}{2}(kq)^2} \qquad \left(k = 2\pi \sqrt{\frac{\nu}{h}}\right),$$

where $H_{s'}$ is a normalized Hermitian polynomial. Moreover, the characteristic energy values are given by

$$E_{s'} = h\nu_s (s' + \tfrac{1}{2}) \qquad (s' = 0, 1, 2, \ldots).$$

Consider next the wave-mechanical problem of an atomic system in interaction with this field. We write the Hamiltonian of the problem in the form

$$H = H_0 - \mathrm{K},$$

where H_0 is the Hamiltonian of the atom and the field, with neglect of their interaction, while K is the interaction term. Let ϕ_1, ϕ_2,... be the normalized characteristic functions of the unperturbed problem, i.e. corresponding to the Hamiltonian H_0 alone, and let E_1, E_2,... be the corresponding characteristic energy values. The most general solution of the unperturbed problem is then

$$\psi = \sum_n a_n \phi_n,$$

where the a_n are constants which we suppose chosen in such a way that $a_n a_n^*$ is just equal to the number of atoms in the assembly which are in the nth quantum state. From the perturbation theory which was developed in a preceding section (see p. 26) we know that the quantities a_m must satisfy the equations

$$\frac{da_m}{dt} - \frac{2\pi i}{h} \sum_n a_n \mathrm{K}_{nm} = 0 \qquad (m = 1, 2...),$$

where

$$\mathrm{K}_{nm} = \int \phi_m^* \mathrm{K} \phi_n \, d\tau, \tag{47}$$

is the matrix of the interaction energy.

Neglecting the square of the vector potential, the term $\frac{1}{2\mu}\left(\mathbf{p} - \frac{e}{c}\mathbf{A}\right)^2$ in the Hamiltonian reduces to $\frac{1}{2\mu}\mathbf{p}^2 - \frac{e}{\mu c}\mathbf{A}\mathbf{p}$. The interaction term in question is therefore

$$\mathrm{K} = \frac{e}{\mu c}\mathbf{A}\mathbf{p}.$$

In the case of a single electron $\mathbf{p}/\mu$ stands for the momentum of a single electron, divided by its mass, that is, in the language of wave mechanics,

$$\frac{p_x}{\mu} = \frac{h}{2\pi i \mu} \frac{\partial}{\partial x}, \quad \text{etc.}$$

In the case of several electrons K is simply the sum of $\frac{e}{\mu c}\mathbf{A}\mathbf{p}$ over all electrons. The vector potential is in each case to be calculated at the positions of the respective electrons. In the following we proceed as if there were but one electron present, since the generalization to several electrons is trivial.

Introducing the expansion for the vector potential given earlier we find

$$\mathrm{K} = \frac{e}{\mu}\sqrt{\frac{8\pi}{\Omega}}\sum_s q_s \sin\Gamma_s(\mathbf{A}_{0s}\mathbf{p}).$$

Expression (47) for the perturbation matrix may be further reduced by the following considerations. First it may be recalled that the wave functions ϕ_n, supposed to be calculated by neglect of the interaction energy, may be written in the form

$$\phi_n = \zeta_{n_0}\zeta_{n_1}\zeta_{n_2}\cdots,$$

where ζ_{n_0} is the wave function of the atom, and ζ_{n_1}, ζ_{n_2}, etc., are wave functions of the harmonic components of the radiation field. Each term in the sum giving K contains a q_s as a factor, and each term will be multiplied by the corresponding product $\zeta_{n_s}\zeta^*_{m_s}$, the quantum states being, say, n_s and m_s. Apart from these particular states, the wave functions will occur in pairs $\zeta_{n_{s'}}, \zeta^*_{m_{s'}}$, and it is therefore clear that each term in the sum giving K_{nm} will be a product of integrals of the form

$$I_{n_s m_s} = \int_{-\infty}^{\infty} \zeta_{n_s} q_s \zeta^*_{m_s}\, dq_s,$$

and

$$I'_{n_s m_s} = \int_{-\infty}^{\infty} \zeta_{n_s}\zeta^*_{m_s}\, dq_s.$$

The latter integral will vanish in all cases except when $n_s = m_s$, when it equals unity. By the theory of the harmonic oscillator the former integral is given by

$$I_{n_s m_s} = \sqrt{\left(\frac{h(n_s+1)}{8\pi^2\nu_s}\right)} e^{2\pi i\nu_s t} \qquad (m_s = n_s+1),$$

or

$$I_{n_s m_s} = \sqrt{\left(\frac{hn_s}{8\pi^2\nu_s}\right)} e^{-2\pi i\nu_s t} \qquad (m_s = n_s-1),$$

where

$$\nu_s = \left|\frac{E_{n_s} - E_{m_s}}{h}\right|.$$

As a consequence of this the series representing a perturbation matrix element reduces to a single term

$$\mathrm{K}_{n_0, n_1 \ldots n_s \ldots; m_0, n_1 \ldots n_s \pm 1 \ldots}$$
$$= \frac{e}{\mu}\sqrt{\left(\frac{h}{\pi\Omega\nu_s}\right)} \int \sin\Gamma_s \zeta^*_{m_0}(\mathbf{A}_{0s}\mathbf{p})\zeta_{n_0}\, d\tau \begin{cases} e^{2\pi i\nu_s t}\sqrt{(n_s+1)} \\ e^{-2\pi i\nu_s t}\sqrt{n_s}, \end{cases}$$

where the upper or lower expression is to be taken according as $m_s = n_s+1$ or n_s-1. Write

$$f_{n_0 m_0}(\nu_s) = \int \sin \Gamma_s \zeta^*_{s m_0}(\mathbf{A}_{0s}\nabla)\zeta_{n_0}\, d\tau.$$

The above expression may then be written in the form

$$\mathrm{K}_{n_0, n_1 \ldots n_s \ldots;\, m_0, n_1 \ldots n_s \pm 1 \ldots} = \frac{h}{2\pi i}\frac{e}{\mu}\sqrt{\left(\frac{h}{\pi\Omega\nu_s}\right)} f_{n_0 m_0}(\nu_s)\begin{bmatrix} e^{2\pi i \nu_s t}\sqrt{(n_s+1)} \\ e^{-2\pi i \nu_s t}\sqrt{n_s} \end{bmatrix}.$$

The equation for the time variation of the a_n-coefficients now becomes

$$\dot{a}_{m_0, n_1 \ldots n_s \ldots} = \sum_{n_0, s}\frac{e}{\mu}\sqrt{\left(\frac{h}{\pi\Omega\nu_s}\right)} f^0_{n_0 m_0}(\nu_s)\{a_{n_0, n_1 \ldots n_s+1}\sqrt{(n_s+1)}\, e^{-2\pi i(\nu_{n_0 m_0}+\nu_s)t} + \\ + a_{n_0, n_1 \ldots n_s-1}\sqrt{n_s}\, e^{-2\pi i(\nu_{n_0 m_0}-\nu_s)t}\},$$

where $f^0_{n_0 m_0}(\nu_s)$ denotes $f_{n_0 m_0}(\nu_s)$ without the complex time factor, while $\nu_{n_0 m_0}$ stands for $(E_{n_0}-E_{m_0})$. These equations may be considered the essential basis of the quantum theory of radiation processes. The special cases of pure absorption or pure scattering are to be understood not as intrinsically different processes, but more as different aspects of one and the same radiation process.

30. Structure of an Emission Line

The simplest possible case is encountered when the system consists of an atom initially in empty space, from which it may suffer transitions to a single lower state under emission of radiation. Denote the upper state by index 2 and the lower by index 1 and write

$$\nu_{21} = -\nu_{12} = \nu, \qquad f^0_{12}(\nu_s) = f_1(\nu_s), \qquad f^0_{21}(\nu_s) = f_2(\nu_s).$$

Neglecting consistently all other terms than the resonance terms, the only surviving equations are

$$\dot{a}_{1,0,\ldots,1_s} = \frac{e}{\mu}\sqrt{\left(\frac{h}{\pi\Omega\nu_s}\right)} f_2(\nu_s) a_{2,0,0,\ldots,} e^{-2\pi i(\nu-\nu_s)t}, \tag{48}$$

$$\dot{a}_{2,0,\ldots,0} = \sum_s \frac{e}{\mu}\sqrt{\left(\frac{h}{\pi\Omega\nu_s}\right)} f_1(\nu_s) a_{1,0,\ldots,1_s} e^{-2\pi i(\nu_s-\nu)t}. \tag{49}$$

Integrating (48) over the time we find

$$a_{1,0,\ldots,1_s} = \frac{e}{\mu}\sqrt{\left(\frac{h}{\pi\Omega\nu_s}\right)} f_2(\nu_s) \int_0^t a_{2,0,\ldots,} e^{-2\pi i(\nu-\nu_s)t}\, dt. \tag{50}$$

Introducing this into (49) we get an integral equation for $a_{2,0,\ldots}$,

$$\dot{a}_{2,0,\ldots} = \sum_s \frac{e^2}{\mu^2}\frac{h}{\pi\Omega\nu_s} f_1(\nu_s) f_2(\nu_s) e^{2\pi i(\nu-\nu_s)t} \int_0^t a_{2,0\ldots,} e^{-2\pi i(\nu-\nu_s)t}\, dt.$$

The sum with respect to s may be changed into an integral over ν_s taken from zero to infinity. Then, using (45),

$$\dot{a}_{2,0,\ldots} = \frac{e^2}{\mu^2}\frac{8h}{c^3} \int_{\nu_s=0}^{\infty} \nu_s f_1(\nu_s) f_2(\nu_s) e^{2\pi i(\nu-\nu_s)t} \left\{ \int_0^t a_{2,0,\ldots,} e^{-2\pi i(\nu-\nu_s)t}\, dt \right\} d\nu_s. \quad (51)$$

Considering the secular decrease in the number of atoms in the upper state, it is clear that this equation will have a secular solution (i.e. a solution smoothed by leaving out periodic terms) of the form

$$a_{2,0,0\ldots} = e^{-\lambda t},$$

where λ is a constant. Introducing this expression of $a_{2,0,0\ldots}$ into equation (51) we find

$$\lambda = \frac{8e^2h}{\mu^2c^3} \int_0^{\infty} \nu_s f_1(\nu_s) f_2(\nu_s) \frac{1-e^{(\lambda+2\pi i(\nu-\nu_s))t}}{\lambda+2\pi i(\nu-\nu_s)}\, d\nu_s.$$

Write this integral in the form

$$I = \frac{1}{2\pi i} \int_0^{\infty} \frac{\nu_s f_1(\nu_s) f_2(\nu_s) e^{(\lambda+2\pi i(\nu-\nu_s))t}}{(i\lambda/2\pi)+\nu_s-\nu}\, d\nu_s - \frac{1}{2\pi i} \int_0^{\infty} \frac{\nu_s f_1(\nu_s) f_2(\nu_s)(\nu_s-\nu)}{(\lambda^2/4\pi^2)+(\nu_s-\nu)^2}\, d\nu_s +$$

$$+ \frac{\lambda}{4\pi^2} \int_0^{\infty} \frac{\nu_s f_1(\nu_s) f_2(\nu_s)}{(\lambda^2/4\pi^2)+(\nu_s-\nu)^2}\, d\nu_s. \quad (52)$$

We may evaluate these integrals if we assume that λ is very small compared with ν (which will always hold in physically interesting cases).

The first question is what becomes of $f_{m_0 n_0}(\nu_s)$ when $\nu_s \to \infty$. In the expression for $f(\nu_s)$ written out in full,

$$f_{m_0 n_0}(\nu_s) = \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} \sin\frac{2\pi\nu_s}{c}(\alpha_s x+\beta_s y+\gamma_s z)\zeta^*_{n_0}(\mathbf{A}_{0s}\nabla)\zeta_{m_0}\, dxdydz,$$

we rotate the axes, till the new z-axis lies in the direction of α_s, β_s, γ_s. Then, performing the integration over the new coordinates x' and y'

we get the expression

$$f_{m_0 n_0}(\nu_s) = \int_{-\infty}^{+\infty} \sin\frac{2\pi\nu_s z'}{c} F(z')\, dz',$$

where

$$F(z') = \mathbf{A}_{0s} \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} \zeta^*_{n_0} \nabla \zeta_{m_0}\, dx'dy',$$

a finite function of z'.

By the integral theorem of Dirichlet:

$$\lim_{\nu\to\infty} \frac{1}{\pi} \int_{-\infty}^{+\infty} f(t+z)\sin\nu z \frac{dz}{z} = f(t),$$

we find

$$\lim_{\nu_s\to\infty} \int_{-\infty}^{+\infty} (0+z')F(0+z')\sin\frac{2\pi\nu_s z'}{c}\frac{dz'}{z'} = 0.F(0) = 0.$$

It is now easily shown that the first† and third of the integrals (52) converge, and it then lies in the nature of the problem that the second must also do so. We see further that the essential contribution to their values is found near the poles $\nu_s = \nu \mp i\lambda/2\pi$. We may, therefore, without perceptibly altering their values extend the lower limits to $\nu_s = -\infty$. As regards the first integral we may join the limits by a path enclosing the negative complex ν_s-plane. By Cauchy's theorem the integral is then equal to $-2\pi i$ times the value of $\nu_s f_1(\nu_s) f_2(\nu_s) e^{(\lambda + 2\pi i(\nu-\nu_s))t}$ at this point, which to a sufficient approximation is

$$-2\pi i \nu f_1(\nu) f_2(\nu).$$

The second integral vanishes owing to the factor $\nu_s - \nu$ in the numerator, the same approximations being assumed as in the preceding case. The third gives, similarly, a contribution $\frac{1}{2}\nu f_1(\nu) f_2(\nu)$. Hence

$$I = -\tfrac{1}{2}\nu f_1(\nu) f_2(\nu) \quad \text{and} \quad \lambda = -\frac{4e^2 h\nu}{\mu^2 c^3} f_1(\nu) f_2(\nu).$$

As the probability of the atom being in the state 2 is $a_{2,0,0,\ldots}\, a^*_{2,0,0,\ldots}$ it is clear that the corresponding Einstein A-coefficient is just $2R\lambda$,‡ or

$$A_{21} = -\frac{8e^2 h\nu}{\mu^2 c^3} R\{f_1(\nu) f_2(\nu)\}.$$

In all optical applications, that is, apart from X-ray spectra, the

† Charlier's convergence test, cf. Whittaker and Watson, *Modern Analysis*, p. 72.

‡ R = real part of. It follows below that λ is actually real, so in the following we need not distinguish $R\lambda$ from λ.

atomic dimensions are very small in comparison with the wave-length of radiation. This means that the sine function in $f_{n_s m_s}(\nu_s)$ may be taken outside the integral sign. The remaining integral is then given by the calculations in § 9:

$$f_{12}(\nu) = \sin\Gamma \int \zeta_2^* \mathbf{A}_0 \nabla \zeta_1 \, d\tau = -\sin\Gamma \frac{4\pi^2 \mu \nu}{h} \mathbf{A}_0 \mathbf{P}_{12},$$

$$f_{21}(\nu) = \sin\Gamma \int \zeta_1^* \mathbf{A}_0 \nabla \zeta_2 \, d\tau = \sin\Gamma \frac{4\pi^2 \mu \nu}{h} \mathbf{A}_0 \mathbf{P}_{12}^*,$$

where $\mathbf{P}_{12}$ is the harmonic amplitude associated with the two states in question. Averaging over all directions of the incident ray we may take $\overline{\sin^2\Gamma} = \frac{1}{2}$ and the average $\overline{|\mathbf{A}_0 \mathbf{P}_{12}|^2} = \frac{1}{3}|\mathbf{P}_{12}|^2$. The final expression of the Einstein A is then

$$A_{21} = \frac{64\pi^4 \nu^3 e^2}{3hc^3} |\mathbf{P}_{12}|^2, \tag{53}$$

in agreement with the result in § 9.

The line profile. We may now also find the distribution of energy in the emission line. This is the same as to find the probability distribution of the states $1,\ldots,1_s$. Since $a_{2,0,\ldots}$ is given by $e^{-\lambda t}$, equation (50) for $a_{1,0,\ldots,1_s,\ldots}$ may be integrated directly, the result being

$$a_{1,0,\ldots,1_s,\ldots} = 4\pi e \sqrt{\left(\frac{\pi}{\Omega \nu_s h}\right)} \mathbf{A}_{0s} \mathbf{P}_{12}^0 \nu \frac{1 - e^{-(\lambda + 2\pi i(\nu - \nu_s))t}}{\lambda + 2\pi i(\nu - \nu_s)} \sin\Gamma_s,$$

where we have introduced the value of $f_{21}(\nu_s)$ found above for the case of long waves. The integration constant has been so chosen as to make $a_{1,0,\ldots,1_s}$ vanish at $t = 0$. From this expression we find directly the probability of the state in question, as it is given by $|a_{1,0,\ldots,1_s}|^2$

In order to find the distribution of frequency, however, we must further multiply by the number of radiation components in the frequency interval ν_s to $\nu_s + d\nu_s$, which is $\Omega(8\pi\nu_s^2/c^3)\, d\nu_s$.

Assuming further the time t to be so large that $2e^{-\lambda t}$ may be neglected in comparison with unity, and writing the resulting distribution in the form $J_{\nu_s} d\nu_s$, we find

$$J_\nu = \frac{2\lambda}{\lambda^2 + 4\pi^2(\nu - \nu_s)^2}.$$

This gives an energy distribution in the line which is the same as for a classical oscillator with damping constant λ.

31. Probability Distribution Law of Weisskopf and Wigner

The above calculation refers only to the special case of the emission of a resonance line, as due to some kind of nondescript excitation. The result is, strictly speaking, not applicable, unless it is verified beforehand that the excitation in question corresponds to the assumptions, and it is not certain that it can be applied without further modifications to the emission of subordinate lines. Further work by Weisskopf and Wigner† seems to show fairly conclusively that the result will hold under very general conditions, when some precautions are kept in mind. If in fact we include a series of states in the calculations, the equations are approximately satisfied by assuming solutions in exponential terms, with suitably chosen damping constants, to be determined from the equations themselves. It turns out that the form of the line to be expected is in most cases that of a classical linear oscillator of the frequency of the line, and with a damping constant equal to half the sum of all transition probabilities from both states involved in the production of the line. This is a result which is of great importance for astrophysical applications.

Weisskopf and Wigner gave an interpretation of this result which is of considerable interest for the systematic development of the radiation theory, as it lends some physical insight into the meaning of the mathematical formalism. It is in fact a little disconcerting to find that the spectral lines have a finite width, although we assume the radiation to take place according to the strict frequency relation $h\nu_{kj} = E_k - E_j$. If we do not want to meddle with this relation, there is only one way to go, which is to admit that the energy values E_k and E_j of the two states involved are not sharply defined, but show a probability distribution round a mean value.

This is also what would be expected on the basis of the Heisenberg uncertainty relation. For let E be the energy of an atom in a stationary state and τ the life period of this state. The uncertainty relation then demands that the uncertainty in E, ΔE say, cannot be made smaller than h/τ. But $1/\tau$ is equal to the probability of transition, A, from the state, so that the uncertainty in the energy is of the order of hA, and that of the frequency of the radiation emitted during a transition from a state 1 to a state 2 is, correspondingly, $A_1 + A_2$.

Weisskopf and Wigner were able to substantiate this general

† V. Weisskopf and E. Wigner, *Zeits. f. Physik*, **63** (1930), 54.

inference, and found the following distribution law for the energy of a state:

$$W_k(E)\,dE = \frac{A_k}{h}\,\frac{dE}{\frac{1}{4}A_k^2+4\pi^2(E-E_k)^2/h^2}, \tag{54}$$

where A_k is the probability of transition from the given state, the mean energy of which is E_k. The state with a definite value E of the energy will be called a sub-state of the given state.

We shall now consider some applications of this formula, which will show its far-reaching adaptability in the radiation theory. We shall first derive the expression given earlier for the profile of an emission line with a nondescript excitation. Let the upper state be denoted by k, and assume the transition to take place to the k-state, with the probability A_{kj}. We assume that the transition ends exactly in the sub-state with energy E' in the close vicinity of E_j. The probability of a transition ending in a particular sub-state is supposed to be independent of the particular sub-state from which the transition took place. The probability that the atom shall have an energy in the interval dE before and in dE' after the transition is therefore

$$W(E_k)W(E_j)\,dE\,dE' = \frac{A_kA_j}{h^2}\,\frac{dE\,dE'}{\{\frac{1}{4}A_k^2+4\pi^2(E-E_k)^2/h^2\}\{\frac{1}{4}A_j^2+4\pi^2(E'-E_j)^2/h^2\}}. \tag{55}$$

The frequency of the radiation emitted (or absorbed, in the case of the converse transition) is

$$\nu = (E-E')/h.$$

The probability that for initial energy E the frequency of the emitted radiation will be found in the interval ν to $\nu+d\nu$ is obtained from (55) by writing $dE' = -h\,d\nu$ and then integrating the probability expression over all values of E, keeping the difference $E-E' = h\nu$ constant. Introducing the notation

$$x = E, \quad p = E_k, \quad q = E_j+h\nu, \quad a = \frac{hA_k}{4\pi}, \quad b = \frac{hA_j}{4\pi},$$

the probability expression assumes the form

$$J_\nu\,d\nu = \frac{h\,d\nu}{\pi^2}\int_{-\infty}^{\infty}\frac{ab\,dx}{\{a^2+(x-p)^2\}\{b^2+(x-q)^2\}}.$$

The integral is most readily evaluated by complex integration, the path of integration being along the real axis in the positive direction,

and then counter-clockwise along a half-circle lying at infinity and joining the positive end of the x-axis with the negative. This latter part of the integral will give no contribution to its value. The whole closed line integral is then by Cauchy's theorem equal to $2\pi i$ times the sum of the residues of the integrand at the poles inside the enclosed area. In this case the integrand has two poles inside this area, which are

$$x_1 = p+ai \quad \text{and} \quad x_2 = q+bi,$$

and the corresponding residues are

$$\frac{-\frac{1}{2}bi}{b^2+(p-q+ai)^2} \quad \text{and} \quad \frac{-\frac{1}{2}ai}{a^2+(q-p+bi)^2}.$$

The sum of these two terms is easily found by observing that each of the denominators may be resolved into a product of two linear factors, of which one, $p-q+i(a-b)$, is common to both residues. The two remaining factors are respectively $p-q+i(a+b)$ and $p-q-i(a+b)$, and the sum of the two residues becomes

$$\frac{-\frac{1}{2}i}{p-q+i(a-b)}\left(\frac{b}{p-q+i(a+b)}+\frac{a}{p-q-i(a+b)}\right) = \frac{-\frac{1}{2}i(a+b)}{(a+b)^2+(p-q)^2}.$$

The final form of the integral is therefore

$$J_\nu d\nu = \frac{h\,d\nu}{\pi}\,\frac{a+b}{(a+b)^2+(p-q)^2} = \frac{(A_k+A_j)\,d\nu}{\frac{1}{4}(A_k+A_j)^2+4\pi^2\{\nu-(E_k-E_j)/h\}^2}.$$

When comparing this result with the one derived directly from the radiation theory it must be remembered that we then assumed the transition probability A_j to be zero. Otherwise the present result is in conformity with the earlier calculations. In order to obtain the intensity of the line it is only necessary to multiply the above expression by $h\nu_{kj}A_{kj}$.

32. Dispersion Theory

Thus far we have only been considering the form of emission lines. In astrophysical problems, however, we are most frequently concerned with absorption processes, and it is of vital importance that the student of astrophysics should be familiar with all aspects of such problems, even if the direct astrophysical application of the result may be excluded by the mathematical complexity of the problem.

Let us first of all consider absorption of radiation in the first

resonance line of the atom, that is, the line connecting the two lowest energy states. We assume a plane monochromatic beam of light of a frequency in the proximity of the relevant line to be passed through a layer of atoms in the gaseous state at a sufficiently low density. According to everyday experience we then know that the influence of the gas on the radiation will only consist in a scattering of the radiation over different directions, without sensible change in frequency or loss in energy. It is of considerable interest to understand in detail how this result is brought about.

Let us consider the case from the point of view of Weisskopf and Wigner. The two relevant states of the atom, A and B say, will each have a probability distribution of the standard form (54). When a light ray of frequency ν strikes the atom it will presumably find this latter in some sub-state of A', say with an energy E. If the atom suffers a transition as an effect of the encounter, it will become transferred to a sub-state of B, with an energy $E' = E + h\nu$. We assume the density of the gas so low that the atom does not have any chance of suffering any collisions with other atomic particles during the time it resides in the upper state, so that the only thing that can happen is that it will return to state A. Here, however, it has a choice between a continuous number of sub-states, each corresponding to a different value of the frequency of the scattered radiation. The scattered radiation cannot therefore take the form of an infinitely sharp line, even if we consider the ideal case when the incident light is strictly monochromatic. This intrinsic width of the re-emitted line puts a limit to the ability of the scattered light to form interference patterns, and, on the whole, to exhibit the phenomena usually called refraction and dispersion. When such phenomena show up terrestrially to such an extent as they really do, it is due to the fact that the life period of an atom in a normal state is under usual experimental conditions, practically speaking, infinitely large. As long as ordinary collisions and the thermal Doppler effect may be neglected, the limit to the life period of this state is put by the intensity of the incident light itself.

Unfortunately, it is not possible to consider this problem properly on using the probability distribution law (54) directly, since we should have to deal with the case when two states coincide, which was supposed not to occur in the derivation of the formulae. Weisskopf, however, carried through the calculations directly, and found the

energy distribution in the re-emitted light to be given by the formula

$$J(\nu', \nu) = \frac{\text{const.}}{[(\gamma_A)^2+(\nu-\nu')^2][(\gamma_B)^2+(\nu_{AB}-\nu')^2]}. \qquad (56)$$

Here $4\pi\gamma_A$ and $4\pi\gamma_B$ are the total transition probabilities from the two states respectively, ν is the frequency of the incident light and ν' that of the scattered light. It will be noticed that the intensity has *two* maxima, one when $\nu' = \nu$ and another when $\nu' = \nu_{AB}$. The process of dispersion, even if the frequency of the incident radiation is well outside the centre of the line, will therefore necessarily also be associated with an emission of the line itself. That this must be so follows immediately from a consideration of the probability distribution of the two states. When, nevertheless, the scattered light usually contains by far the greater part of the energy, it is due to the much greater damping constant of the upper state. Assume then that ν is well outside the region of the line, so that the two maxima of (56) are clearly separated. The scattered light will then definitely pose as a doublet line, and the intensity will be the sum of the intensities of the components. Consider for instance the component with ν as a centre. We may then obtain a fair approximation by putting $\nu' = \nu$ in the second factor of the denominator, and integrate the expression over ν' for all frequencies. Since

$$\int_{-\infty}^{\infty} \frac{dx}{\gamma_A^2+x^2} = \frac{\pi}{\gamma_A},$$

this gives for the intensity

$$J_\nu = \frac{1}{\gamma_A} \frac{\text{const.}}{(\gamma_B)^2+(\nu-\nu_{AB})^2}.$$

Proceeding in the same way with the other component we find its intensity to be

$$J_{AB} = \frac{1}{\gamma_B} \frac{\text{const.}}{(\gamma_A)^2+(\nu-\nu_{AB})^2}.$$

The intensities of the doublet will thus be sensibly in the ratio of the corresponding transition probabilities, and since γ_B is very much larger than γ_A the radiation in the line itself will be insignificant, so that practically the whole energy will appear in the line round the frequency of the incident light.

These considerations serve to show how intimately the ordinary phenomena of dispersion are connected with the fact that the first resonance line of the atoms only is involved. Although we do not

possess a detailed theory as yet, it is clear that we must expect the atoms to scatter light in the vicinity of subordinate lines in a very different manner. $4\pi\gamma_A$ will still continue to denote the total probability of transition from the lower state of the line, also including the probability of spontaneous transitions, or any other influence shortening the life period of the state, and likewise for $4\pi\gamma_B$. In the case of subordinate lines γ_A will usually be larger than γ_B. In consequence of this the scattered light will be mostly re-emitted as if it were due to ordinary spontaneous transitions from the upper state. The line surrounding the frequency of the incident light will be relatively broad, and of inferior intensity. Further it must be remembered that the sum of the intensities of these doublets will no longer give the energy extracted from the incident beam, as an uncontrollable amount of radiation may be diverted into lines ending in lower energy states. It is clear, therefore, that the extension of results obtained by applying the ordinary theory of dispersion to subordinate lines must be eyed with suspicion, or, at any rate, administered with very great caution.

32. 1. Classical Theory of Dispersion

We have thus far made it rather a point to develop the consequences of the atomic theory directly from the wave-mechanical equations. At the present place it will be desirable, however, for the sake of historical continuity to consider also the classical dispersion theory. This latter assumes two different forms, according as one considers directly the emission of radiation from a system of electric particles with a given polarization, or considers the motion of an individual electron subjected to a damping force, due to the emission of radiation.

Let us consider the latter case first, and assume the electron to vibrate elastically, except for the decay which is due to its emission of radiation. The latter is known to be equivalent to a force, proportional to the rate of change of the acceleration:

$$\frac{2}{3}\frac{e^2}{c^3}\dddot{x},$$

where x is the linear coordinate determining the position of the electron. The equation of motion is then

$$\mu\ddot{x} = -\mu(2\pi\nu_0)^2 x + \frac{2}{3}\frac{e^2}{c^3}\dddot{x},$$

where μ and e denote as usual mass and charge of the electron, and its proper frequency of vibration. Writing in this equation

$$x = \text{const.}\, e^{i\sigma t}$$

we find

$$\sigma^2 = (2\pi\nu_0)^2 + i\frac{2e^2\sigma^3}{3\mu c^3}.$$

Considering the imaginary term in this equation as a small correction only, we find to the first order

$$i\sigma = 2\pi i\nu_0 - \frac{4\pi^2 e^2 \nu_0^2}{3\mu c^3}.$$

The solution found is thus a damped oscillation. The time of decay, which is usually denoted by 2τ, is given by

$$2\tau = \frac{3\mu c^3}{4\pi^2 e^2 \nu_0^2}.$$

Then τ is the time of decay of the energy of the electron. For a frequency of the order 10^{15} sec.$^{-1}$, τ is of the order 10^{-8} sec. Expressed in terms of τ, the damping force due to the emission of radiation becomes

$$\frac{\mu\dddot{x}}{4\pi^2\nu_0^2\tau}.$$

Consider next the transmission of a monochromatic plane and plane-polarized light wave through a gaseous medium. The gas atoms are assumed to contain only one electron each, which when undisturbed performs damped harmonic oscillations with a frequency ν_0 and with a time of decay τ. Under the steady influence of the light wave the electrons will perform forced oscillations with the frequency of the incident wave. Neglecting a possible interaction between the electrons themselves the equation of motion of an electron will be

$$\mu\ddot{x} = eE - \mu(2\pi\nu_0)^2 x + \frac{\mu\dddot{x}}{4\pi^2\nu_0^2\tau}. \tag{57}$$

Here the electric field is supposed to have the intensity E, and to be directed along the x-axis, so that the steady motion of the electron will be along this axis only, x being the displacement of the electron from its position of equilibrium.

Considering the case in which the wave-length of the electric field is long in comparison with the amplitude of the electronic oscillations, we may assume E to be independent of x, and to be proportional to $e^{2\pi i\nu t}$. In the final steady state x will then be proportional to E, and

equation (57) reduces to

$$\left(\nu_0^2-\nu^2+\frac{i\nu^3}{2\pi\nu_0^2\tau}\right)x=\frac{eE}{4\pi^2\mu}.$$

In practice ν will always be so close to ν_0 that we may write $\nu=\nu_0$ in the imaginary term, and write $\nu_0^2-\nu^2=2\nu_0(\nu_0-\nu)$. Then

$$x=\frac{eE}{8\pi^2\mu\nu_0(\nu_0-\nu+i/4\pi\tau)}. \tag{58}$$

Let N denote the number of atoms per unit volume. The current density due to the motion of the electrons will then be $J=Ne\dot{x}$, or, by (58),

$$\mathbf{J}=\frac{Ne^2}{8\pi^2\mu\nu_0(\nu_0-\nu+i/4\pi\tau)}\frac{\partial\mathbf{E}}{\partial t}.$$

This current will generate a macroscopic electromagnetic field, which in combination with the incident wave will satisfy the Maxwellian equations

$$\begin{aligned}\operatorname{curl}\mathbf{H}&=\frac{1}{c}\frac{\partial\mathbf{E}}{\partial t}+\frac{4\pi}{c}\mathbf{J},\\ \operatorname{curl}\mathbf{E}&=-\frac{1}{c}\frac{\partial\mathbf{H}}{\partial t}.\end{aligned} \tag{59}$$

Eliminating $\mathbf{H}$ and introducing the above expression of $\mathbf{J}$, we find

$$\frac{\partial^2\mathbf{E}}{\partial x^2}=\frac{1}{c^2}\left(1+\frac{Ne^2}{2\pi\mu\nu_0[\nu_0-\nu+i/4\pi\tau]}\right)\frac{\partial^2\mathbf{E}}{\partial t^2}, \tag{60}$$

and a corresponding equation for the magnetic field strength $\mathbf{H}$. In this elimination process use is made of the fact that both $\operatorname{div}\mathbf{E}$ and $\operatorname{div}\mathbf{H}$ vanish, since the wave carries no electricity. The above equation represents a damped wave, and we may write the solution in the form

$$\mathbf{E}=\text{const.}\,e^{2\pi i\nu(t-nx/c)-\frac{1}{2}\kappa x},$$

where n is the index of refraction, and κ the absorption coefficient of the wave, relative to the flux of energy. Introducing this expression in (60) we find the following equation for the determination of κ and n:

$$(\tfrac{1}{2}\kappa+2\pi i\nu n/c)^2=-\frac{4\pi^2\nu^2}{c^2}\left(1+\frac{Ne^2}{2\pi\mu\nu_0(\nu_0-\nu+i/4\pi\tau)}\right);$$

or, separating real and imaginary quantities,

$$n^2-1=\frac{Ne^2(\nu_0-\nu)}{2\pi\mu\nu_0\{(\nu_0-\nu)^2+1/16\pi^2\tau^2\}}+\frac{\kappa^2c^2}{16\pi^2\nu^2},$$

and

$$n\kappa=\frac{1}{4\pi\mu c\tau}\frac{Ne^2}{(\nu_0-\nu)^2+1/16\pi^2\tau^2}.$$

Thus far we have been concerned with one electron per atom only. Let there now be f identical electrons to each atom. These electrons will, in the steady state, move in unison as if they formed one electron of mass $f\mu$ and electric charge fe. Hence the correct generalization is obtained by writing $f\mu$ for μ and fe for e in the above equations. Noting that the time of decay is then divided by f, it appears that the correct generalization is

$$n^2-1 = \frac{Ne^2f}{2\pi\mu\nu_0}\frac{\nu_0-\nu}{(\nu_0-\nu)^2+(f/4\pi\tau)^2}+\frac{\kappa^2c^2}{16\pi^2\nu^2}$$

and

$$n\kappa = \frac{Ne^2f^2}{4\pi\mu c\tau}\frac{1}{(\nu_0-\nu)^2+(f/4\pi\tau)^2} = \frac{4\pi Ne^2f^2}{\mu c\tau\{16\pi^2(\nu_0-\nu)^2+f^2/\tau^2\}}. \tag{60a}$$

The further generalization to atoms containing groups of electrons with different proper frequencies is not easily made by the above procedure, since the representation of the reaction of the radiation field by a force $\mu\dddot{x}/4\pi^2\nu_0^2\tau$ breaks down in that case. The point of interest in the above result is that, neglecting the term in κ^2 and the damping term in the denominator, we find n^2-1 to be proportional to f while the product $n\kappa$ is proportional to f^2. The total absorption in a line, on the other hand, will be proportional to f only. We find, namely, on assuming the index of refraction to be of order unity, even in the region of the line

$$\int_0^\infty \kappa\, d\nu \sim \pi f\frac{Ne^2}{\mu c}.$$

Comparing these results with those derived from the quantum theory, it is clear that the classical damping constant f/τ plays the part of the sum of the probabilities of transition from the two levels of the line, while $\pi fe^2/\mu c$ plays the part of the Einstein B-coefficient of absorption.

The general form of the absorption coefficient may also be found in a somewhat simpler way, which perhaps gives a better insight into the physical processes at work. We know that the radiation from a system of electrons is determined by the time derivative of the electrical momentum

$$P_x = fex, \text{ etc.},$$

according to the formula

$$\text{rate of loss of energy} = \frac{2}{3c^3}\ddot{\mathbf{P}}^2. \tag{61}$$

The variable electric momentum means that there is a current density $N\dot{\mathbf{P}}$ in the gas, which, when introduced into the Maxwellian equations (59) gives

$$\nabla^2\mathbf{E} = \frac{1}{c^2}(\ddot{\mathbf{E}} + 4\pi N\ddot{\mathbf{P}}).$$

The electrical momentum will, to a first approximation, satisfy the equation of motion

$$\mu\ddot{\mathbf{P}} = fe^2\mathbf{E} - \mu(2\pi\nu_0)^2\mathbf{P},$$

which gives in the usual way for the steady state ($P \propto E \propto e^{2\pi i\nu t}$)

$$\mathbf{P} = \frac{fe^2\mathbf{E}}{\mu(2\pi)^2(\nu_0^2 - \nu^2)}. \tag{62}$$

This makes $\mathbf{P}$ real, so the approximation is not sufficient for the determination of the absorption coefficient from the Maxwellian equations. But we may proceed some way by using (62), on reasoning as follows: The energy lost from the light per unit time and volume is by (61) and (62)

$$\frac{2(2\pi\nu)^4}{3c^3}\left(\frac{fe^2\mathbf{E}}{\mu 4\pi^2(\nu_0^2 - \nu^2)}\right)^2 N = \frac{2}{3}\frac{f^2e^4\nu^4\mathbf{E}^2N}{c^3\mu^2(\nu_0^2 - \nu^2)^2}.$$

Dividing this expression by the flux of radiation, which is here given by the Poynting vector $c\mathbf{E}^2/4\pi$, we find the absorption coefficient to be

$$\kappa = \frac{8\pi}{3}\frac{e^4f^2N}{\mu^2c^4}\frac{\nu^4}{(\nu_0^2 - \nu^2)^2}.$$

This expression is identical with the one derived earlier, N denoting the number of atoms per unit volume, on allowing the same approximations and neglecting the damping term in the denominator. However, we have above made no assumption with regard to the frequency ν and may in particular assume it to be very large. Then the expression approaches the limit

$$\kappa = \frac{8\pi}{3}\frac{e^4}{\mu^2c^4}f^2N,$$

which is Thomson's formula for the coefficient of scattering by free electrons. If the atom contains a series of electrons with different proper frequencies $\nu_1, \nu_2, \ldots$, and the corresponding f's being denoted by $f_1, f_2, \ldots$, the absorption coefficient becomes obviously

$$\kappa = \frac{8\pi}{3}\frac{e^4}{\mu^2c^4}N\left(\sum_i \frac{f_i\nu^2}{\nu_i^2 - \nu^2}\right)^2. \tag{63}$$

32. 2. Quantum Theory of Dispersion

Although it does not lead to any new results, it is of some interest to show how the above procedure may be translated into the language of the quantum theory, step for step. The radiation field will then be treated as a small perturbation, just as in the case of the theory of the Einstein coefficients considered in an earlier section.

We then assume that an atom with its centre at the origin of coordinates is exposed to the influence of a plane and plane-polarized light wave of frequency ν. The atom is supposed to contain one electron only, moving in a field of potential energy V, and the vector potential of the light wave is supposed to be directed along the x-axis and to be given by

$$A = 2A_0 \cos 2\pi\nu t = A_0(e^{2\pi i\nu t} + e^{-2\pi i\nu t}).$$

This means that the dimensions of the atom are small in comparison with the wave-length of light, otherwise a term linear in y and z should have been added to t in the above expression, which would have complicated the calculations very much.

The wave equation of the problem is (§ 9)

$$\nabla^2\psi - \frac{4\pi ie}{hc} A \frac{\partial\psi}{\partial x} - \frac{8\pi^2\mu}{h^2} V\psi + \frac{4\pi i\mu}{h} \frac{\partial\psi}{\partial t} = 0.$$

The wave function appropriate to a stationary state k we assume to be split into the unperturbed function

$$\phi_k e^{-(2\pi i E_k/h)t}$$

and the perturbation term u_k; ϕ_k being independent of the time. The functions ϕ_k will be assumed to be orthogonal and normalized in the usual way. Neglecting the product Au_k, which is of the second order, it is clear that u_k will satisfy the equation

$$\nabla^2 u_k - \frac{8\pi^2\mu}{h^2} V u_k + \frac{4\pi i\mu}{h} \frac{\partial u_k}{\partial t} = \frac{4\pi ie}{hc} A \frac{\partial\phi_k}{\partial x} e^{-(2\pi i E_k/h)t}.$$

As is usual in perturbation theory, we assume without further proof that the perturbation term to the right may be expanded into a convergent series of the functions ϕ_k of the form

$$\frac{4\pi ie}{hc} A \frac{\partial\phi_k}{\partial x} e^{-(2\pi i E_k/h)t} = \sum_n a_n (e^{-2\pi i((E_k/h)+\nu)t} + e^{-2\pi i(E_k/h-\nu)t})\phi_n,$$

where the constants a_n are given by

$$a_n = \frac{4\pi ei}{hc} A_0 \int \phi_n^* \frac{\partial\phi_k}{\partial x} d\tau.$$

It was shown earlier (§ 9, p. 28) that

$$\int \phi_n^* \frac{\partial \phi_k}{\partial x}\, d\tau = \frac{4\pi^2\mu}{h} \nu_{kn} (P_{kn}^0)_x, \qquad \text{that is,} \qquad a_n = \frac{16\pi^3 ie\mu}{h^2 c} A_0 \nu_{kn} (P_{kn}^0)_x,$$

where $(P_{kn}^0)_x$, which from now on will be denoted by P_{kn} simply, is the harmonic amplitude of the atom in the direction of the x-axis. We have already implicitly assumed the atom to contain no permanent electric moment, so the one we are considering here will be in the direction of the external field.

As a next step we assume u_k to have the form

$$u_k = v_k e^{-2\pi i(E_k + h\nu)t/h} + w_k e^{-2\pi i(E_k - h\nu)t/h},$$

where v_k and w_k depend on the coordinates only. The above form of u_k is clearly demanded by the dependence of the perturbation term on time. The two functions v_k and w_k will each satisfy a separate wave equation, but these equations will only differ from each other by the sign of ν. It is therefore sufficient to write down and solve the equation satisfied by v_k, which is easily found to be

$$\nabla^2 v_k + \frac{8\pi^2\mu}{h^2}(E_k + h\nu - V) v_k = \sum_n a_n \phi_n.$$

Assuming an expansion of the form

$$v_k = \sum_n b_n \phi_n$$

we find easily, by making use of the unperturbed equation for ϕ_n,

$$b_n = \frac{h}{8\pi^2\mu} \frac{a_n}{\nu_{kn} + \nu} = \frac{2\pi ieA_0}{hc} \frac{\nu_{kn} P_{kn}}{\nu_{kn} + \nu}.$$

The corresponding expansion coefficient of w_k is obviously obtained by writing $-\nu$ instead of ν in this expression. We may therefore at once write down the complete form of u_k

$$u_k = \frac{2\pi ieA_0}{hc} \sum_n \nu_{kn} P_{kn} \phi_n \left\{ \frac{e^{-2\pi i(E_k + h\nu)t/h}}{\nu_{kn} + \nu} + \frac{e^{-2\pi i(E_k - h\nu)t/h}}{\nu_{kn} - \nu} \right\}. \tag{64}$$

From this expression we may now form the induced electrical moment of the atom, which we shall denote by M_{kk}. Since the momentum is directed in the x-direction, we have, on neglecting second-order terms and assuming the unperturbed electric moment to be zero,

$$(M_{kk})_x = e \int (\phi_k u_k^* e^{-2\pi i E_k t/h} + \phi_k^* u_k e^{2\pi i E_k t/h}) x\, d\tau.$$

Introducing the expression of u_k given by (64) we find

$$\mathbf{M}_{kk} = \frac{2e^2}{h}\frac{4\pi\nu}{c}\mathbf{A}_0 \sin 2\pi\nu t \sum \frac{\nu_{kn} P_{kn}^2}{\nu_{kn}^2 - \nu^2}.$$

The emission of energy per unit time from the atom is now

$$\frac{2}{3}\frac{\ddot{\mathbf{M}}_{kk}^2}{c^3} = \frac{4\pi\nu^2}{c}\mathbf{A}_0^2 \sin^2 2\pi\nu t \frac{8\pi}{3}\frac{e^4}{\mu^2 c^4}\left(\sum_n \frac{8\pi^2\nu^2\mu\nu_{kn} P_{kn}^2}{h(\nu_{kn}^2 - \nu^2)}\right)^2.$$

Denoting the electric field intensity by $\mathbf{E}$, the flux of radiation is given by the Poynting vector $c\mathbf{E}^2/4\pi$. Now

$$\mathbf{E} = -\frac{1}{c}\frac{\partial \mathbf{A}}{\partial t} = \frac{4\pi\nu}{c}\mathbf{A}_0 \sin 2\pi\nu t,$$

so that

$$\frac{c\mathbf{E}^2}{4\pi} = \frac{4\pi\nu^2}{c}\mathbf{A}_0^2 \sin^2 2\pi\nu t.$$

Dividing the emission of energy by the above expression of $c\mathbf{E}^2/4\pi$ we obtain the coefficient of absorption of the atom,

$$\kappa = \frac{8\pi}{3}\frac{e^4}{\mu^2 c^4}\left(\sum_n \frac{8\pi^2\nu^2\mu\nu_{kn} P_{kn}^2}{h(\nu_{kn}^2 - \nu^2)}\right)^2.$$

Comparing this with the classical formula (63) it appears that the quantum theory equivalent of the 'number of dispersion electrons' f associated with a line is given by

$$f_{kn} = \frac{8\pi^2\mu\nu_{kn}}{h} P_{kn}^2.$$

In the classical theory the f's were naturally assumed to be integral numbers, but in the quantum theory this is no longer necessary, a feature which is confirmed by experiment. In fact, the f's cannot be integers, a feature which is brought out distinctly by the Thomas-Kuhn sum rule. This rule may be derived in a direct way from the quantum theory, but it also follows from the above formulae. In fact, the above expression of the scattering coefficient of an atom should hold the better the higher the frequency ν of the incident light. But if ν is sufficiently large, we may neglect ν_{kn} in the denominator, and there remains

$$\kappa = \frac{8\pi}{3}\frac{e^4}{\mu^2 c^4}\left(\sum_n f_{kn}\right)^2.$$

For sufficiently high frequencies we should expect the electron to

scatter as if it were free, and the above expression of κ to coincide with the Thomson formula. For this to be the case we must have

$$\sum_n f_{kn} = 1.$$

If the atom contains several electrons, the sum will be equal to the number of electrons in the atom. The sum is in all cases extended to all possible quantum states, the continuous states, when the electron is essentially free, being included. This is the Thomas-Kuhn sum rule.

33. Influence of the Doppler Effect

Since the atoms of a gas are in a state of thermal agitation, it is necessary to consider how the foregoing results are to be modified before applying them to actual gases.

Let m be the mass of an atom, u its velocity in the line of sight, and ν_0 its proper frequency. Relative to the observer this proper frequency will appear to be

$$\nu_0' = \nu_0(1-u/c).$$

The absorption coefficient of this atom for radiation of the frequency ν relative to the observer will then assume the form

$$\kappa' = \frac{K\delta}{(\nu-\nu_0')^2+\delta^2},$$

when the frequency ν is close to the apparent proper frequency ν_0' of a characteristic line of the atom. Further, $4\pi\delta$ is the damping constant f/τ, and K summarizes the rest of the constant factor of the coefficient.

By the above expression of ν_0' as a function of ν_0, the absorption coefficient assumes the form

$$\kappa' = \frac{K\delta}{(\nu-\nu_0+\nu_0 u/c)^2+\delta^2}.$$

The number of atoms per unit volume of the gas which have a velocity component in the interval u to $u+du$ is, by Maxwell's law,

$$dP = N\sqrt{\frac{m}{2\pi kT}}e^{-\frac{1}{2}mu^2/kT}\,du,$$

where N is the number of atoms in unit volume and T the temperature. Integrating over all possible velocities we find the average

absorption coefficient to be given by

$$\kappa = \int \kappa' \, dP = NK\delta \sqrt{\frac{m}{2\pi kT}} \int_{-\infty}^{\infty} \frac{e^{-\frac{1}{2}mu^2/kT} \, du}{(\nu - \nu_0 + \nu_0 u/c)^2 + \delta^2}.$$

In order to bring the integral into a simpler form we introduce the notation

$$\zeta = u\sqrt{\frac{m}{2kT}}, \qquad \eta = \frac{\nu_0 - \nu}{\nu_0} c\sqrt{\frac{m}{2kT}}, \quad \text{and} \quad D = \frac{\delta c}{\nu_0}\sqrt{\frac{m}{2kT}}.$$

The absorption coefficient then assumes the following form

$$\kappa = \frac{NK\delta}{\sqrt{\pi}\nu_0^2} \frac{mc^2}{2kT} \int_{-\infty}^{\infty} \frac{e^{-\zeta^2} \, d\zeta}{(\zeta - \eta)^2 + D^2}.$$

Two limiting cases are of special interest. First of all it is clear that if η is sufficiently large, ζ may be neglected in the denominator, in which case the original form of the absorption coefficient is obtained. In that case the thermal motion has no influence on the absorption of radiation. This holds at sufficient distances from the centre of the line.

Next we may consider the case when η and D both are small. The denominator will then have its minimum value ($\zeta = \eta$) very near zero, while the numerator will be relatively large, and change very slowly. Near $\zeta = \eta$ the integrand then runs up to a sharp maximum and down again during an interval of ζ in which the numerator does not change very much. In that case we may put $\zeta = \eta$ in the exponential term, take it outside the integral sign, and perform the remaining integration. The result is

$$\kappa = NK\sqrt{\pi}\frac{D}{\delta}e^{-\eta^2}. \tag{65}$$

The original profile is in this case completely lost. What remains is only a reflection of the Maxwellian velocity distribution of the atoms.

An approximate expression also for intermediate values of η is obtained as follows. We first integrate by parts, and obtain

$$\int_{-\infty}^{\infty} \frac{e^{-\zeta^2} \, d\zeta}{(\zeta - \eta)^2 + D^2} = \frac{2}{D} \int_{-\infty}^{\infty} \zeta e^{-\zeta^2} \arctan \frac{\zeta - \eta}{D} \, d\zeta,$$

since the integrated part vanishes at both limits. The limits of the

above integral may be changed from $-\infty$ and $+\infty$ to 0 and ∞ by writing

$$\arctan\frac{\zeta-\eta}{D}+\arctan\frac{\zeta+\eta}{D}=\arctan\frac{2\zeta D}{D^2+\eta^2-\zeta^2}$$

instead of $\arctan\{(\zeta-\eta)/D\}$ in the integrand. We further assume η to be so much larger than D that D^2 may be neglected in comparison with η^2.

For large values of ζ the above arctan-expression approaches the value π. For small ζ it may be approximated to by $2\zeta D/\eta^2$. The division between these regions is not very sharply defined, but may be taken at $\zeta = \eta$, since this corresponds to half the maximal value of the function. The consequent approximate form of the integral is

$$\int_{-\infty}^{\infty}\frac{e^{-\zeta^2}\,d\zeta}{(\zeta-\eta)^2+D^2}=\frac{\pi}{D}e^{-\eta^2}+\frac{4}{\eta^2}\int_0^{\eta}\zeta^2e^{-\zeta^2}\,d\zeta,$$

and the corresponding form of the absorption coefficient is

$$\kappa = N\mathrm{K}\sqrt{\pi}\frac{D}{\delta}\left(e^{-\eta^2}+\frac{4D}{\pi\eta^2}\int_0^{\eta}\zeta^2e^{-\zeta^2}\,d\zeta\right).$$

34. Experimental Tests

It is not easy to devise experiments for a direct test of the theoretical predictions regarding the dependence of the absorption coefficient on frequency in the region of a line. The theoretical connexion between absorption and dispersion provides means which are useful at sufficiently large distances from the core. As was pointed out by Ladenburg and Reiche,† a further test is provided by the total absorption in a line, a quantity which can be measured fairly easily. Consider, namely, the transmission of a parallel beam of continuous radiation of initial intensity I through a homogeneous medium. We assume the intensity I to be initially independent of frequency in the region of the line to be studied. The intensity of frequency ν after the beam has passed a distance s through the medium will be given by $Ie^{-s\kappa}$. The total loss of energy in the beam at this distance is therefore

$$J = I\int_0^{\infty}(1-e^{-s\kappa})\,d\nu. \tag{66}$$

† Ladenburg and Reiche, *Ann. d. Physik*, **42** (1913), 181; W. Schutz, *Zeits. f. Astrophys.* **1** (1930), 300.

The integral will assume different forms, according to the functional form of κ in the region where the exponential term drops most rapidly from its maximal value to zero. Consider first the case when κ in this critical region has the form demanded by the theory of damping by radiation alone, the width of the line being considerably greater than the damping constant, so that we may write

$$\kappa s = \frac{Nsf\delta e^2}{\mu c(\nu-\nu_0)^2},$$

ν_0 being the frequency at the centre of the line. Instead of ν it is now advantageous to take

$$u = \frac{1}{\sqrt{(s\kappa)}} = (\nu-\nu_0)\sqrt{\frac{\mu c}{Nsf\delta e^2}}$$

as an integration variable. The upper limit of the integral will then be $u = \infty$, while the lower limit will be $-\frac{\nu_0}{e}\sqrt{\frac{\mu c}{Nsf\delta}}$. By a suitable choice of N or s this quantity may easily be made so large numerically that we may put the lower limit equal to minus infinity. The quantity J now assumes the form

$$J = I\sqrt{\frac{Nsf\delta e^2}{\mu c}}A,$$

where A is a numerical constant of order unity

$$A = \int_{-\infty}^{\infty}(1-e^{-u^2})\,du.$$

It appears that the total absorption is proportional to the square root of $Ns\delta f$, which is the result first derived by Ladenburg and Reiche.

Consider next the case when the absorption line is so narrow that the influence of the Doppler effect is predominating for the form of the absorption coefficient. Using expression (65) for this latter we find the following expression for the total absorption:

$$J = I\int_0^{\infty}(1-e^{-ae^{-\eta^2}})\,d\nu = I\frac{\nu_0}{\eta_0}\int_{-\eta_0}^{\infty}(1-e^{-ae^{-\eta^2}})\,d\eta,$$

where $$\eta_0 = c\sqrt{\frac{m}{2kT}}, \quad \text{and} \quad a = \sqrt{\pi}N\mathrm{K}\frac{D}{\delta}s.$$

The quantity η_0, being of the order of the ratio of the velocity of light to that of a moving particle, is so large that the lower limit of

the integral may be taken equal to infinity. The dependence of the total absorption on N, s, and K is through the constant a only. An exact integration of the expression is difficult, and neither is it necessary, since it is clear that an approximate value of the integral will be given by twice the value of η which makes the exponential term of the order $\frac{1}{2}$. This η is given by

$$\eta' = \sqrt{\log\left[\frac{a}{\log 2}\right]},$$

so that

$$J = 2I\frac{\nu_0}{\eta_0}\sqrt{\log\left[\frac{a}{\log 2}\right]},$$

approximately. Very large variations in N, s, or the intrinsic strength of the line will thus be necessary in order to produce a sensible increase in total absorption in this region.

We may finally mention also the trivial case when κs is small in comparison with unity. The exponential in (66) may then be expanded, and the expansion reduced to the two first terms only. It is then evident that the total absorption will be proportional to $\int_0^\infty s\kappa\, d\nu$. But this case is too difficult to realize experimentally to be of much value.

The important result of these considerations is the fact that the total intensity of broad lines will follow the square root law, while narrow lines, which show the effect of Doppler broadening, will rather tend to be independent of density and optical path. This result depends in the first instance on the functional form of the absorption coefficient in the two cases, and we shall therefore see later on that the same result is obtained even in cases when the external conditions are so much different that other results would have seemed more likely.

IX

TRANSFER OF RADIATION IN A STAR

35. The Equation of Radiative Transfer

In the preceding section we have considered the detailed mechanism of interaction processes involving at the same time both atoms and a field of radiation from the point of view of atomic theory. In order to make this information available for the interpretation of astrophysical facts we need, however, a macroscopic complement to this theory, consisting in a study of the transformation suffered by a field of radiation sifting through a medium of varying temperature and chemical composition. For this macroscopic problem it is unnecessary to retain the machinery of quantization of the radiation. It is sufficient to consider a parallel beam of radiation and to control its depletion or growth as it moves along, quantum theory being consulted only when it comes to the introduction of the expressions for the coefficients of absorption and emission in the final formula.

Consider thus the transmission of radiation through an absorbing and emitting medium, which we assume to be permanently at rest. The case when the medium is in motion is also of importance, but may be postponed for later treatment. We shall also in the preliminary discussions assume the refractive index of the medium to be unity, which means a great simplification. Let

$$E_\nu \, d\nu d\omega dV$$

denote the energy of radiation within the frequency interval ν to $\nu+d\nu$, within the solid angle $d\omega$, emitted per second at the time t, from matter enclosed in the infinitesimal volume element dV, supposed to be situated at the point x, y, z. We denote further by $\varkappa_\nu$ the absorption coefficient of matter for radiation of the given quality, defined in such a way that $\varkappa_\nu ds$ is the relative weakening of a linear pencil of rays passing a distance ds through the medium. The behaviour of the radiation field is then sufficiently described by the definition that

$$I_\nu \, d\nu d\omega d\sigma$$

is the radiation energy in the frequency interval ν to $\nu+d\nu$ and within the solid angle $d\omega$ which per second passes perpendicularly through the infinitesimal surface element $d\sigma$. The coefficients I_ν and E_ν are characteristic of the radiation field and the emissivity of

matter. They are functions of x, y, z, t and the direction cosines of $d\omega$, which we shall denote by α_x, α_y, α_z. We shall call I_ν the intensity of ν-radiation, and E_ν the emissivity of matter in the same frequency interval.

Consider now the motion of an individual packet of radiation as it is propagated along the direction of the infinitesimal line element ds with the velocity c. After an infinitesimal interval of time dt it has moved a distance $cdt = ds$. On this path it has lost an amount of energy

$$I_\nu \, d\nu d\omega d\sigma \, x_\nu c dt$$

by absorption, and gained by emission an amount

$$E_\nu \, d\nu d\omega d\sigma \, c dt,$$

$d\sigma \, cdt$ being just the volume of the matter which can contribute to the specific energy in the time interval in question. The requirement of conservation of energy leads therefore necessarily to the equation

$$\frac{d}{dt} I)_\nu d\nu d\omega d\sigma) = (E_\nu - x_\nu I_\nu) \, d\nu d\omega d\sigma \, c,$$

which will be called the equation of motion of the radiation. The symbol d/dt refers to a moving light packet. Relative to a coordinate system at rest we must substitute for d/dt the hydrodynamic operator

$$\frac{d}{dt} = \frac{\partial}{\partial t} + c \frac{\partial}{\partial s},$$

in virtue of which the final form of the equation of motion of radiation becomes

$$\frac{1}{c} \frac{\partial I_\nu}{\partial t} + \frac{\partial I_\nu}{\partial s} = E_\nu - x_\nu I_\nu. \tag{67}$$

36. The Flux of Radiation

We define the monochromatic flux of radiation as a vector, the components of which are given by the expressions

$$F_{\nu x} = \int I_\nu \alpha_x \, d\omega; \quad F_{\nu y} = \int I_\nu \alpha_y \, d\omega; \quad F_{\nu z} = \int I_\nu \alpha_z \, d\omega. \tag{68}$$

The integrations must be extended to all directions of the unit sphere. Evidently this vector describes the net flux of radiation in the frequency interval ν to $\nu + d\nu$, which per second passes through unit surface placed perpendicularly to its direction. On using the equation of motion this vector may be brought into a different form, which is of some interest. Write equation (67) in the form

$$I_\nu = \frac{1}{x_\nu} \left\{ E_\nu - \alpha_x \frac{\partial I_\nu}{\partial x} - \alpha_y \frac{\partial I_\nu}{\partial y} - \alpha_z \frac{\partial I_\nu}{\partial z} - \frac{1}{c} \frac{\partial I_\nu}{\partial t} \right\},$$

and introduce this expression of I_ν into (68). We then find

$$x_\nu F_{\nu x} = \int E_\nu \alpha_x \, d\omega - \frac{\partial}{\partial x} \int I_\nu \alpha_x^2 \, d\omega - \frac{\partial}{\partial y} \int I_\nu \alpha_x \alpha_y \, d\omega - \\ - \frac{\partial}{\partial z} \int I_\nu \alpha_x \alpha_z \, d\omega - \frac{1}{c} \frac{\partial F_{\nu x}}{\partial t}. \quad (69)$$

This equation brings out the importance of the *radiation tensor*, the components of which are given by

$$P_{ik} = \int I_\nu \alpha_i \alpha_k \, d\omega,$$

as it is the divergence of this tensor that appears in equation (69). The first term on the right-hand side of (69) will be proportional to $F_{\nu x}$, etc., and may be left out by a suitable re-definition of the absorption coefficient, which we assume to be done. The physical reason for this is that the emission by matter taking place independently of the radiation field will be symmetrical with respect to the flux and contribute nothing to the expression in question. The emission stimulated by the flux will, on the other hand, be proportional to the intensity of the flux and be in its direction of motion, according to Einstein's original theory of stimulated emission. Considering, moreover, stationary conditions, the term involving a derivative with respect to time is to be neglected, and equation (69) reduces to the simple statement that the product of the absorption coefficient and the flux of radiation is precisely equal to the negative divergence of the radiation tensor. Now $x_\nu \mathbf{F}_\nu / c$ is the force exerted by radiation upon unit volume of matter. Equation (69) shows therefore that, with the limitations stated, this force is the divergence of a tensor, which only reduces to a simple isotropic pressure—the light pressure—under special conditions.

37. Digression on the Stellar Interior

The neglect of the time derivative with respect to other terms in equation (69) is found to be justifiable in nearly all cases. For this term to become of importance the radiation field will have to suffer considerable variations in time intervals of the order $1/x_\nu c$, and this will usually be a very small fraction of a second, and hence be negligible, both in the stellar interior as well as in the stellar atmosphere.

Another question of more real importance is to find the conditions necessary for the radiation tensor to reduce to an isotropic pressure. For this purpose we split I_ν in two parts, I'_ν and I''_ν, with the under-

standing that I'_ν is the linear average of I_ν for all directions,

$$4\pi I'_\nu = \int I_\nu \, d\omega,$$

or, what is the same, $4\pi I'_\nu/c$ is the energy density of the radiation field. The radiation tensor will now split up into a corresponding sum $P'_{ik}+P''_{ik}$, which are determined by the equations

$$P'_{ik} = \int I'_\nu \alpha_i \alpha_k \, d\omega = \frac{4\pi}{3} I'_\nu \delta_{ik} \qquad \left(\delta_{ik} = \begin{cases} 1, & i = k \\ 0, & i \neq k \end{cases}\right),$$

$$P''_{ik} = \int I''_\nu \alpha_i \alpha_k \, d\omega.$$

The tensor P'_{ik} corresponds consequently to a hydrostatic pressure, while P''_{ik} corresponds to oblique stresses in a continuous medium. Let us estimate the relative importance of the two tensors for the case of the sun. The flux of radiation is evidently some measure of the asymmetry of the radiation field. It is not probable that this flux will increase very much from the surface towards the interior. At most we may postulate an increase inversely as the square of the distance from the centre through the outer half of the star, while it is likely again to decrease by further approach towards the centre. This will then also be the probable course of the tensor P''_{ik}. The tensor P'_{ik}, on the other hand, will behave quite differently. From the very definition of I'_ν it follows, in fact, that it will increase proportional to Planck's function, or, integrated over all frequencies, proportional to the fourth power of the temperature, according to the law of Stefan and Boltzmann. While the temperature increases from several thousand degrees on the stellar surface to some million degrees in the interior, the radiation tensor will increase several billion times. At the surface of the sun the two tensors are of about the same magnitude, but in the main interior the tensor P'_{ik} will exceed P''_{ik} by a factor of the order 10^{12}. From the equation of motion (67) we see that the average value of I_ν must be very closely given by E_ν/x_ν, which we shall denote by B_ν:

$$B_\nu = \frac{1}{x_\nu} E_\nu,$$

so that

$$\frac{1}{c} P'_{ik} = \frac{4\pi}{3c} B_\nu \delta_{ik},$$

which is the light pressure of elementary theory. The flux multiplied by $\frac{1}{c} x_\nu$ is now just equal to the gradient of this pressure.

The quantity B_ν must, generally speaking, be a function of temperature only. In the case of thermodynamic equilibrium we have, according to the laws of Kirchhoff and Planck,

$$\frac{4\pi B_\nu}{c} = \frac{8\pi h\nu^3}{c^3}(e^{h\nu/kT}-1)^{-1},$$

and the question arises how far this result will be influenced by deviations from thermal equilibrium in the stars. The spatial changes in temperature cannot influence the problem at all, as these variations inside the star are not likely to be comparable with changes in temperature known from laboratory experiments to be necessary in order to influence seriously the laws of thermal equilibria. Moreover, the stellar substance is likely to be rather hazy, ensuring an intimate interaction between matter and radiation, which serves to equalize their temperatures.

Various authors have contended that the irreversible generation of energy, which we believe to take place in the interior of the stars, will perceptibly influence the laws of thermal equilibrium. It is, however, mostly overlooked that the deviations in question must depend intimately upon the special mode of generation of energy, and may even be expected to vanish completely, such as would be the case if the energy were supplied by a slow shrinkage of the star as a whole, as was suggested by Helmholtz. It lies in the nature of things, however, that this irreversible emission will never entail any serious deviations from the laws of thermal equilibrium as regards the stellar interior. This follows from a line of reasoning entirely similar to that just applied in the estimate of the relative importance of P'_{ik} and P''_{ik}. For the flux of radiation from the surface includes everything of irreversible emission in the star, and since the energy carried by this flux is negligibly small in comparison with the energy residing in the radiation field of the stellar interior, the question is to be decided in the negative.

When the absorption coefficient is independent of the frequency we may drop ν as an index. Integrating over the whole spectrum, the total flux of radiation becomes

$$\mathbf{F} = -\frac{4\pi}{3x}\nabla\int_0^\infty B_\nu\, d\nu = -\frac{c}{3x}\nabla u,$$

where $$u = aT^4, \qquad a = \frac{8\pi k^4}{h^3c^3}\int_0^\infty \frac{x^3\, dx}{e^x-1} = \frac{8\pi^5k^4}{15h^3c^3},$$

according to Stefan-Boltzmann's law. If x does depend on the frequency, we may use the same expression for the flux provided we define x by the equation

$$\frac{1}{x} = \frac{\pi}{caT^3} \int_0^\infty \frac{\partial B_\nu}{\partial T} \frac{d\nu}{x_\nu}.$$

When x_ν is averaged in this way it is called the *stellar* absorption coefficient.

38. The Case of an Atmosphere

Consider the problem of integration of the equation of radiative transfer under conditions approximating those to be expected in a stellar atmosphere. To simplify the formulae it is convenient to leave out ν as an index from all quantities depending on ν, as long as there is no direct danger of serious misunderstandings. We shall further always assume the star to be spherically symmetrical, so that the radiation field will have rotational symmetry round the direction to the centre. Denote by r the distance to the centre, and by θ the angle between the direction to the centre and the line of sight. We have then

$$\frac{\partial}{\partial s} = \frac{\partial}{\partial r}\frac{dr}{ds} + \frac{\partial}{\partial \theta}\frac{d\theta}{ds}.$$

By simple geometry—cf. Fig. 3—we find

$$\frac{dr}{ds} = \cos\theta, \qquad \frac{d\theta}{ds} = -\frac{\sin\theta}{r}.$$

The equation of radiative transfer therefore becomes

$$\cos\theta \frac{\partial I}{\partial r} - \frac{\sin\theta}{r}\frac{\partial I}{\partial \theta} = x(B-I). \tag{70}$$

In most problems of stellar atmospheres it is legitimate to assume the atmosphere to be stratified in plane parallel layers. This presupposes that we are considering cases when the line of sight really passes through the star proper, and not merely through the atmosphere, because the curvature of the star will necessarily be of decisive importance very close to the case of grazing incidence, such as is met with in the study of the chromosphere and the corona during solar eclipses. Leaving such cases out of consideration, however, as well as abnormal stars like the O-stars, the B-stars with emission line spectra, and the novae, the assumption of plane-layer atmospheres appears to give a close approximation to the truth. For

this case the angle θ may be treated as a constant in equation (70) and the second term involving the derivative with respect to θ may be neglected altogether.

Instead of r it is convenient to introduce as a new variable the *optical depth* of the atmosphere, defined by the equation

$$d\tau = -\varkappa dr.$$

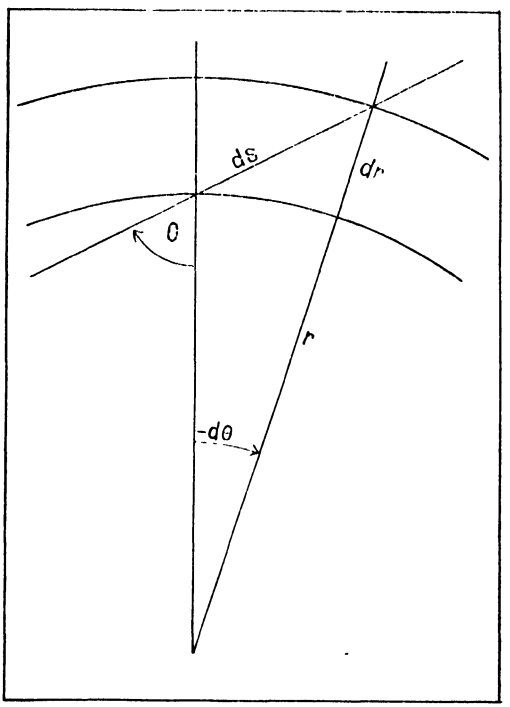


FIG. 3. Diagram showing the spherical coordinates used in the equation of radiative transfer.

Denoting further $\cos\theta$ by μ, the equation of radiative transfer becomes

$$\mu \frac{dI}{d\tau} = I - B. \tag{71}$$

On account of the assumed spherical symmetry the element of solid angle $d\omega$ may be taken to be $2\pi d\mu$, and the resultant flux of radiation becomes, therefore,

$$F = 2\pi \int_{-1}^{1} I\mu \, d\mu,$$

the perpendicular components vanishing because of the symmetry.

We further define the total intensity of radiation by the equation

$$J = 2\pi \int_{-1}^{1} I \, d\mu.$$

Both these expressions hold whether I represents the intensity in a small frequency interval or the intensity integrated over all frequencies.

From the differential equation for I we shall now derive two equations involving F and J which are of considerable importance in the following. First multiply equation (71) as it stands by $2\pi \, d\mu$ and integrate between the limits -1 and $+1$. This gives, on assuming the emission to be independent of the direction,

$$\frac{dF}{d\tau} = J - 4\pi B. \tag{72}$$

Next multiply by $2\pi\mu \, d\mu$ and integrate between the same limits. This gives

$$F = \frac{\partial}{\partial \tau} \int_{-1}^{1} 2\pi I \mu^2 \, d\mu.$$

Since the mean value of μ^2 over the unit sphere is $\frac{1}{3}$, it was suggested by Eddington as a suitable approximation to take

$$2\pi \int_{-1}^{1} I\mu^2 \, d\mu = \tfrac{1}{3} 2\pi \int_{-1}^{1} I \, d\mu = \tfrac{1}{3} J,$$

or

$$F = \frac{1}{3} \frac{\partial J}{\partial \tau}. \tag{73}$$

Introducing this expression of F into equation (72) we find the following ordinary differential equation for the total intensity:

$$\frac{1}{3} \frac{\partial^2 J}{\partial \tau^2} = J - 4\pi B.$$

It is of interest to consider a little more closely what the above approximation really means. It holds exactly when the intensity is uniformly the same in all directions. It also holds exactly when the intensity is uniform, I_1 say, over one hemisphere, and uniform of intensity I_2 over the other hemisphere. It holds also, therefore, for a uniform emission from a half-plane into a vacuum without back radiation. Consider, on the other hand, the case when the intensity

is I_1 over a cone corresponding to $\mu = \mu_0$, and I_2 over the remaining directions. Then

$$\int_{-1}^{1} I\mu^2 \, d\mu = \frac{2\pi}{3}\{I_1(1-\mu_0^3)+I_2(1+\mu_0^3)\},$$

and

$$J = 2\pi\{I_1(1-\mu_0)+I_2(1+\mu_0)\}.$$

Take for instance the rather extreme case $I_1 = 2I_2$ and $\mu_0 = -\frac{1}{2}$. We then find

$$\int_{-1}^{1} I\mu^2 \, d\mu = \tfrac{25}{28}.\tfrac{1}{3}J,$$

so even in this case the error in the approximation (73) is small.

39. The Formation of the Continuous Spectrum

The simplest case surpassing the bounds of the strict equilibrium theory occurs when the state of *matter* corresponds at each point to that of local thermal equilibrium, including its ability of emitting and absorbing radiation. Still, the variation of temperature from one place to another will leave some traces in the field of radiation, but even there deviations must be confined to the surface regions, so that proceeding inwards into the star a state of strict thermal equilibrium is rapidly approached.

As a second simplification we shall consider first the case when the absorption coefficient is independent of frequency. This case will give a first insight into the conditions for the formation of the continuous spectrum in the radiation from a star which seems to approach rather close to the truth. The consideration of selective absorption has no particular interest in this connexion, as it is fairly evident that the formation of stellar absorption lines is intimately connected with a departure from a state of local thermodynamic equilibrium, as we shall see later.

Neglecting further the generation of heat energy by viscosity or otherwise, and the transfer of heat by molecular conduction or by convection, it follows from the principle of conservation of energy that the integral flux of radiation must be a constant. By integral flux we mean the monochromatic flux integrated over all frequencies. Since the absorption coefficient is assumed to be independent of the frequency, equation (72) holds provided the total intensity J and the emission function B are also assumed to be integrated over all frequencies. Since F is a constant, $dF/d\tau$ vanishes, and we have

$$J = 4\pi B. \tag{74}$$

By (73) it follows further on integration that

$$J = 3F\tau + J_0, \tag{75}$$

where J_0 is an integration constant, which must be chosen in such a manner that there is no inward radiation at the boundary of the star, when $\tau = 0$.

The simplest way to determine the approximate value of this constant is to assume the radiation field on each point separated into two parts, each of constant intensity. Taking the outward intensity equal to I_1 and the inward intensity equal to I_2 we have

$$I = I_1 \quad (1 \geqslant \mu \geqslant 0), \qquad I = I_2 \quad (0 \geqslant \mu \geqslant -1).$$

We then find that the following relations hold:

$$J = 2\pi(I_1 + I_2); \qquad F = \pi(I_1 - I_2); \qquad 2\pi \int_{-1}^{1} I\mu^2 \, d\mu = \tfrac{1}{3}J.$$

The last equation shows that the present approximation is consistent with the one introduced earlier, on which our equation (75) is directly based. We see further that the incident radiation vanishes when

$$J_0 = 2F,$$

which is therefore the value of the constant J_0 we have been looking for. Hence, finally,

$$J = F(2 + 3\tau). \tag{76}$$

Since we have assumed from the outset that matter is in local temperature equilibrium, we have

$$\frac{4\pi B}{c} = aT^4.$$

Equation (74) then shows that the integral intensity will at each point correspond to the local temperature of matter. This is no general consequence of the assumption of local temperature equilibrium but solely due to the much more special assumption that the absorption coefficient is independent of frequency.

It is customary to express the flux in terms of the effective temperature T_e by the equation

$$F = \tfrac{1}{4}caT_e^4, \tag{77}$$

corresponding to the idealized case of emission from a sharply bounded boundary surface, kept at a temperature T_e and radiating into empty space without back radiation. The occurrence of the

factor $\frac{1}{4}$ in this relation, which at first sight may appear out of place, has a very simple origin. For caT_e^4 is the total intensity of radiation, independent of direction. Half of it is radiated back to the star, leaving one-half for the outward radiation, and accounting for a factor $\frac{1}{2}$. The other half is due to the integral over μ giving the directed flow of outward radiation.

From (74), (76), and (77) we have now the final relation

$$T^4 = T_e^4(\tfrac{1}{2}+\tfrac{3}{4}\tau).$$

The limiting temperature of the atmosphere thus becomes

$$T_0 = \frac{1}{\sqrt[4]{2}} T_e = 0{\cdot}841 T_e.$$

The optical thickness of the atmosphere is, on the other hand,

$$\tau = \tfrac{2}{3},$$

since this makes $T = T_e$.

As emphasized before, the above results depend essentially on the assumption that the absorption coefficient is independent of frequency. Dropping this assumption, the results may turn out rather different. In order to elucidate the degree of arbitrariness involved in the problem from this cause, we shall consider a few simple cases. It is sufficient to consider directly the exchange of radiation at the top of the atmosphere. Since the emergent radiation is practically speaking uniform over the lower half-plane and otherwise zero, the total amount of radiation absorbed is $2\pi x_\nu B_\nu(T_e)$. We here add ν as an index to emphasize the dependence on ν. The corresponding amount of radiation emitted is, on the other hand, $4\pi x_\nu B_\nu(T_0)$. In the case of radiative equilibrium the total amount of radiation emitted and absorbed must be equal, or

$$2\pi \int_0^\infty x_\nu B_\nu(T_e)\, d\nu = 4\pi \int_0^\infty x_\nu B_\nu(T_0)\, d\nu. \tag{78}$$

Three different cases are interesting.

1. **The atmospheric absorption is confined to the infra-red part of the spectrum,** so that the emissivity is given by Rayleigh's law:

$$B_\nu \propto \nu^2 T.$$

The integrals over ν will then balance on both sides and there remains

$$T_0 = \tfrac{1}{2} T_e.$$

The limiting temperature is then just one-half the effective tempera-

ture of the star. This is to be regarded as a lower limit of the temperature of a plane atmosphere in radiative equilibrium.

2. **The absorption is confined to the far ultra-violet part of the spectrum.** As a typical case of this sort we may take

$$x_\nu = 0 \quad (\nu < \nu_m), \qquad x_\nu = \text{const.}\, \nu^{-3} \quad (\nu > \nu_m).$$

The choice of the negative third power for the frequency in the above relation was first suggested by X-ray absorption experiments, and was later also justified by quantum-theory calculations. Since the absorption is to take place in the ultra-violet, we are justified in taking $h\nu_m/kT_e$ large in comparison with unity. It is then permissible to use the Wien approximation for the intensity of radiation, and equation (78) assumes the following form:

$$T_e e^{-h\nu_m/kT_e} = 2T_0 e^{-h\nu_m/kT_0}.$$

This relation may also be written in the form

$$\frac{T_e - T_0}{T_0} = \frac{kT_e}{h\nu_m} \log\left(\frac{2T_0}{T_e}\right),$$

which shows that T_0 will approach indefinitely near to T_e as $h\nu_m/kT_e$ approaches infinity.

3. **The absorption coefficient is independent of frequency.** We then have the case treated in detail before, and the condition (78) reduces to

$$T_e = \sqrt[4]{2}\, T_0$$

in complete agreement with the earlier result. This serves as a sort of control of the method by showing that its degree of approximation is of the same order as that of the more complete calculation.

Knowing nothing definite about the absorption coefficient, we can only assert that the atmospheric temperature will assume a value somewhere between the effective temperature of the star as an upper limit, and half this value as a lower limit.

Having calculated the emergent radiation to a first approximation we may proceed a step farther and find the distribution of radiation on different directions and frequencies. For this purpose it is sufficient to introduce the expression for the emissivity B which results from (76), viz.

$$B = \frac{1}{4\pi} F(2+3\tau),$$

into the equation of radiative transfer (71), and integrate. Multiplying

this equation by $e^{-\tau/\mu}$ it may be written

$$\mu \frac{d}{d\tau}(Ie^{-\tau/\mu}) = -Be^{-\tau/\mu},$$

or integrating from the observer, that is $\tau = 0$, to an optical depth τ_0, we find

$$I = I_0 e^{-\tau_0/\mu} + \int_0^{\tau_0} Be^{-\tau/\mu}\, d\tau/\mu,$$

where the intensity at optical depth τ_0 has been denoted by I_0. If the intensity does not increase exponentially with τ it is clear that the first term on the right-hand side will soon become negligible in comparison with the second term. This is all the more certain if we suppose the integration carried right through the star, so that I_0 will denote the radiation incident from without. Supposing no incident radiation to be present and the integration to be extended to an infinite distance from the observer, we may therefore without hesitation write

$$I = \int_0^{\infty} Be^{-\tau/\mu}\, d\tau/\mu. \tag{79}$$

We shall now introduce into (79) the first-order expression derived above for B as a function of τ, and thus obtain the corresponding expression of the emergent intensity as a function of direction. Since B is a linear function of τ only, it is clear that the contributions to the integral will come from quite superficial layers only. We find

$$I = \frac{1}{4\pi} F(2+3\mu). \tag{80}$$

This formula refers to the intensity of the integrated spectrum. If we want to find the distribution of radiation in frequency in the emergent intensity, we must use the Planck expression for B as a function of frequency,

$$B = \frac{2h\nu^3}{c^2}(e^{h\nu/kT}-1)^{-1},$$

and express the temperature as a function of optical depth by the formula

$$T = T_e(\tfrac{1}{2}+\tfrac{3}{4}\tau)^{\frac{1}{4}},$$

or

$$I = \frac{2h\nu^3}{c^2} \int_0^{\infty} \frac{e^{-\tau/\mu}\, d\tau/\mu}{e^{\frac{h\nu}{kT_e}(\frac{1}{2}+\frac{3}{4}\tau)^{-\frac{1}{4}}}-1}.$$

This formula, which was first given by Lindblad and Milne, deviates

only slightly from the Planck formula—so little, in fact, that the difference is of no particular importance for astrophysical applications.

As regards the observational side of the problem it is a fact that the continuous background of the spectrum of the sun and most of the stars corresponds roughly to the energy curve of a black body at a certain temperature. This shows at once that the assumption of a state of local temperature equilibrium for the stars cannot be so far from the truth, and also that very great variations with frequency in the coefficient of general absorption are unlikely.

However, certain definite features of the continuous spectra of the stars, and especially the sun, seem to be established beyond doubt. Although their theoretical interpretation is far from clear yet it is still of interest to sum them up briefly on this occasion.

The solar spectrum looks on the whole like a black-body spectrum at a temperature of about 6,000°. The maximum intensity in the spectrum when plotted as a function of wave-length occurs at about 4,700 A., which, according to Wien's law,

$$\lambda_m T = 0{\cdot}289 \text{ deg. cm.},$$

corresponds to a temperature of 6,150° K. Determining, on the other hand, the total energy emitted per square cm. of the solar surface per second, and using Stefan-Boltzmann's law as a means of obtaining the temperature, one finds a value of only 5,750° K. The difference between these results is probably connected with the fact that the solar energy curve rises more steeply to a maximum than an ordinary Planck curve. The infra-red part of the solar curve thus corresponds to a temperature of only 5,600°, while the ultra-violet part corresponds, according to Fabry, to a temperature of 5,900°. Very careful measurements of H. H. Plaskett lead to a somewhat higher value of the temperature in the ultra-violet. On comparing these different investigations on solar temperature it must always be remembered that the subject is one of the most difficult in astrophysics, both observationally and theoretically, and one should be very careful in drawing theoretical conclusions from the results obtained up to this time.

In the case of the sun there is an interesting check on the theory in the variation of the intensity from the centre of the disk to the limb, which is the same thing as the variation of the intensity with different angles θ. It is very interesting that for the integrated radia-

tion of all wave-lengths the simple cosine law

$$I(\mu) = I(1)(1-u+u\mu)$$

gives a good approximation to the observations. Here u is a constant called the coefficient of darkening. The value found for u observationally is about 0·56. The theoretical law derived earlier (80) corresponds to a coefficient of darkening of 0·60, so the agreement is fairly good, at least as good as might be expected.

This agreement between theory and observation with regard to darkening of the solar radiation towards the limb indicates that the temperature distribution in the solar atmosphere does really correspond to that demanded by the simple theory of radiative equilibrium, giving a linear relation between T^4 and the optical depth. The various attempts made in order to push theory one step farther, and also to explain the deviations from a Planck curve as well as the minor differences between theory and observations with regard to the darkening towards the limb, have thus far not led to concordant results.

The continuous background of stellar spectra is much less investigated than for the sun. It should be mentioned, however, that a careful analysis of the light curve of eclipsing variables has revealed the fact that these stars too are strongly darkened towards the limb. A cosine relation of the kind found to hold for the sun works well also for these stars, but the relation cannot be said to have been put to a real test in this case.

Although it is scarcely possible to stress the observational evidence very far, it seems established beyond doubt that for a limited number of stars, whose general energy distribution in the continuous spectrum indicates a relatively low temperature, the spectral type is such as is usually associated with a very much higher temperature. To a minor degree this is a characteristic difference between giant and dwarf stars, a dwarf demanding a higher colour temperature in order to show the same general spectral features as a giant. It is therefore natural to conclude that abnormally red or yellow stars—abnormal because their spectral lines would indicate a higher temperature—have the giant characteristics in an excessive degree. Some giant characteristics at least must be accentuated far beyond the normal scope.

It is then interesting to inquire if this characteristic may be just

an abnormal extension of the stellar atmosphere; or, slightly differently put, if a large extent of the atmosphere will tend to produce a marked difference between the colour temperature and the temperature determining the spectral class of the star. This is a question which cannot properly be answered until the theory of the spectral classes has been developed; and we shall therefore postpone this problem for later consideration.

40. Formation of Absorption Lines. General Considerations

The theory of absorption lines cannot be based simply on the assumption of local thermodynamic equilibrium to the same extent as the theory of the continuous spectrum. It is very important to realize this before entering on a more detailed analysis. If, in fact, this were the case, the intensity of light, in the region of a line, would be given by a slight modification of formula (79) for the continuous spectrum. This modification consists in the proper distinction between the optical depth in the line and in the adjacent continuous spectrum. We make this distinction by denoting the optical depth in the line by τ' and in the continuous background by τ. The formula for the emergent intensity in the line is then by (79)

$$I' = \int_0^\infty B(\tau) e^{-\tau'/\mu} \, d\tau'/\mu.$$

This means essentially that the intensity I' will equal approximately the temperature emission function B at an optical depth $\tau' = \mu$. If the line absorption coefficient is much larger than the continuous absorption coefficient, the optical depth $\tau' = \mu$ will lie at a level far above the photosphere, and hence correspond to a lower temperature and to a correspondingly smaller intensity. The central line-intensity in the middle of the solar disk should thus correspond very nearly to the limiting temperature $T_0 = T'_e/\sqrt[4]{2} = 0{\cdot}841 T'_e$. As a rule, however, the observed residual intensities are much smaller than those corresponding to this limit. Still more striking in this connexion is the behaviour of the absorption lines at the solar limb. At the limb the whole radiation, continuous or in the lines, corresponds to the temperature T'_0, which means that the spectral lines should vanish. This, however, is far from being the case, the absorption lines at the limb being nearly as prominent as near the centre of the solar disk. It is therefore to be considered a settled matter that the theory of

absorption lines must embody features essentially different from those displayed in a state of local thermodynamic equilibrium.

The reason why this must be so is interesting, as it is intimately connected with the theory of atomic structure. While there is no doubt that the far interior of a star is in local thermodynamic equilibrium it is because of the large density and high temperature which are prevalent there. When an atom absorbs a quantum of radiation in the interior the atom is likely to suffer a collision with some free particle before having the chance of re-emitting the absorbed radiation. The distribution of the atoms over the various quantum states will therefore be determined in the random, nondescript manner characteristic of collisional processes, which usually lead to a state approaching local thermal equilibrium. Proceeding sufficiently far out into the atmosphere, however, a state will be reached when collisions are so rare that an absorbed quantum is likely to be re-emitted *by the same atom* before an effective collision takes place. The radiation will then be re-emitted either nearly in the same frequency, or in the form of a limited number of spectral lines, depending upon the state of the atom under consideration. In this case, which may be termed a state of pure radiative equilibrium, there are much fewer possibilities open for the redistribution of the absorbed radiation over different frequencies, and the consequent state of the radiation field will differ in several important respects from that corresponding to local temperature equilibrium, as we shall see presently. While the state of temperature equilibrium tends to hide the individual atomic processes in play, so to speak, behind a veil, the state of radiative equilibrium tends to draw aside this veil, so that the individual processes may reveal themselves in the observed phenomena. It is solely by this feature of radiative equilibrium that we have a chance at all to make that incisive analysis of stellar atmospheres which is actually made in astrophysics.

In this field it is necessary to be rather patient, as we are really concerned with an enormously complicated problem, and any simple solution must be eyed with suspicion. All we can hope to do at present is to make the situation clear in a general way by interpreting the fact that the solar and stellar lines are more marked than the theory of local temperature equilibrium would lead us to expect. At the same time it will be a crucial point to find out how the lines can remain unaffected to such a large degree on approaching the limb of the sun.

The key to the understanding of the phenomena in question is provided by the proper consideration of possible *cycles* of transitions which an atom may perform under the influence of a radiation field. Two sorts of cycles will come into consideration, the one being a successive set of transitions between quantum states in the ordinary sense, the other transitions between the continuous set of states which, in the theory of Weisskopf and Wigner, constitutes each of the ordinary states, and which must be considered closely in any theory of the line profiles. Both sorts of cycles will play a predominant part in the theory of stellar lines, as we shall see in the following.

41. Coherent Scattering

The simplest case imaginable is met with when the atom contains only two states, and, moreover, the lowest state is considered as *infinitely sharp*. When an assembly of such atoms are exposed to the influence of a monochromatic field of radiation of frequency ν, the scattered radiation must necessarily also have the frequency ν, provided the Doppler effect due to the recoil velocity acquired by the atom during the scattering process is neglected. We shall call an absorption process with subsequent emission according to this scheme *coherent scattering*, because its coherence in frequency is the physical basis of experiments involving dispersion of light in the ordinary sense. In such experiments the dispersive substance is to a very high degree in the normal quantum state, which is practically speaking infinitely sharp. It is broadened only by transitions to higher states, and under ordinary experimental conditions these take place so infrequently that the resultant broadening is negligible. If the lowest energy state has an appreciable width, as will usually be the case when considering subordinate lines, the scattered light will be redistributed in frequency over the spectral line in question according to a certain law, which may be stated quantitatively. There will then be reason to suspect that dispersion in the ordinary sense will be weak, and that the problem must be handled in a different way from that of coherent scattering. Provisionally we shall only be concerned with this latter case.

Let as usual I denote the intensity of the radiation field, and let σ denote the scattering coefficient per unit volume of the atoms for frequency ν, which is supposed to be well within the region of the

spectral line in question. In the case of coherent scattering the intensity of the re-emitted light will depend on the angle between the incident and the scattered beam. Denoting this angle by θ we must consider σ as a function of ν and θ, besides the optical depth in the atmosphere, measured in the continuous region, which we denote by τ as usual. The dependence of σ on θ is given simply by the factor $1+\cos^2\theta$, an expression which was first derived by Rayleigh. The average value of $\cos^2\theta$ is $\frac{1}{3}$, so the fluctuation over different directions is not large. For this reason it is usually neglected in astrophysical theories, although it must be retained in other cases, particularly in the theory of the scattering of the sunlight in our atmosphere, which is a spectacular instance of coherent scattering.

We may thus write down directly the emission of radiation per unit volume and unit solid angle as

$$E = \frac{\sigma J}{4\pi},$$

where J as usual is the total intensity of the field, and the equation of radiative transfer becomes (z being the height above a certain level):

$$\mu \frac{dI}{dz} = \frac{\sigma J}{4\pi} - \sigma I.$$

Since no radiation can enter or leave the frequency interval it once occupies, granted it is only transformed by scattering processes, the flux of radiation through a scattering atmosphere will remain constant, not only for all wave-lengths but also for each infinitesimal frequency interval. The equation derived in the case of pure absorption with a constant absorption coefficient for the integral flux and intensity may therefore be taken over unaltered for the monochromatic flux and intensity in the present case. By exactly the same system of approximations we further find the total monochromatic intensity to be related to the monochromatic flux by the equation

$$J = F(2+3\tau).$$

It must be remembered that the quantities are calculated for a definite frequency of radiation, and that the optical depth τ is defined as

$$\tau = -\int \sigma \, dz,$$

where σ may depend on the frequency in any manner.

Assume that a purely scattering atmosphere is superposed on a

photosphere emitting radiation corresponding to a definite temperature. Let τ signify the total thickness of the atmosphere for the given frequency. J/c will be the energy density of radiation at the photospheric level, and be given by Planck's formula, corresponding to the effective temperature of the star. If the atmosphere had been absent, the relation between the flux and the energy density would have been $F = \frac{1}{4}J$. This formula, however, cannot be directly compared with the one given above for the relation between F and J, because we have now included in J an amount $\frac{1}{2}J$ of inward radiation which is not supposed to be present in the other case. Leaving this out we are left with a flux $\frac{1}{2}J$ when there is no scattering atmosphere overlying the photosphere, and $J/(2+3\tau)$ when there is one of optical thickness τ. The flux is thus cut down by a factor $(1+\frac{3}{2}\tau)^{-1}$ on account of the scattering.

When the scattering coefficient is independent of the frequency the effect of a scattering atmosphere is simply a general decrease in the flux. However, in a practical case the superposition of such an atmosphere on a star would not be likely to influence the resultant flux in the slightest degree, because the radiation held back by the atmosphere would heat the star, and would go on doing so until J were increased by the factor $1+\frac{3}{2}\tau$ so as to precisely cancel the effect of the scattering in question. The case under consideration has therefore only an interest in the case of selective scattering, or a uniform scattering with selective transmission bands.

It is further of interest to find the dependence of intensity on direction. We then proceed in the same way as before when deriving the dependence of the emergent radiation on direction. The only difference is that we now integrate through the scattering atmosphere only. The expression of the intensity may then be written down at once as follows:

$$I(0,\mu) = I(\tau,\mu)e^{-\tau/\mu} + \int_0^\tau \frac{F}{4\pi}(2+3\tau)e^{-\tau/\mu}\,d\tau/\mu,$$

or, carrying out the integration,

$$I(0,\mu) = I(\tau,\mu)e^{-\tau/\mu} + \frac{F}{4\pi}\left\{2(1-e^{-\tau/\mu}) + 3\mu\left[1-\left(\frac{\tau}{\mu}+1\right)e^{-\tau/\mu}\right]\right\}.$$

This equation may be further simplified by observing that we have, approximately,

$$I(\tau,\mu) = \frac{1}{4\pi}J = \frac{1}{4\pi}F(2+3\tau), \tag{81}$$

in consequence of which we find

$$I(0,\mu) = \frac{F}{4\pi}(2+3\mu-3\mu e^{-\tau/\mu}).$$

If the optical thickness of the atmosphere is large in comparison with unity, the last term in this expression may be neglected, and there remains only

$$I = \frac{F}{4\pi}(2+3\mu).$$

This result is rather interesting, as it shows that spectral lines produced by pure scattering will be darkened towards the limb in about the same manner as is observed for the absorption lines in the solar spectrum. The argument raised earlier against the suggestion that Fraunhofer lines arise as a result of absorption and emission in a state of local thermal equilibrium would thus not be effective against the suggestion that they arise by selective scattering.

Although it is clear from general considerations that the conditions of pure scattering cannot apply even approximately under stellar conditions, it is interesting to have this fact confirmed by experiment. This is done in the present case by the fact that the above theory would predict spectral lines with nearly zero intensity in the centre, which must be contrasted with the fact that residual intensities of 10 per cent. or more of the intensity in the adjacent continuous spectrum are common in the sun. As was shown earlier, the scattering coefficient given by the ordinary dispersion theory is of the form const.$(\nu-\nu_0)^{-2}$, ν_0 being the frequency of the centre of the line. For $\nu = \nu_0$ the optical depth τ would then be infinitely large, so that for given photospheric intensity J, equation (81) would require a zero flux F. Adding a small damping term to $(\nu-\nu_0)^2$ would formally remove the infinity, but practically the conclusion would be the same.

This conclusion would further be only slightly modified by taking into account the Doppler effect produced by the thermal motion of the atoms. The Doppler effect will tend to increase the central intensity to a quite marked degree, but it is very far from being able to explain the large residual intensities observed in the sun and the stars. The obvious conclusion is therefore that pure scattering is, as far as stellar conditions are concerned, a fictitious case, which at least must be considerably modified before it can come seriously into consideration even as a first approximation to facts.

42. Absorption and Scattering Combined

An obvious suggestion for an improved approximation is to superpose on the scattering some thermal emission, both of a selective nature, operating in the region of the spectral lines, and of a non-selective nature, responsible for the continuous background.† We are then led to write the selective emissivity of the substance in the form

$$E = \sigma \frac{J}{4\pi} + \kappa B,$$

where σ and κ are certain coefficients, which we shall call coefficients of selective scattering and absorption respectively. It is clear that σ and κ must be proportional to the coefficient of scattering given by quantum theory as regards the dependence on frequency, but both may be affected by coefficients depending on the optical depth. This side of the question cannot well be carried into greater detail since the whole present formulation of the problem is a little hazy.

On forming the equation of radiative transfer for the present case we must add the absorption and emission of a non-selective nature, which is responsible for the formation of the continuous spectrum. Let κ_0 be the coefficient of continuous absorption. The absorbed radiation is then $\kappa_0 I$ and the emissivity is $\kappa_0 B$. The consequent form of the equation of radiative transfer is

$$\mu \frac{dI}{dz} = -(\sigma+\kappa+\kappa_0)I + \sigma \frac{J}{4\pi} + (\kappa+\kappa_0)B.$$

Let τ denote the total optical depth of the atmosphere from the top down to the point in question:

$$\tau = -\int (\sigma+\kappa+\kappa_0)\, dz,$$

and write

$$1-\lambda = \frac{\sigma}{\sigma+\kappa+\kappa_0}.$$

The equation then assumes the new form

$$\mu \frac{dI}{d\tau} = I - (1-\lambda)\frac{J}{4\pi} - \lambda B. \tag{82}$$

In order to integrate this equation in successive approximations we proceed in the same way as before, and derive the exact expressions

$$\frac{dF}{d\tau} = \lambda J - 4\pi\lambda B; \qquad F = \frac{d}{d\tau}\int I\mu^2 2\pi\, d\mu. \tag{83}$$

† Cf. E. A. Milne, *Monthly Not.* **88** (1928), 493. Sir Arthur Eddington, ibid. **89** (1929), 620.

In the last equation we substitute for μ^2 its mean value $\frac{1}{3}$, which makes

$$F = \frac{1}{3}\frac{dJ}{d\tau}.$$

Introducing this value into the first equation (83) we find

$$\frac{d^2J}{d\tau^2} = 3\lambda J - 12\pi\lambda B. \tag{84}$$

Consider first the case in which λ is constant and B a linear function of τ. By this hypothesis we obtain some insight into the influence of a temperature gradient on the emergent radiation. In this case the above equation may evidently be written in the form

$$\frac{d^2}{d\tau^2}(J - 4\pi B) = 3\lambda(J - 4\pi B),$$

the general solution of which is

$$J = 4\pi B + \alpha e^{\sqrt{(3\lambda)}\tau} + \beta e^{-\sqrt{(3\lambda)}\tau}.$$

The arbitrary constants α and β are fixed by the conditions to be satisfied by the solution at the top of the atmosphere and in the far interior. Since conditions must approach rapidly towards temperature equilibrium in the interior, J must approach rapidly towards $4\pi B$ as τ becomes very large. A glance at the above solution then shows that the constant α must be put equal to zero.

For the emergent radiation we have the usual condition:

$$J = 2F = \frac{2}{3}\frac{dJ}{d\tau}$$

as τ approaches zero. Writing B in the form

$$B = a + b\tau$$

we find as an expression of the above condition

$$4\pi a + \beta = -\tfrac{2}{3}\sqrt{(3\lambda)}\beta + \frac{8\pi}{3}b,$$

or solved with respect to β

$$\beta = \frac{\frac{8}{3}\pi b - 4\pi a}{1 + \frac{2}{3}\sqrt{(3\lambda)}}.$$

This approximate solution may now be used for the calculation of I to a second approximation. Just as before we introduce the first-order expression of J into equation (82) for the radiation intensity and perform the integration involved. This gives an expression

which for large values of the upper limit of integration converges towards the value

$$I(0,\mu) = a+b\mu+\frac{(1-\lambda)(\frac{2}{3}b-a)}{\{1+\frac{2}{3}\sqrt{(3\lambda)}\}\{1+\mu\sqrt{(3\lambda)}\}}. \tag{85}$$

For abnormally small values of λ, that is, for dark lines, this expression reduces approximately to

$$I(0,\mu) = \tfrac{2}{3}\sqrt{(3\lambda)}a(1+\tfrac{3}{2}\mu),$$

where the term in b has been neglected because it is very small in comparison with a in this case. We notice again that the dependence of the intensity of the emergent radiation on μ is such that the intensity near the centre of the line decreases in the same manner as the adjacent continuous spectrum on approaching the limb. For stars the dependence on μ is usually not observable. What then becomes of importance is the total flux F, which in the present case by (85) is found to be

$$F = 2\pi\left(\tfrac{1}{2}a+\tfrac{1}{3}b+\frac{(1-\lambda)(\frac{2}{3}b-a)}{1+\frac{2}{3}\sqrt{(3\lambda)}}\,\frac{\sqrt{(3\lambda)}-\log\{1+\sqrt{(3\lambda)}\}}{3\lambda}\right).$$

43. More General Solutions

It is very simple to find the solution corresponding to constant λ, while the dependence of B on optical thickness is left arbitrary. We shall write t instead of $\sqrt{(3\lambda)}\tau$ and put

$$J = e^t\int u\,dt,$$

where u is a new function of t. Introducing this expression of J in equation (84) we find the following equation for u:

$$\frac{du}{dt}+2u+4\pi Be^{-t} = 0,$$

which gives, when multiplied by e^{2t} and integrated,

$$e^{2t}u = u(0)-4\pi\int_0^t Be^x\,dx.$$

The corresponding solution of (84) is consequently

$$J = \alpha e^t+\beta e^{-t}-4\pi e^t\int_0^t e^{-2z}\,dz\int_0^z B(x)e^x\,dx,$$

where α and β are arbitrary constants. In order to find the second approximation to I it is only necessary to introduce the above expression of J into (82) and perform the integration.

It is further of importance to know some solutions corresponding to a variation of λ with height in the atmosphere, as will frequently be the case in nature, due to variations in the state of ionization of the stellar gases with height. It is of some interest in this connexion that it is possible to find a solution in which λ varies exponentially with the optical height, which was first shown by Eddington.† A solution of this type is not strictly applicable to the case where σ/κ is independent of frequency. But it is possible that such a solution may nevertheless be used in an approximate treatment of the problem.

Let us thus consider the case where

$$3\lambda = e^{\alpha\tau+\beta},$$

where α and β are independent of τ. Equation (84) then assumes the form

$$\frac{d^2J}{d\tau^2} = e^{\alpha\tau+\beta}(J-4\pi B).$$

It is further advantageous to use

$$y = \frac{1}{\alpha^2}e^{\alpha\tau+\beta}$$

as an independent variable instead of τ. The resulting form of our equation is then

$$\frac{d}{dy}\left(y\frac{dJ}{dy}\right) = J-4\pi B.$$

From the general theory of linear differential equations of the second order it is known that the solution of an inhomogeneous equation can always be found provided the corresponding homogeneous equation has been solved. This latter is, in the present case,

$$\frac{d}{dy}\left(y\frac{dJ}{dy}\right)-J = 0,$$

which is transformed into an equation of a Bessel function of zero order by the transformation

$$y = -\left(\frac{x}{2}\right)^2.$$

In fact, we find in the new variable

$$\frac{d}{dx}\left(x\frac{dJ}{dx}\right)+xJ = 0.$$

The solution of this equation, which remains finite for $x = 0$, is,

† *Monthly Not.* **89** (1929), 620.

apart from a constant factor, given by the convergent power series

$$J = 1-\left(\frac{x}{2}\right)^2+\left(\frac{x^2}{2.4}\right)^2-\left(\frac{x^3}{2.4.6}\right)^2+\cdots,$$

which is a Bessel function of zero order. Expressed with y as an argument the series is

$$J = 1+y+y^2/2^2+y^3/2^2.3^2+\cdots.$$

When B is a linear function of y the particular solution in question is

$$J_0 = 4\pi\left(B+\frac{dB}{dy}\right),$$

which, when added to the proper Bessel function, gives the complete solution of the problem.

44. An Integral Equation for J

The problems of radiative transfer are also susceptible of solution in terms of integral equations, as was first shown by W. H. Jackson,† and later developed by L. V. King,‡ Schwarzschild,‖ and Hopf.††

In order to show the possibilities of this method we shall consider the problem of mixed scattering and absorption from this point of view. We then start from the differential equation

$$\mu\frac{dI}{d\tau} = I-(1-\lambda)\frac{J}{4\pi}-\lambda B,$$

and introduce the new dependent variable

$$\phi = (1-\lambda)\frac{J}{4\pi}+\lambda B,$$

which will be the function figuring in the integral equation we are going to construct. It is important to notice that ϕ depends on τ only, and not on μ. The differential equation now assumes the form

$$\mu\frac{dI}{d\tau} = I-\phi.$$

Assuming that there is no radiation incident upon the star from outside, the solution of this equation assumes the form

$$I(z) = \int_z^\infty \phi\, e^{-(\tau-z)/\mu}\frac{d\tau}{\mu}.$$

† *Bull. Amer. Soc. Math.*, June 1910, 473.
‡ *Phil. Trans. Roy. Soc.* A, **212** (1912), 375.
‖ *Sitz. d. preussischen Akad. d. Wiss.* **2** (1914), 1183.
†† *Cambridge Tracts*, No. **31**.

The integration is supposed to be carried backward from the point $\tau = z$ to infinity. If it is a case of *emergent* radiation, so that μ is positive, the integral will extend right through the star, so that the limits of the integral, expressed in terms of the optical thickness, will be z and ∞. We shall denote the intensity of the emergent radiation by a dash, that is,

$$I'(z) = \int_z^\infty \phi\, e^{-(\tau-z)/\mu} \frac{d\tau}{\mu} \qquad (0 < \mu < 1).$$

For radiation directed inwards into the star the integration may only involve a moderate optical thickness, that is, assuming as we want to do that we are in the neighbourhood of the stellar surface. Assuming the zero-point of the optical depth to be the top layer of the atmosphere, as we have done earlier, the integration will only cover a range from 0 to z, so that the intensity is given by the expression

$$I''(z) = -\int_0^z \phi\, e^{-(\tau-z)/\mu} \frac{d\tau}{\mu} \qquad (-1 < \mu < 0).$$

From these expressions we now form the total intensity J. We have obviously

$$J = 2\pi\left(\int_0^1 I'\, d\mu + \int_{-1}^0 I''\, d\mu\right),$$

or using the above expressions of I' and I''

$$J = 2\pi\left\{\int_z^\infty \phi\, d\tau \int_0^1 e^{-(\tau-z)/\mu} \frac{d\mu}{\mu} - \int_0^z \phi\, d\tau \int_{-1}^0 e^{-(\tau-z)/\mu} \frac{d\mu}{\mu}\right\}.$$

This expression for J may be brought into a more compact form by introducing the notation

$$\mathrm{K}(x) = \int_x^\infty e^{-u} \frac{du}{u}.$$

For if we use

$$x = \frac{\tau - z}{\mu}$$

as an integration variable in the above integrals, we have

$$\int_0^1 e^{-(\tau-z)/\mu} \frac{d\mu}{\mu} = \int_{\tau-z}^\infty e^{-x} \frac{dx}{x} = \mathrm{K}(\tau - z),$$

and $$\int_1^0 e^{-(\tau-z)/\mu}\frac{d\mu}{\mu} = -\int_{z-\tau}^{\infty} e^{-x}\frac{dx}{x} = -\mathrm{K}(z-\tau).$$

Hence, finally,

$$J = 2\pi\left\{\int_z^{\infty}\phi\mathrm{K}(\tau-z)\,d\tau + \int_0^z \phi\mathrm{K}(z-\tau)\,d\tau\right\} = 2\pi\int_0^{\infty}\phi\mathrm{K}(|\tau-z|)\,d\tau.$$

This expression for J may now be introduced into the equation defining ϕ, and we then obtain

$$\phi(z) = \tfrac{1}{2}(1-\lambda)\int_0^{\infty}\phi(\tau)\mathrm{K}(|\tau-z|)\,d\tau + \lambda B,$$

which is an ordinary linear, inhomogeneous integral equation for $\phi(z)$.

We shall not pursue this line of investigation any further, since we have already derived the most relevant results in other ways. Furthermore, the method of integral equations as far as it is of interest in radiation problems has been amply represented by L. Hopf in the book referred to before.

45. Solutions by Numerical Methods

It will be clear from the preceding discussion that analytical solutions of the equation of transfer will become very cumbersome when passing on to more complex cases, where the various coefficients depend more intricately on the optical depth. In order to obtain some information about conditions even in such cases it is necessary to have recourse to numerical solutions of the differential equations.

An attempt in this direction was made by Pannekoek.† Expressing Pannekoek's formulae in the language of the formalism used here, it may be said that he finds solutions of the equation of transfer when λ is a function of optical depth, given by

$$\lambda = \frac{Ax}{Bx+1},$$

where A and B are constants and where x is the optical depth measured by the combined action of selective absorption and scattering. The coefficient of general absorption is assumed proportional to the square of total pressure, as a consequence of which there will exist a quadratic relation between x and our τ, and the

† *Monthly Not.* **91** (1930), 139. A modification of Pannekoek's solution was given by Minnaert, *Zeits. f. Astrophys.* **10** (1935), 40.

emissivity E will be a quadratic function of x without a linear term. Pannekoek's equations differ, however, from those used above, in assuming the selective absorption to be due to collisions, while we have assumed it to be due to cyclical transitions. This limits the applicability of this solution seriously, since in actual atmospheres the effect of collisions is necessarily very small. For this reason we shall refrain from entering further on these calculations.

X

PROFILES OF ABSORPTION LINES

46. Observations of Line Profiles and their Interpretation

THE first measurements of line profiles in the solar spectrum appear to be those made by Schwarzschild,† which were concerned with the *H* and *K* lines of ionized calcium. These lines are fully twenty

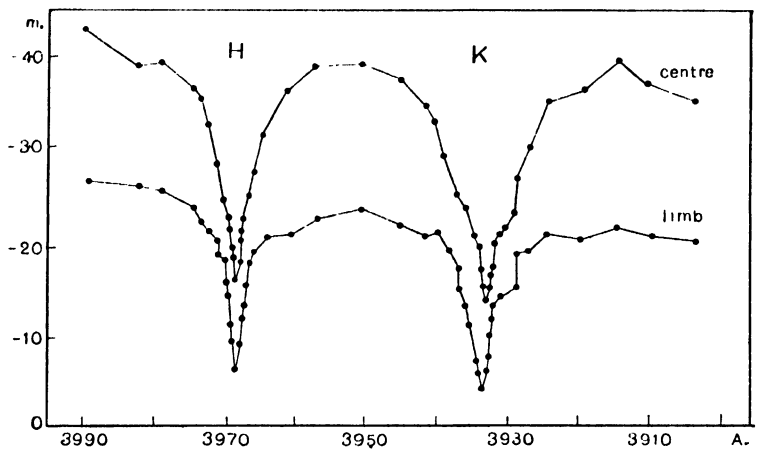


FIG. 4. The calcium *H* and *K* lines at the centre and the limb of the sun according to Schwarzschild (*Sitz. d. preussischen Akad. d. Wiss.* **2** (1914)). Ordinates are $m = -0{\cdot}4$ log intensity.

angstroms broad, so there is no trouble with the intrinsic width, although it is very difficult to obtain a correct estimate of the residual intensity in the centre of the lines. It appears, however, that later work has not led to serious divergences from the work of Schwarzschild even in this respect. Schwarzschild was a pioneer in the field of the theory of radiative equilibrium of the solar atmosphere, and his measurements were intended as a test of his theoretical work, particularly as regards the darkening in the lines on approaching the solar limb.

The work on line profiles was next taken up again in the late twenties, first by v. Klüber and then by Unsöld, Minnaert, H. H. Plaskett, and others. By now there is available much material on line profiles both in the spectrum of the sun and the stars.

† *Sitz. d. preussischen Akad. d. Wiss.* **2** (1914), 1183.

The amount of information hidden in observational material on line profiles is very large, and only part of it is as yet unearthed. We shall in the following try to show step by step what kind of information is to be expected, and the degree of reliability which may be placed on it.

First of all it is important to find out the nature of the processes responsible for the most common line profiles. We have already quite formally classified these processes as scattering processes on one hand and absorption processes on the other, and have developed a theory in which the ratio between the corresponding coefficients is

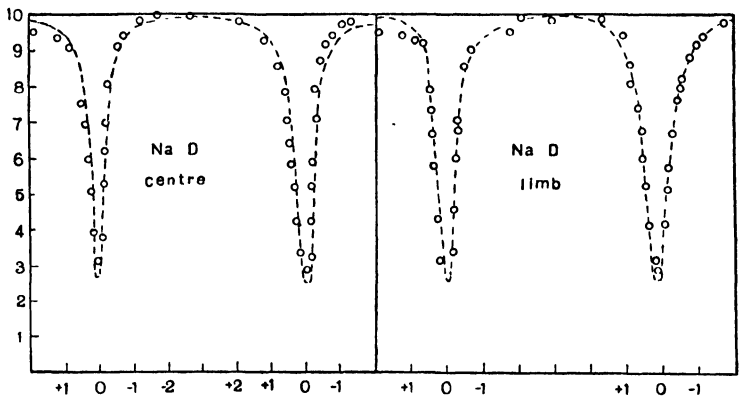


FIG. 5. Profiles of sodium D lines in the sun (Unsöld).

either independent of optical depth or a simple function of this depth. We shall now see that this theory goes far to render an account of the observed profiles of absorption lines in the sun and the stars.

Let us therefore again consider the theory of a line profile when the ratio of the selective absorption coefficient κ to the coefficient of scattering σ is independent of optical depth, and write

$$\kappa = \eta\sigma.$$

It is to be expected that η is independent of frequency too, as we have been using the terms 'absorption' and 'scattering' for intrinsically the same process.

The quantity λ giving the ratio of absorption to extinction now assumes the form

$$\lambda = \frac{\kappa_0+\eta\sigma}{\kappa_0+(\eta+1)\sigma}.$$

The new quantity η measures in some way the effect of cyclical

transition on the emissivity of the particular radiation in question. To give a theoretical deduction of the value to be ascribed to this quantity does not appear possible at present. An attempt at what virtually amounts to such a theory was made by Unsöld† and by Pannekoek.‡ A more rational way of approach, however, was indicated by Woolley,‖ as we shall see later on. At present η is considered to be an adjustable constant for each line. Further we shall assume the ratio of selective to general absorption to be independent of optical depth, so that λ itself is also independent of optical depth. Using the simplest solution for mixed scattering and absorption, in which B is assumed to be linear in τ, we have the intensity formula (85). The coefficients a and b in this formula behave rather differently, in that a is practically constant inside a line, while b varies very much. In order to determine how b depends on the frequency we write

$$B = a + b_0 \tau_0,$$

where τ_0 denotes the optical depth in the adjacent continuous spectrum, and b_0 is independent of frequency in the line. Since we have implicitly assumed all coefficients to have the same dependence on density and temperature we have

$$\tau_0 = \frac{\kappa_0 \tau}{\kappa_0 + (\eta+1)\sigma},$$

and consequently

$$b = b_0 \frac{\kappa_0}{\kappa_0 + (\eta+1)\sigma} = b_0\{\lambda - \eta(1-\lambda)\}.$$

Approximate formulae for a and b_0 may be derived as follows. Let T_0 be the boundary temperature of the star as before. The considerations at the end of § 39 made it plausible to assume that the formula

$$T^4 = T_e^4(\tfrac{1}{2} + \tfrac{3}{4}\tau_0) = T_0^4(1 + \tfrac{3}{2}\tau_0)$$

applies also in this case. Hence it follows that

$$\left(\frac{dT}{d\tau_0}\right)_{\tau_0=0} = \tfrac{3}{8}T_0.$$

We have further by definition

$$a = \frac{2h\nu^3}{c^2}(e^{h\nu/kT_0} - 1)^{-1},$$

† *Zeits. f. Physik*, **44** (1927), 481.
‡ Ibid. **46** (1928), 765.
‖ *Monthly Not.* **94** (1934), 631.

and by Taylor's theorem

$$b_0 = \left\{\frac{dT}{d\tau_0}\frac{\partial}{\partial T}\left(\frac{2h\nu^3}{c^2}(e^{h\nu/kT}-1)^{-1}\right)\right\}_{\tau_0=0} = \tfrac{3}{8}a\frac{y}{1-e^{-y}},$$

where

$$y = \frac{h\nu}{kT_0}.$$

Hence, finally,

$$b = a\frac{3y}{8(1-e^{-y})}\{\lambda-\eta(1-\lambda)\}.$$

It should now be clear how the profiles are to be calculated on the hypothesis η = const. for each line profile.

However, this does not mean that the approximation works well under all circumstances. The formation of a line profile is, in fact, an integral effect, in which the contributions from various levels in the solar atmosphere are summed up, and the result may look simple on account of mutual compensations of very complicated variations.

But the variation with level is not out of control. It may, in fact, be controlled in different ways, and first of all by considering the variations of the profiles with the angle of emergence. The nearer to the limb, the higher the effective level of formation of the line. By deriving the average value of η for the same line, but at successive values of the angle of emergence, it is possible to obtain a first idea of the functional dependence of η on height in the solar atmosphere.

A convenient material for testing the theoretical formula is at hand in observations of the so-called b-lines of magnesium, which correspond to the transitions $1\,^3P-1\,^3S$. The lower state of the lines, $1\,^3S$, should be metastable according to the transition rule for normal multiplets, since the next, and lowest, state of magnesium is a $1\,^1S_0$ state (see Fig. 6). But the intercombination line 4571·15 is actually observed. However, it is so weak that the ^3P state is not far from being metastable. The b-lines should therefore be intermediate in character between resonance lines and subordinate lines.

This is fortunate, since we should expect the theory to give most unambiguous results for resonance lines. In this case the chances for the setting up of uncompensated cycles are small, and may, in certain aspects of the problem, be left out of consideration. In this way the main uncertainty involved in the theory is reduced to a minimum.

Measurements of these magnesium lines have been carried out by

several astronomers. d'Azambuja,† working at Meudon, has made a special study of the line 5183·62 A. at the centre of the solar disk. Minnaert and Mulders‡ studied the components 5183·62 A. and

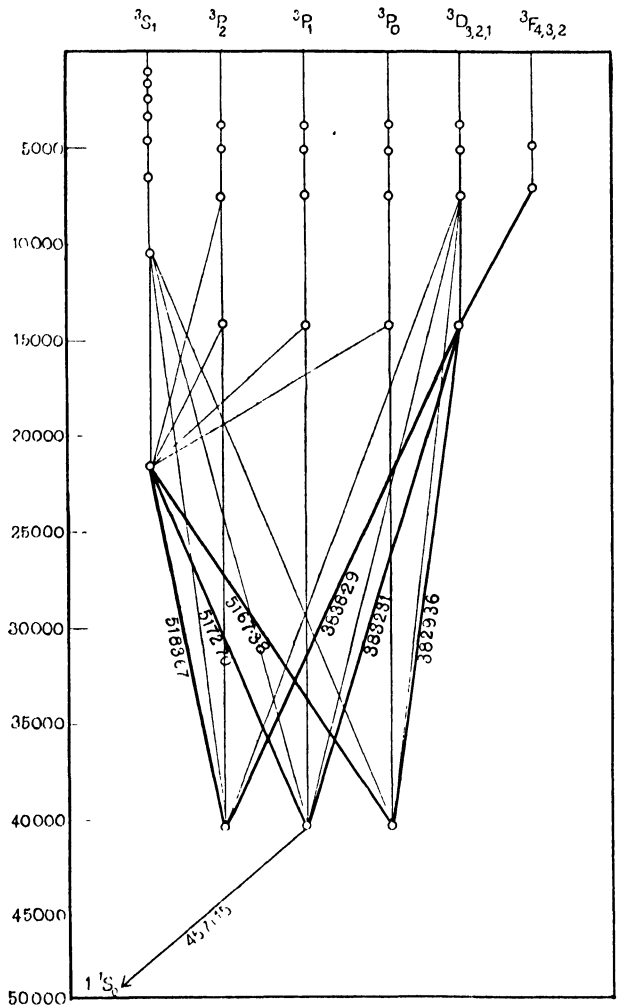


FIG. 6. Level diagram of the magnesium *b*-lines 5167, 5173, and 5184.

5172·70 A., also at the centre of the disk. These same two lines were also the subject of study by Righini,‖ both at the centre and at the

† L. d'Azambuja, *Ann. d. l'Obs. d. Paris*, 8 (1930), 59, 65.
‡ M. Minnaert and G. F. W. Mulders, *Zeits. f. Astrophys.* 1 (1930), 192.
‖ G. Righini, *Mem. d. Soc. Astron. Italia*, 5 (1931), 283.

limb. But the most careful study of the whole triplet is due to H. H. Plaskett.† The results of his work on the two components

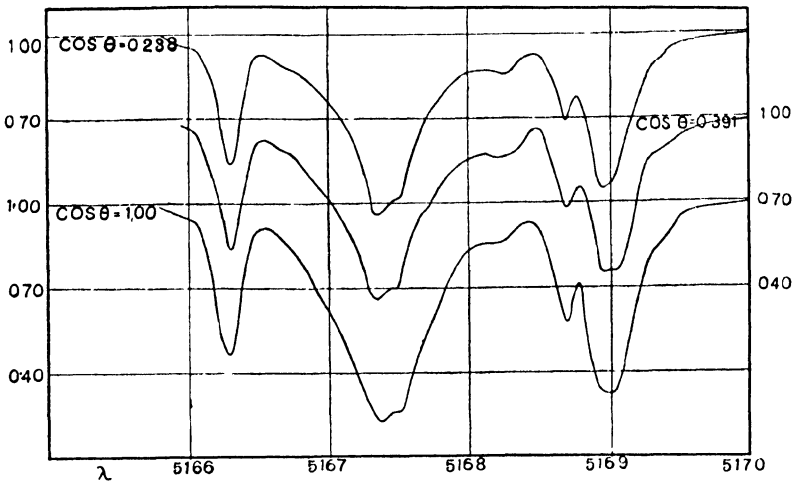


FIG. 7. Profiles of the magnesium *b*-line 5167·33 according to H. H. Plaskett. Ordinates are the ratio of the intensities in the line to that of the adjacent continuous spectrum. The angle θ denotes heliocentric angular distance from the centre of the solar disk. The symbol of this line is $1\,^3P_0-1\,^3S_1$.

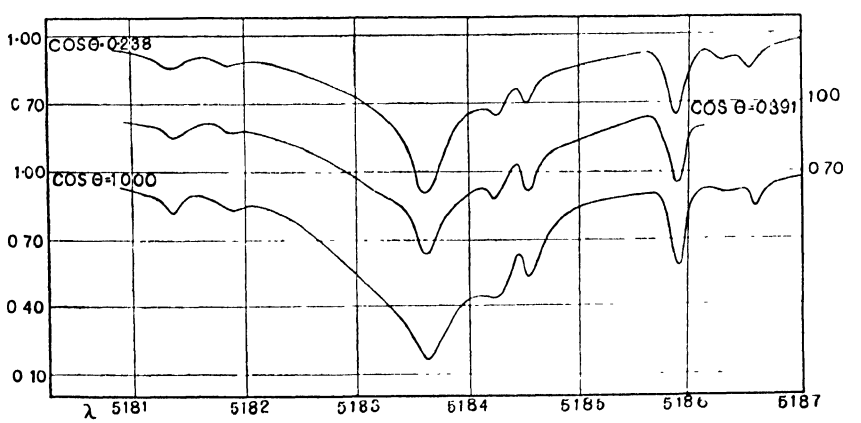


FIG. 8. Profiles of the magnesium *b*-line 5183·62 according to H. H. Plaskett. The symbol of this line is $1\,^3P_2-1\,^3S_1$. For other particulars cf. Fig. 7.

$^3P_0-{}^3S_1$ and $^3P_2-{}^3S_1$ are condensed in Figs. 7 and 8, which should be self-explanatory. The line 5183·62 shows the cleanest profile although even that is deformed by intruders.

The analysis of these profiles on using the intensity formula with

† *Monthly Not.* **91** (1931), 870.

constant η shows that it is possible by a suitable choice of the constants to represent any one of the profiles fairly well. But far from being a constant for different angles of emergence, η is found to increase very much with increasing angle. Adapting, for instance, this formula to Plaskett's line profiles for the Mg line 5183 it is found that η increases from about 0·05 when $\cos\theta = 1$ to about unity when $\cos\theta = 0{\cdot}238$.

A different way of using the equations for constant η in the representation of these profiles was introduced by Plaskett† and further refined by Wiles.‡ This method consists essentially in dropping our assumption that η is independent of the wave-length, and using the latitude thus acquired to conserve the constancy of η at different levels of the atmosphere. The improvement introduced by Wiles consisted in avoiding the use of physically doubtful values of the coefficients, which had occurred in Plaskett's investigation, such as negative values of the coefficients of absorption or scattering. The agreement between theory and observation achieved in this way may be judged from the curves in Fig. 9 which are reproduced from Wiles's paper.

The weakness of this method of analysis lies in the circumstance that it introduces a dependence of the coefficient η on wave-length which is difficult to justify physically. Until a definite physical mechanism has been suggested for this dependence, this method remains essentially formal, and the proper procedure will be to study further cases of dependence of η on optical depth.‖

47. Theory of Cycles

Thus far the residual intensity in the line centres has been accounted for by introducing the empirical quantity η. We shall now consider the theory of the central intensities a little closer by studying the various uncompensated cycles of transitions which are likely to occur in stellar atmospheres.

In an exact theory it would be necessary to inquire in what manner an atom has been brought into a definite quantum state. There is reason to suspect that a complete theory of radiative equilibrium will make the intensity-distribution in a line depend on the route

† *Monthly Not.* **91** (1931), 870. ‡ Ibid. **92** (1932), 401.

‖ An attempt in this direction was made by R. v. d. R. Woolley, ibid. **92** (1932), 482; ibid. **93** (1933), 691.

of the atom before entering the initial state of the line. This side of the question will, however, be neglected in the following. We shall proceed as if the route is irrelevant, and assume that the width of the line is in all cases determined by the rule of Weisskopf and Wigner, obtained by assuming the energy levels to be

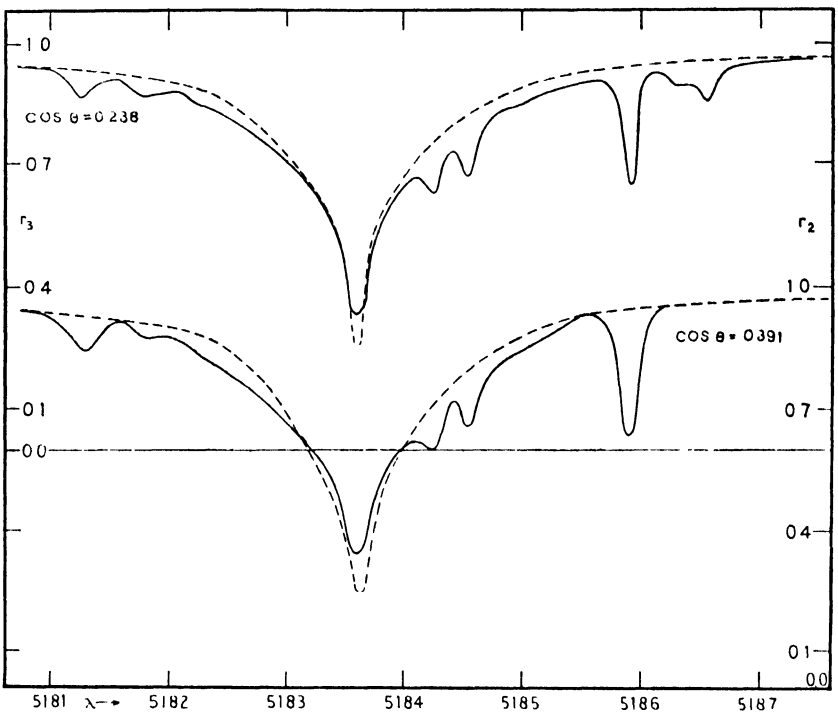


FIG. 9. Theoretical representation of Plaskett's profiles of the magnesium line 5184 by Wiles. The figure shows only the profiles at cos θ = 0·391 and 0·238. The third profile obtained at the centre of the solar disk has been used as a calibration curve for the derivation of the ratio of the coefficients of absorption and scattering, which is used in the representation of the above curves (*Monthly Not.* **92** (1932), 401).

distributed around their mean values according to the statistical law (54).

In the theory of coherent scattering it was assumed that the number of transitions from the lower to the upper level of a line was exactly counterbalanced by the number of transitions going the opposite way. Denoting the two states by i and k and the corresponding numbers of atoms and transition probabilities as usual by n_i, n_k, a_{ik}, and a_{ki} we have the condition

$$n_i a_{ik} = n_k a_{ki}.$$

This is a special case of what is called the principle of detailed balancing, which is supposed always to hold in a state of thermodynamic equilibrium, but which may or may not hold in other cases. It will hold in all cases of a steady state of atoms having but two quantum states, but otherwise it will usually be limited to thermal equilibrium. Considering therefore a closed cycle of transitions of an atom through a given chain of quantum states $i_1, i_2, \ldots, i_1$, the probability that an atom will pass through this cycle will be different for any two opposite directions. The difference between two such probabilities will, in fact, give a sort of measure of the deviation of the state in question from a state of thermal equilibrium.

Let us first prove that if the principle of detailed balancing holds for any two pair of states, it will hold for any closed circuit of transitions. This follows simply by writing down the series of equations:

$$\begin{aligned} n_{i_1} a_{i_1 i_2} &= n_{i_2} a_{i_2 i_1} \\ n_{i_2} a_{i_2 i_3} &= n_{i_3} a_{i_3 i_2} \\ \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \\ n_{i_{n-1}} a_{i_{n-1} i_n} &= n_{i_n} a_{i_n i_{n-1}} \\ n_{i_n} a_{i_n i_1} &= n_{i_1} a_{i_1 i_n}. \end{aligned}$$

Multiplying these equations as they stand, all the n_i's drop out, leaving as a result

$$a_{i_1 i_2} a_{i_2 i_3} \cdots a_{i_n i_1} = a_{i_1 i_n} a_{i_n i_{n-1}} \cdots a_{i_2 i_1},$$

which shows that the probabilities of an atom completing the circuit in opposite directions are equal.

In the stellar atmospheres, and still more in the nebulae, the radiation does not correspond to any definite temperature, so that any two cycles will not have the same probability in the above sense. This means that the radiation absorbed in a given line will partly be re-emitted in other lines, a process which in usual terminology is called fluorescence.

When the distribution of atoms over the different quantum states can no longer be calculated from the principle of detailed balancing, we must set up, for each state, an equation expressing the fact that the increase in the number of atoms in this state is the difference of the number of atoms entering the state and those leaving it, all per unit time. Considering the k-state, for example, the number of

atoms leaving it per second is, in the usual notation,

$$\sum_i n_k a_{ki},$$

the sum being extended to all quantum states of the atom. The number of atoms entering the state per second is, similarly,

$$\sum_i n_i a_{ik},$$

and the net increase in n_k per second is, consequently,

$$\frac{dn_k}{dt} = \sum_i (n_i a_{ik} - n_k a_{ki}).$$

The summation on the right-hand side of this equation does not involve a term of index k, since any such term would have to appear both in the negative and the positive sum, and to cancel out in the end. This fact may be used to write the equation in a more symmetrical form. We may, in fact, write

$$a_{kk} = -\sum_i a_{ki} \qquad (k \lesseqgtr i),$$

in consequence of which the above equation assumes the form

$$\frac{dn_k}{dt} = \sum_i n_i a_{ik}.$$

Under stationary conditions the left-hand member of this equation is zero, and the distribution is found by solving the following system of *ordinary* linear equations

$$\sum_i n_i a_{ik} = 0 \qquad (k = 1, 2, \ldots).$$

According to the theory of linear equations the solution may be written in the form

$$n_i = \lambda_m a^{im},$$

where a^{im} is the minor corresponding to a_{im} in the determinant of all a_{ik} (the algebraic complement of a_{im}), m is an arbitrary index, while λ_m is an arbitrary constant. This constant will be determined if the sum of all n_i's is given, for then we have obviously

$$\lambda_m = \sum_i n_i \Big/ \sum_i a^{im}.$$

For the above solution to be possible the determinant a_{ik} must vanish. This is identically the case in virtue of the definition of a_{kk}. This, namely, makes the determinant a_{ik} belong to the type in which the elements of a row are connected with the elements in parallel

rows by the same linear relation. Such a determinant may obviously be written as a sum of determinants with two identical rows, which are known to vanish identically. While the *formal* solution is thus easily found, it is quite a different matter to find a desired *numerical* solution. To calculate a determinant, and its various minors, may be feasible as long as it is of low order, but the amount of work required increases enormously with increasing order of the determinant, and soon becomes prohibitive. But this is another side of the question—let us just assume the operations possible, and consider the form of the coefficients a_{ik}.

As was shown by Einstein, and also shown earlier in this book, a coefficient a_{ik} leading from a higher to a lower state has, in the case of thermal equilibrium, the form

$$a_{ik} = \alpha_{ik}\left(1 + \frac{c^2 J}{8\pi h \nu_{ik}^3}\right), \tag{86}$$

where J is the *total* intensity of radiation per unit frequency range near the centre of the line, ν_{ik} the frequency of the line, and α_{ik} the probability of spontaneous transitions $i \to k$. The coefficient a_{ki} has the form

$$a_{ki} = \alpha_{ik}\frac{\sigma_i}{\sigma_k}\frac{c^2 J}{8\pi h \nu_{ik}^3}, \tag{87}$$

where σ_i and σ_k denote the weights of the states i and k. The above relations also hold in other cases, provided only the intensity of the radiation field does not vary sensibly over the main region of the lines.

When the intensity varies very rapidly in the line it may be necessary to express the transition probability as a function of the monochromatic absorption coefficient κ. We have then to a sufficient approximation

$$a_{ki} = \frac{1}{h\nu_{ik}}\int_0^\infty \kappa J \, d\nu.$$

For the sake of comparison it is advantageous to normalize the absorption coefficient by writing

$$\bar{\kappa} = \frac{\kappa}{\alpha_{ik} c^2}\frac{\sigma_k}{\sigma_i} 8\pi\nu_{ik}^2,$$

and then to write

$$\bar{J} = \int_0^\infty \bar{\kappa} J \, d\nu.$$

The transition probabilities then assume the form

$$a_{ki} = \alpha_{ik}\frac{\sigma_i}{\sigma_k}\frac{c^2}{8\pi h\nu_{ik}^3}\bar{J}; \qquad a_{ik} = \alpha_{ik}\left(1+\frac{c^2}{8\pi h\nu_{ik}^3}\bar{J}\right).$$

48. Cycles in Dilute Temperature Radiation

When the radiation field corresponds to that of temperature radiation at a temperature T, *except* for a constant multiplier W less than unity, the field will be said to be one of dilute temperature radiation:

$$J = W\frac{8\pi h\nu^3}{c^2}(e^{h\nu/kT}-1)^{-1}. \tag{88}$$

This case corresponds in a general way to the radiation field to be found in the stellar atmospheres, and still more in the nebulae. The quantity W has then the physical significance of the solid angle subtended by the radiating star in units of 4π, and so is always a quantity less than unity. Since such a field is in a certain sense representative for the fields to be reckoned with in the analysis of stellar atmospheres and nebulae, it is of general interest to study the cycles it will generate.

The simplest non-trivial case of a cycle is given by considering three states only, 1, 2, and 3 say, taken in the order of increasing energy. Selecting the minors in the first row ($k = 1$) of the determinant a_{ik} as expressions for the numbers n_1, n_2, n_3 we find easily, on leaving out the multiplier,

$$n_1 = a_{21}(a_{31}+a_{32})+a_{23}a_{31},$$

$$n_2 = a_{12}(a_{31}+a_{32})+a_{13}a_{32},$$

and

$$n_3 = a_{13}(a_{21}+a_{23})+a_{12}a_{23}.$$

From these expressions we find

$$n_2a_{21}-n_1a_{12} = n_3a_{32}-n_2a_{23} = n_1a_{13}-n_3a_{31} = \Delta,$$

where

$$\Delta = a_{13}a_{32}a_{21}-a_{12}a_{23}a_{31}.$$

Using the expressions for the transition coefficients which result from the combination of (88) with (86) and (87) we find

$$a_{12}a_{23}a_{31} = ya_{13}a_{32}a_{21},$$

where

$$y = W\frac{F_{31}}{F_{32}F_{21}},$$

$$F_{ik} = 1-(1-W)e^{-h\nu_{ik}/kT}.$$

When W is less than unity, as always in problems of interest to astrophysics, y is positive and less than unity. Hence the quantity

Δ is positive. It is then easier for the atom to perform a cycle $1 \to 3 \to 2 \to 1$, involving one absorption act and two emissions, than it is to go the opposite way with two absorptions and one emission. The presence of the atoms in the radiation field will therefore tend to increase the intensity of the field in the lines corresponding to the long wave-length transitions $3 \to 2$ and $2 \to 1$, and to reduce by a corresponding amount the energy in the region of the short wave-length line $3 \to 1$.

This holds for the total field, independent of direction. It may be well to point out, as was emphasized by Woolley, that the result cannot be applied indiscriminately to the intensity in a definite direction. It may well be that the effect on a parallel beam of light will still be to create three absorption lines, since the emitted radiation is distributed over all directions. Only when the cycle has had the chance to deepen the absorption lines beyond a certain limit will the excess emission become great enough to counterbalance the directional effect.

The limitation to three states is a matter of pure convenience, and it is clear that cycles involving a larger number of states will show qualitatively the same result. In practical cases the limitation to three states is not so fortunate as it is not possible to find three ordinary quantum states between which all transitions are possible without violation of the selection principle for the inner quantum number l. But this difficulty is lessened by the fact that ordinarily the ionized state will play an important part in the cycles, and since the ordinary selection principles do not apply in the same strict sense to this state, it may happen that a three-state cycle will be of service in the elucidation of some points.

The effect of cycles, as far as stellar atmospheres are concerned, will be to fill out long wave-length lines and to darken those of short wave-lengths, or the continuous absorption beyond the series limits. It is probably this effect which accounts for the relatively high central intensity of the resonance lines for which the theory of coherent scattering would predict a zero central intensity. Most likely it plays a similar part in the case of many subordinate lines.

49. Woolley's Theory of the Hydrogen Lines

Very little work has been done on the actual calculation of the effect of cycles on the intensities of the lines in the solar spectrum or

in the stars. Some preliminary work was done by Unsöld† and by Pannekoek,‡ but the first paper of any real consequence in this field was that of Woolley‖ on the fluorescence in H_α and H_β. The objection may at once be raised against the paper that Woolley combines the theory of cycles with the formulae of coherent scattering in his calculations, although H_α and H_β are typically subordinate lines. This objection can scarcely be disregarded as unimportant, and it is to be expected that Woolley's result will suffer considerable revision when subjected to a more critical theoretical analysis. The paper nevertheless embodies so many important points that it deserves to be reported in full here as a valuable pioneering work, which brings to light some of the complexities to be coped with in the analysis of stellar atmospheres.

The smallest number of states which can be considered in a theory of H_α and H_β is *four*, that is, the three states involved in the production of the two lines plus the ground state. Woolley, however, discovered that transitions to the ionized state also affected the result considerably, so this had to be included as a fifth state. All other states were neglected.

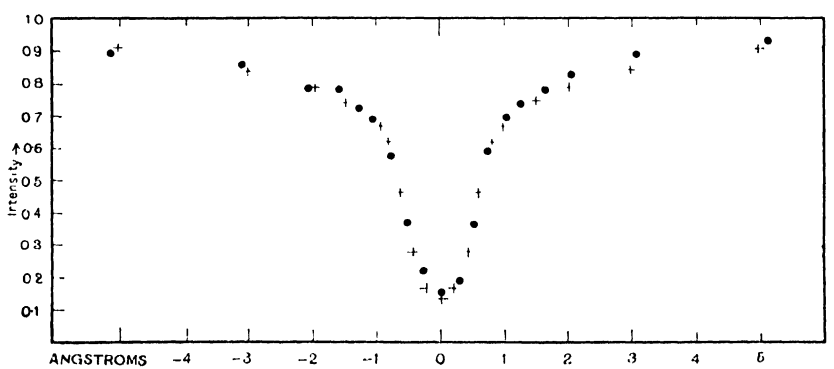


FIG. 10. H_α profile. Filled circles are observations by Thackeray (*Monthly Not.* **95**), and crosses by Minnaert (*Zeits. f. Physik*, **45**).

In the case of hydrogen the transition coefficients are known from theory.†† In the case of transitions from the continuum into one of the states considered some difficulty is encountered due to the fact

† A. Unsöld, *Zeits. f. Astrophys.* **2** (1931), 199.
‡ *Bull. Astron. Inst. Netherlands*, **7** (1933), 151.
‖ *Monthly Not.* **94** (1934), 631.
†† Ibid.

that the temperature of the free electrons is involved. But it so happens that in the range involved, that is, 6,000–7,000°, the transition probabilities are rather insensitive to changes in temperature. The temperature adopted by Woolley in his calculations was 6,500°, which is the colour temperature of the sun as estimated by H. H. Plaskett† for integrated sunlight. The effective temperature as estimated from the total radiation is, on the other hand, nearer to 6,000°. The difference between these two temperature values should obviously be taken into account when the deviation of sunlight from that of thermal radiation is described by a coefficient of dilution simply. Woolley does this by reducing the dilution coefficient in the ratio $(60/65)^4$. This coefficient should otherwise have the value $\frac{1}{2}$, since radiation in the solar atmosphere is preferentially coming from directions within a solid angle 2π. The final value adopted by Woolley was thus

$$W = \frac{1}{2}\left(\frac{60}{65}\right)^4 \sim \frac{1}{3}.$$

The correction in W due to the difference between effective temperature and colour temperature is just such as to give the right values for these two quantities, but otherwise the schematic radiation field adopted may differ from the actual solar field, and the correction can at most be considered a first approximation. From these data the following tables were calculated.

TABLE 1

Continuous Absorption and Emission in Hydrogen

$T = 6{,}500°$; $W = \frac{1}{3}$

State	a_{n5}	a_{5n}
1*s*	0·0378	113
2*s*	$1{\cdot}45.10^5$	17·9
2*p*	$7{\cdot}87.10^4$	9·67
3*s*	$7{\cdot}06.10^5$	5·04
3*p*	$5{\cdot}23.10^5$	3·64
3*d*	$2{\cdot}41.10^5$	1·72
4*s*	$7{\cdot}32.10^5$	1·48
4*p*	$6{\cdot}32.10^5$	1·39
4*d*	$4{\cdot}40.10^5$	0·97

† *Pub. Dom. Astrophys. Obs. Victoria*, **2** (1923), 213.

TABLE 2

Line Transitions in Hydrogen. $T = 6{,}500°$; $W = \frac{1}{3}$ $(i > k)$

State	a_{ik}	a_{ki}
$1s-2p$	$6{\cdot}32.10^8$	$8{\cdot}59$
$1s-3p$	$1{\cdot}69.10^8$	$0{\cdot}080$
$1s-4p$	$6{\cdot}8.10^7$	$0{\cdot}010$
$2s-3p$	$2{\cdot}28.10^7$	$8{\cdot}19.10^5$
$2p-3s$	$0{\cdot}648.10^7$	$2{\cdot}64.10^4$
$2p-3d$	$6{\cdot}55.10^7$	$1{\cdot}30.10^6$
$2s-4p$	$0{\cdot}978.10^7$	$1{\cdot}06.10^5$
$2p-4s$	$0{\cdot}259.10^7$	$3{\cdot}14.10^3$
$2p-4d$	$2{\cdot}08.10^7$	$1{\cdot}26.10^5$

The ratio of downward to upward transitions in a line is denoted by

$$\alpha = \frac{n_i a_{ik}}{n_k a_{ki}},$$

the i-state being as usual supposed to have the higher energy. In order to evaluate α for the two lines H_α and H_β we must replace n_i and n_k by their equivalent minors in the determinant of all a_{ik}. This determinant is in the present case of the fifth order, and thus already rather difficult to handle. But a short cut to the evaluation of the minors is afforded by the fact that some terms are zero and others negligible, although really finite. Working out the strongest transitions for $H_\alpha(2p \to 3d)$ we have $a_{13} = a_{31} = 0$, and also $a_{32} \gg a_{34}+a_{35}$, so that $a_{32} = -a_{33}$ very nearly. Similarly, $a_{21} = -a_{22}$. Selecting the minors in the bottom row of the determinant a_{ik} as equivalents of the number of atoms in the different states, we need for H_α the minors a^{25} and a^{35}.

These are $a^{25} = a_{32}(a_{12}+a_{15})(a_{52}+a_{53})+a_{12}a_{32}a_{51}$,

and $a^{35} = a_{12}a_{23}(a_{51}+a_{52})+a_{15}(a_{21}a_{53}+a_{23}a_{52})$,

where the approximations indicated above have been introduced. We can further neglect a_{15} in $(a_{12}+a_{15})$, $a_{23}a_{52}$ in $(a_{21}a_{53}+a_{23}a_{52})$, and a_{53} in $(a_{51}+a_{52}+a_{53})$. This gives then

$$\frac{a^{35}}{a^{25}} = \frac{a_{12}a_{23}(a_{51}+a_{52})+a_{15}a_{21}a_{53}}{a_{12}a_{32}(a_{51}+a_{52})},$$

and finally $$\alpha = \frac{a^{35}a_{32}}{a^{25}a_{23}} = 1+\frac{a_{15}a_{53}a_{21}}{a_{12}a_{23}(a_{51}+a_{52})}.$$

The corresponding expression for H_β is

$$\alpha = 1 + \frac{a_{15}a_{54}a_{21}}{a_{12}a_{24}(a_{51}+a_{52})}.$$

Introducing numerical values for the transition probabilities Woolley finds

$$\text{for } H_\alpha : \alpha = 1{\cdot}03; \text{ for } H_\beta : \alpha = 1{\cdot}17.$$

These values of α differ so little from unity that it will be understood at once that the cycles will not appreciably affect the exchange of radiation in the region of the two lines. The form assumed for the radiation field can, however, only be supposed to apply in the region of the lines at some depth below the surface, and on approaching the surface the absorption lines will begin to form as a result of ordinary selective absorption and scattering. This will serve to make α increase. It is important to note in this connexion that since the Lyman α-line has a much larger absorption coefficient than the other lines involved, its effect on the radiation field will be confined to a thinner surface layer than in the case of the other lines, a fact which was well brought out in the theory of coherent scattering, and which must hold generally. It is therefore largely permissible to assume the radiation field in the region of Lyman α to correspond to dilute radiation, while the lines H_α and H_β are well under way to become deep absorption lines.† We shall therefore neglect the formation of Lyman lines as absorption lines. The average intensity in the lines H_α and H_β, in units of the adjacent continuous spectrum will be denoted by $\bar{J}_\alpha$ and $\bar{J}_\beta$ respectively. The process of averaging involved was defined earlier. We have then

$$H_\alpha : \alpha - 1 = \frac{a_{53}a_{21}a_{15}}{\bar{J}_\alpha a_{23}a_{12}(a_{51}+a_{52})}; \qquad H_\beta : \alpha - 1 = \frac{a_{54}a_{21}a_{15}}{\bar{J}_\beta a_{24}a_{12}(a_{51}+a_{52})}.$$

These expressions bring out clearly how the formation of the absorption lines tends to strengthen the cyclical transitions, and it is clear that for some value of α the further deepening out of the absorption line will be prevented by the cycle itself.

The value of α at which an absorption line ceases to grow further must be in the vicinity of 2. The argument leading to this result is the following. Considering the case of ordinary coherent scattering, the reason for the growth of an absorption line is to be found in the

† The force of this argument is brought out in a more precise manner in the theory of nebulous envelopes. See § 121.

continuous transformation of radiation coming from directions within a solid angle 2π, approximately, into radiation distributed over 4π of solid angle. The re-emitted radiation is therefore only half as intense as the radiation absorbed, and an absorption line is the inevitable result. Suppose, however, that the cycle brings in some extra radiation into the emitted flow. As long as this radiation does not make up for the whole lost half of the incident radiation, the absorption line will continue to grow. The limiting state of no further growth is obviously reached when the extra radiation makes up for the full half, which corresponds to a value 2 of α.

This argument holds approximately in an ordinary stellar atmosphere, where the absorbed radiation emerges from a solid angle 2π only. In the case of abnormal stars, say *O*-stars or planetary nebulae or novae, conditions may be very different.

Putting $\alpha = 2$ in the above expressions and solving for $\bar{J}_\alpha$ and $\bar{J}_\beta$ we find at once

$$\bar{J}_\alpha = 0{\cdot}03 \text{ and } \bar{J}_\beta = 0{\cdot}17.$$

These are minimum values for the average intensities in these lines. It will be realized that the averaging process is such as to make the average intensity correspond very closely to the intensity near the centre of the line, since the weight function used in the averaging process has an extremely sharp maximum at the centre. The above minimum values are well below those actually found in the solar spectrum. The theory is therefore on the safe side so far.

To push the approximation a step farther it will first of all be necessary to correct for the gradual depletion of the radiation in the Lyman α-line. This correction will be in the direction of increasing the central intensities, since a weakening of the radiation in this line will provide fewer atoms in the second state, and consequently also fewer absorption transitions in the H_α and H_β lines. As these calculations are rather lengthy and do not bring in any essentially new points, we shall only give the result. It seems that the effect is to bring the central intensity of H_α very nearly up to that of H_β, or a little more, both being of the order 0·20, which again is not far away from the observational estimate, which varies from 0·25† to 0·15.‡

A number of other points were discussed by Woolley in the same connexion, indicating that cycles were at work. He called attention

† A. Unsöld, *Zeits. f. Astrophys.* **2** (1931), 199.

‡ A. D. Thackeray, *Monthly Not.* **95** (1935), 296.

to the fact that mottling of the solar surface in integrated light (integrated over all frequencies) is accompanied by a corresponding

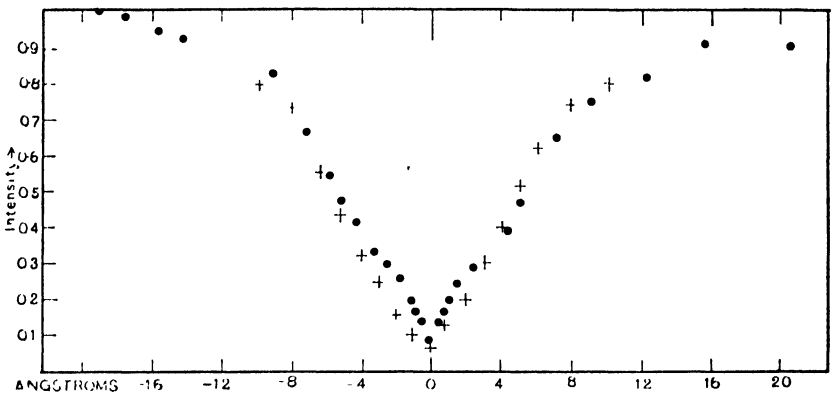


FIG. 11. Profile of Ca^+ *K* line. Filled circles observations by Thackeray (*Monthly Not.* **95** (1935), 297), and crosses by Minnaert (*Zeits. f. Physik*, **45**).

mottling of the surface when seen in the light of the H_α-line by a spectrohelioscope, or photographed by a spectroheliograph. This means that a local increase in temperature of the surface tends to fill out the H_α-line. This is, however, in conformity with Woolley's calculations of the correction to the intensity due to the formation of Lyman-α, which we have omitted. The same calculations indicated no corresponding effect for H_β, and none is observed. It is further suggested that the reason why hydrogen prominences usually show up as dark markings, instead of bright, is that they lie above the fluorescent level, so that the light reaching them has already been depleted in ultra-violet light. The peculiar form of the calcium lines may also be interpreted by considering a similar division of the solar atmosphere in successive layers of absorption and fluorescent emission. On the whole, the calculations may be taken to indicate that it is necessary to take the cycles into account in the theory of all solar absorption lines. This was already suggested explicitly by the present author some years ago.†

In a recent paper Strömgren‡ has shown how it is possible to formulate a general theory of central intensities, based on cycles involving electron captures, for the case of principal lines.

The importance of the theory of cycles for the interpretation of bright line spectra will be considered later on.

† *Astrophysik auf atomtheoretischer Grundlage*, by S. Rosseland, Berlin, 1930, p. 144.

‡ *Zeits. f. Astrophys.* **10** (1935), 237.

50. Subordinate Lines

There is at present one essential feature missing, which must be supplied before a rigorous theory of subordinate lines can be formulated. This is the equivalent of Weisskopf's theory of resonance radiation and dispersion. A first investigation in this field was made by Breit and Lowen.† We do not know at present in what proportion radiation scattered in a subordinate line is reflected without essential change in wave-length, or how much is diverted into essentially different spectral regions. Before this question is definitely settled, further progress is difficult. An application of Weisskopf and Wigner's probability law would indicate that conditions in a subordinate line will differ radically from that of a resonance line in this respect, in the sense that the light scattered coherently will carry only a small part of the absorbed energy, and that this part of the radiation is re-emitted in the form of a rather broad line, with a width at least as large as the width of the initial state of the line. There will then be no question of dispersion in the ordinary sense of the word, and a first approximation might be had by proceeding as if it were a question of spontaneous emissions and the reverse in the ordinary sense of the word, as in the case of radiation under electron impacts, for example. For the present, however, this is little more than guesswork, which must be substantiated or disproved by exact calculations. But the arguments are sufficiently strong to inspire distrust in an indiscriminate application of the formula of coherent scattering to the formation of subordinate lines.

51. The Equation of Ionization

Disregarding for the moment the elements of uncertainty involved in the theory of line profiles, we turn to the question of what information may be obtained about the state of the solar and stellar atmospheres from the present state of the theory.

The principal points are best brought out by considering the ideal case of an atmosphere in which the line profiles may be analysed on the simple theory characterized by constant λ and η in our earlier notation. This means, furthermore, that σ/κ_0 is also independent of optical height.

When each of these coefficients is proportional to density, the ratio σ/κ_0 is equal to the ratio of the corresponding optical depths τ/τ_0.

† *Phys. Rev.* **46** (1934), 590.

The adaptation of the theoretical intensity formula to an observed line profile gives then a definite value of σ/κ_0 or τ/τ_0. Supposing τ_0 to signify the optical depth of the atmosphere, τ receives the significance of the optical depth of the atmosphere in the given frequency interval, for selective scattering only. Dividing this quantity by the coefficient of scattering of a single atom we obtain the number of atoms per square cm. above the photosphere. This number will be denoted by N in the following, and will be considered the most direct physical datum about the state of the atmosphere obtainable from line profiles.

If a reliable theory of continuous absorption were also in existence, it would be possible to obtain more information from the simultaneous discussion of τ and τ_0. We shall consider a case of this kind later on. But the origin of the continuous absorption in the atmosphere is still in a much more obscure state than is the case of the selective absorption. This fact need not surprise us. For the lines may be immediately identified from their wave-lengths and series relationship. But there is nothing, or only very little, about the continuous absorption which indicates an identification with particular elements responsible for the absorption processes. And when we do not know the absorbing element it is not possible to propose a definite theory for the general absorption.

When it happens, as it does in the solar spectrum in the case of calcium and strontium, that resonance lines of the neutral and the ionized element are present, it is possible to determine the state of ionization of the particular element. A first indication of this state is, namely, immediately obtained from the ratio N^+/N of the corresponding N values. In the case of calcium this ratio is, for instance, of the order 400, so that the ionized calcium atoms dominate completely over the un-ionized atoms. From this ratio it is possible from the theory of dissociation to calculate a mean density of free electrons in the reversing layer, or, by a further elaboration of the theory, the distribution of this density from level to level in the atmosphere. This result is of great importance for stellar theory.

In order to show this clearly it is necessary to write the equations of dissociative equilibrium in a somewhat different form. We neglect molecular combinations and consider first the equation of dissociative equilibrium for two constituents, which may be taken as nuclei and electrons. Let λ_0 and λ_1 denote the corresponding parameters

entering equations (11) of dissociative equilibrium. The numbers per unit volume of free electrons n_e and of atoms with p bound electrons n_p are then given by

$$n_e = f_e e^{\lambda_1} \quad \text{and} \quad n_p = f_p e^{p\lambda_1 + \lambda_0},$$

f_p and f_e denoting the corresponding partition functions. Deviations from the ideal gaseous state are neglected. Eliminating the parameters we obtain the equation

$$\frac{n_e n_p}{n_{p+1}} = \frac{f_e f_p}{f_{p+1}}. \tag{89}$$

The partition function is defined as the integral

$$f = \int e^{-E/kT}\, d\sigma,$$

where $d\sigma$ is an element of volume in the phase space and E the energy of the system. The integration is to be extended to the total accessible part of the phase space.

In the quantum theory this integration reduces to a summation over all quantized states, each state extending over a volume of phase space of h^s, s being the number of degrees of freedom of the system. That this is the correct procedure is evident from the quantization rules, particularly in the old form given by Sommerfeld, and this choice is amply confirmed by experiments.

The partition function as defined above really involves an arbitrary scale factor, and in view of the requirements of quantum theory it is convenient to fix this factor by defining the partition function as follows

$$f = \int e^{-E/kT} \frac{d\sigma}{h^s},$$

which has the additional advantage of making f zero-dimensional. For a free particle of mass m, rectangular coordinates x, y, z, and corresponding velocity components u, v, w we have then

$$f = \int e^{-m(u^2+v^2+w^2)/2kT} \frac{m^3}{h^3}\, du dv dw = \frac{(2\pi mkT)^{\frac{3}{2}}}{h^3}.$$

Since the translational motion of an atom is independent of its internal state, each of the functions f_p and f_{p+1} will contain $(2\pi mkT)^{\frac{3}{2}}$ as a factor.

The partition function corresponding to the internal, quantized states of an atom assumes the simple form

$$f' = \sum_i w_i e^{-E_i/kT},$$

where w_i is called the weight of the ith state, and is simply equal to the number of intrinsically different quantum states which have the same energy E_i (degenerate systems). It is frequently convenient to count quantum states in groups in this way even though their energies may be slightly different and the systems consequently non-degenerate.

When introducing for E_i the usual expression for the energy levels in the hydrogen atom, it will be found that the series diverges violently. This difficulty, however, is purely formal, and is due to the fact that the usual Balmer expression for the hydrogen terms will break down in an actual gas when sufficiently high quantum states are considered. From a certain state onwards the interaction between the moving electron and the surrounding gas atoms will become so great that the electron cannot any longer be counted as bound to any particular atom, but must be considered as free. This argument may be put in a more rigorous form by an application of Debye and Hückel's theory of the interaction between charged particles in a gas,† but the above general arguments should suffice for the present purpose.

We may now write down the complete expressions of the partition functions f_e and f_p. The free electron has no internal energy to be considered, but its weight is 2, owing to the two equivalent possible orientations of the spin. Its partition function is therefore

$$f_e = \frac{2(2\pi\mu kT)^{\frac{3}{2}}}{h^3}$$

simply.

Let M denote the mass of an atom. Since M is in all cases much bigger than μ, it is permissible to use the same value of M in all stages of ionization. Then

$$f_p = \frac{(2\pi MkT)^{\frac{3}{2}}}{h^3} \sum_i w_i e^{-E_i/kT}.$$

It is convenient to take the exponential factor of the ground term in the series outside the summation sign, and to define the energies of the excited states relative to the ground state, by putting

$$E_i - E_0 = X_i.$$

The partition function of the atom then assumes the form

$$f_p = \frac{(2\pi MkT)^{\frac{3}{2}}}{h^3} e^{-E_0^p/kT} \sum_i w_i^p e^{-X_i^p/kT}.$$

† *Physikal. Zeits.* **24** (1923), 1.

We have added an index p to X_i and E_0 so as to be able to handle different stages of ionization. We further introduce a special symbol for the series in the above expression:

$$B_p = \sum_i w_i^p e^{-X_i^p/kT}.$$

Using this notation the equation of ionization (89) becomes finally

$$\frac{n_p n_e}{n_{p+1}} = \frac{2(2\pi\mu kT)^{\frac{3}{2}}}{h^3} \frac{B_p}{B_{p+1}} e^{-\chi/kT}, \tag{90}$$

where

$$\chi = E_0^p - E_0^{p+1}$$

is just the energy of ionization of the atom with $p+1$ electrons.

In Fig. 12 there is given a graphic representation of the above equation. The dependence of B_p/B_{p+1} on temperature is neglected, and the electron density is taken to be a constant and to have such a value at a temperature of about 6,000° K. that an element of ionization potential of 8·5 volts is just 50 per cent. ionized. This

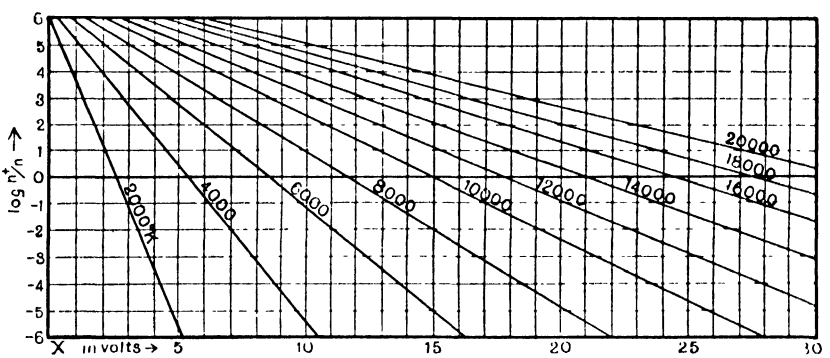


Fig. 12. Graphical representation of the state of ionization in an atmosphere. Ordinates are values of $\log(n_p/n_{p+1})$, and abscissae are values of the ionization energy in volts. Although the diagram involves several rough approximations, it forms a useful means of obtaining a rapid survey of the state of ionization to be expected under given circumstances.

should correspond nearly to conditions in the reversing layer of the sun, as will be apparent later. The diagram then brings out fairly well how rapidly the ionization diminishes in elements with higher ionization potentials, or increases for smaller ionization potentials. Calcium ($\chi = 6{\cdot}09$ volts) is seen to be more than 99·5 per cent. ionized, while in hydrogen ($\chi = 13{\cdot}6$ volts) there is scarcely one ionized atom per fifty thousand neutral atoms. These figures cannot claim any great accuracy, but they show well how, at a definite

temperature, the state of ionization is determined by the temperature and the relevant ionization potential.

The equation of ionization as considered above corresponds to matter in thermal equilibrium. In a stellar atmosphere, however, and still more in a nebula, there are definite departures from a state of thermal equilibrium, which must affect the state of ionization as well. The case of dissociative equilibrium in the presence of a temperature gradient was considered by Dirac.† The deviations from thermal equilibrium in a stellar atmosphere are, however, of a slightly different kind. The ionizing processes are then likely to be due to a photoelectric effect, and the rate of these processes will therefore depend primarily on the radiation field. The rate of the recombination processes, on the other hand, is determined by the velocity of the free electrons. Describing the state of the radiation field by a temperature T' while that of the free electrons is T, it is easy to show that to the first order the corresponding form of the equation of ionization is

$$\frac{n_p n_e}{n_{p+1}} = \frac{2(2\pi\mu kT)^{\frac{3}{2}}}{h^3} \frac{B_p}{B_{p+1}} e^{-\chi/kT'}$$

as was first shown by Pannekoek.‡ The dependence of B_p/B_{p+1} on temperature is neglected. Since we are going to consider this question more in detail later (§ 115), we refrain from pursuing the matter further here.

The ionization formula has thus far referred to the density of atoms and electrons. The quantities given by observation, however, are rather the amount of matter per square cm. above a given level (the photosphere). Let n_0 and n^+ denote the densities of neutral and singly ionized atoms, and let N_0, B_0 and N^+, B^+ correspond to n_0 and n^+. The ratio N^+/N_0 then gives a kind of integrated mean value of n^+/n_0, and since the density diminishes very rapidly with height, it is to be expected that N^+/N_0 will nearly correspond to n^+/n_0 at a low level. We are thus able to define a corresponding average of the electron density, which we denote by $\bar{n}_e$, and which then is given by

$$\bar{n}_e = \frac{N_0}{N^+} \frac{2(2\pi\mu kT)^{\frac{3}{2}}}{h^3} \frac{B^+}{B_0} e^{-\chi/kT'}.$$

† *Proc. Cambridge Phil. Soc.* **22**, ii (1924), 132.

‡ *Bull. Astron. Inst. Netherlands*, **3** (1926), 207. Cf. also Woolley, *Monthly Not.* **93** (1933), 696.

This result is interesting, as it enables us to obtain an estimate of the electron density in the photospheric layer of a star without postulating any definite theory of general opacity.

On the other hand, if a satisfactory theory of general opacity is available, this will give a value for $\bar{n}_e$, and the above equation of ionization may be used to obtain the effective temperature. Considering the difficulties which up to this time have beset the problem of effective temperature, this may prove a useful check method. Thus Woolley† has tried to obtain an independent determination of the effective temperature of the sun by postulating a definite theory of the continuous absorption coefficient. The result proved to be in good agreement with determinations of effective temperature by other methods.

† *Monthly Not.* **93** (1933), 691.

XI

THE TOTAL INTENSITY OF ABSORPTION LINES

52. Theory of Total Intensities

THE observation of a clean profile in stellar absorption lines is a rather rare event. As regards the sun it is estimated† that the equipment available at present cannot give true profiles for lines weaker than 4 on the Rowland scale. The total intensity of a line, however, is independent of the profile, and is not affected very much by the resolving power of the equipment in use. It is therefore very important to bring out the kind of information which is contained in observations of total absorption in the lines.

This problem has already been discussed by several authors, both from the theoretical and the observational side, and it appears that in principle the theory is much farther advanced than the theory of line profiles. Estimates of the intensities of solar lines were given by Rowland‡ along with his measurement of the wave-lengths of the lines. These estimates were used by Russell‖ in an attempt purporting to obtain a wholesale determination of relative abundance of the elements in the solar atmosphere. Further refinement of this work is due to Minnaert and Mulders,†† and Minnaert and Slob,‡‡ who also gave the first comprehensive discussion of the theoretical basis of the problem.

While the problem is already important for solar work, it becomes much more so when turning to the stars. First of all, the interpretation of the spectral classes of the stars will rest upon the theory of total intensities, both as regards the run of the individual series as well as the splitting of the classification into several parallel series and all spectral absolute magnitude effects.

Since the problem thus has wide implications, it will be well to consider it from every angle. We begin by considering the total intensity of lines when the emergent intensity is sufficiently described by the simple theory repeatedly considered before, which is charac-

† Cf. for instance A. Unsöld, *Astrophys. J.* **75** (1932), 111, or Minnaert and Mulders, *Zeits. f. Astrophys.* **2** (1931), 173.

‡ *Astrophys. J.* **1** and **2** (1894).

‖ Ibid. **68** (1928), 279.

†† *Zeits. f. Astrophys.* **1** (1930), 192, and **2** (1931), 165.

‡‡ *Proc. K. Akad. v. Wetens. Amsterdam*, **34** (1931), no. 4.

terized by λ constant in space. In that case the emergent intensity in a line is essentially a function of λ. For this reason it is convenient to change the variable from the frequency ν to a linear function of λ, the most convenient variable for this case being probably

$$X = 1-\lambda.$$

It is further convenient to write

$$\eta' = \eta+1$$

and to express in the usual way σ/κ_0 by the number N of atoms above the photosphere by the relation†

$$\sigma/\kappa_0 = N\bar{\sigma},$$

$\bar{\sigma}$ being the coefficient of scattering of a single atom. We then find

$$X = \frac{N\bar{\sigma}}{1+\eta'\bar{\sigma}N}. \tag{91}$$

Whatever the detailed form of X as a function of $\bar{\sigma}$, it must, in general, fall from a sharp maximum at the centre of the line, $\nu = \nu_0$, to practically zero well outside the line. At the centre $N\eta'\bar{\sigma}$ is usually so large in comparison with unity that the central value of X may be taken equal to $1/\eta'$, while the limiting value well outside the line is obviously zero. It is now convenient to assume the intensity normalized in such a way that it is always unity outside a line. The energy lost in the line at frequency ν is then $1-I$, and the total loss of energy in the line, that is, the intensity of the absorption line, is hence given by

$$R = 2\int_0^{1/\eta'} (1-I)\frac{d\nu}{dX}dX.$$

In order to be able to calculate $d\nu/dX$, the functional form of $\bar{\sigma}$ must be specified. Consider first the limiting case when the line is so wide that the influence of the thermal Doppler effect is confined to an insignificant part of the width, and we may write $\bar{\sigma}$ in the form

$$\bar{\sigma} = \frac{Kf\delta}{(\nu_0-\nu)^2},$$

where the symbols have their usual meaning. Introducing this expression in (91) we find

$$\nu_0-\nu = \sqrt{(KNf\delta)}\sqrt{(X^{-1}-\eta')},$$

† A factor of order unity, due to the optical depth in the continuous spectrum, is left out.

and hence $$\frac{d\nu}{dX} = \frac{1}{2X^2}\sqrt{\left(\frac{KNf\delta}{X^{-1}-\eta'}\right)}$$

and the intensity R assumes the form

$$R = \sqrt{(KNf\delta)}\int_0^{1/\eta'} \frac{(1-I)\,dX}{X^2\sqrt{(X^{-1}-\eta')}}.$$

The integral is now a function of η' only, so that R has the form

$$R = \sqrt{(KNf\delta)}F(\eta'),$$

where $F(\eta')$ is some function of η'. For the sun, when the observation is made at a definite point at the solar disk, F will also be a function of the heliocentric latitude.

The above formula states that for lines corresponding to the same value of η' the intensity will be proportional to $\sqrt{(Nf\delta)}$. Unfortunately, in order to be able to decide whether any two lines correspond to the same η' or not, it is necessary to know the profile, since η' is essentially a measure of the true central intensity in the line. If small variations in η' would entail very large variations in F, the above result would therefore be more or less devoid of interest. From the cases in which the profile is known we know, however, that the variations in η' to be expected are so small as to be negligible in a first survey of the field. Consider for instance an approximate representation of the profile by the formula:

$$1-I = X = 1-\lambda.$$

This formula suggests itself as a fairly good approximation by a numerical discussion of the formula in § 46. Using the same expression of $\bar{\sigma}$ as we did above, we find easily

$$R = \pi\sqrt{\left(\frac{KNf\delta}{\eta'}\right)}. \tag{92}$$

The residual intensity in solar lines to which the above formula may be applied will be of the order 10 to 20 per cent., which means that η' will be confined to the interval 1 and 1·25, and hence only influence the total intensity of the line very little. For weak lines, this conclusion must be accepted with some reserve. However, this does not mean very much, since the whole above calculation only holds for rather strong lines.

Consider next what happens when the line shrinks in intensity because of diminishing N, so that the influence of the thermal Doppler

effect first becomes sensible and finally dominant. For an exact analysis it will then be rather important to use the complete expression of the coefficient, in which radiation damping and Doppler effect are simultaneously taken into account. But in order to bring out the typical features of the problem it is sufficient to consider the limiting case when the Doppler effect is dominant.

We thus consider the case when the coefficient of scattering is given by an expression of the form

$$\bar{\sigma} = Kfe^{-q(\nu-\nu_0)^2},$$

where K and q are constants which need not be further specified here. The quantity X then assumes the form

$$X = \frac{KNfe^{-q(\nu-\nu_0)^2}}{1+\eta' KNfe^{-q(\nu-\nu_0)^2}}.$$

In the present case it is not certain that KNf will be large in comparison with unity, and it is therefore not possible to specify $1/\eta'$ as an upper limit for X. For this reason there is no particular advantage to be drawn from a change of variable from ν to X in the integral giving the total intensity. This has the effect that it is not possible to draw general conclusions about the total intensity as was done for broad lines. The only definite fact of this kind which remains is the obvious conclusion that N and f will enter the expression for the intensity as a product Nf, and not occur separately, and further that the intensity will be independent of the damping constant.

If one wants to go farther, it is necessary to consider special cases. A simple case of this sort, which is of considerable importance, is again the case when the absorbed radiation for a definite frequency is proportional to X, which holds fairly well for several solar lines. Then

$$R = \int_0^\infty X\, d\nu,$$

or, writing x for $\sqrt{q}(\nu-\nu_0)$ and substituting for X,

$$R = \frac{1}{\sqrt{q}} \int_{-\infty}^{\infty} \frac{KNfe^{-x^2}\, dx}{1+\eta' KNfe^{-x^2}} = \frac{1}{\eta'\sqrt{q}} \int_0^\infty \frac{2\eta' KNfe^{-x^2}\, dx}{1+\eta' KNfe^{-x^2}}.$$

The lower limit of the first integral should strictly speaking have been $-\sqrt{q}\,\nu_0$, but in all practical cases this quantity will be so large in

comparison with unity that no perceptible error is made by assuming it infinitely large.

For a very weak line $\eta' KNf$ will be small in comparison with unity, and the exponential term may be neglected in the denominator. The expression is then readily integrated, giving an intensity proportional to Nf. When $\eta' KNf$ becomes comparable with unity this law begins to fail. The integrand will then be closely equal to unity, until x becomes of the order $\sqrt{\{\log(KNf\eta')\}}$, and then it drops very rapidly to zero. The value of the integral is therefore approximately

$$R = \frac{1}{\eta'}\sqrt{\left(\frac{\log(KNf\eta')}{q}\right)}.$$

In this case the dependence of R on N is thus much less marked than for very weak or very strong lines. Plotting R as a function of N, we thus find proportionality with Nf when N is sufficiently small. When a certain value of N is reached the curve bends over and tends towards proportionality of R with $\frac{1}{\eta'}\sqrt{\left(\frac{\log(KNf\eta')}{q}\right)}$. For still further increase in N the line widens out so that the line width is no longer determined by thermal Doppler effect, and the curve rises again and approaches asymptotically proportionality with $\sqrt{(KNf\delta)}$. It is important to notice how the dependence of R on the atomic constants changes from one end of the curve to the other.

The preceding considerations serve to show the qualitative features of the intensity curve. On the other hand, Minnaert and Mulders† and Minnaert and Slob‡ have made extensive calculations of the curve connecting total absorption with the concentration of atoms and the oscillatory strengths.

The physical assumptions underlying the two papers in question are somewhat different, and merit separate consideration. In both cases the selective extinction is supposed to be due to radiative damping, with the possible effect of collisions or absorption on the damping term, and the coefficient to be affected by the thermal Doppler effect in the usual manner. The difference between the two papers consists essentially in the assumptions made concerning the coefficient of general absorption and the coefficient of selective absorption and emission in their dependence on density.

In the paper of Minnaert and Mulders theoretical expressions are

† *Zeits. f. Astrophys.* **1** (1930), 192, and **2** (1931), 165.
‡ *Proc. K. Akad. v. Wetens. Amsterdam*, **34** (1931), no. 4.

worked out on the assumptions (1) of an exponential absorption of radiation in the chromosphere, (2) that the intensity is proportional to $1-\lambda$ in our earlier notation, (3) of combined scattering and absorption in the manner first given by Eddington.† The resulting curves brought out very clearly the three different parts discussed above, and qualitatively they did not show great intrinsic difference.

The calculations of Minnaert and Slob, on the other hand, were based on the theoretical expression for the emergent monochromatic intensity derived earlier by Pannekoek, and which was discussed briefly above. The same qualitative course of the curves was established, although some differences showed up in detail.

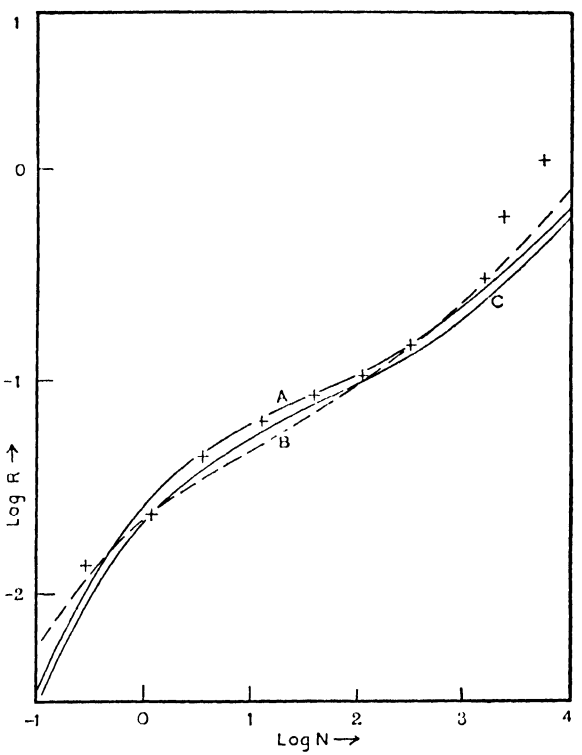


FIG. 13. Total intensity, as a function of the number N of absorbing atoms per cm.^{2} of the sun's photosphere, according to Minnaert and Mulders (*Zeits. f. Astrophys.* **2**). A = exponential absorption; B = combined absorption and scattering (Eddington); C = pure scattering. Crosses represent observations.

† *Monthly Not.* **84** (1929), 620.

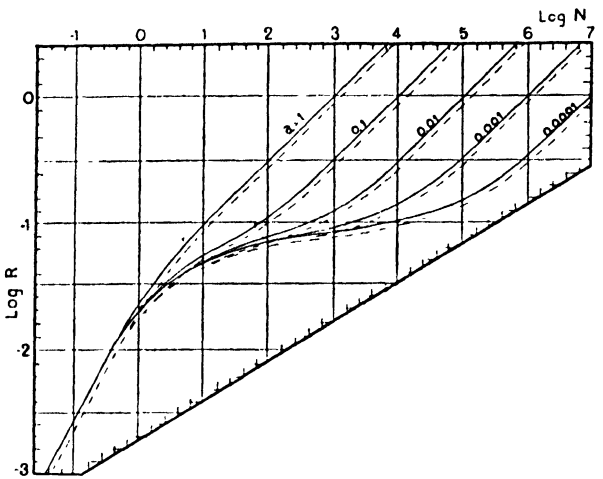


FIG. 14. Total intensity R of Fraunhofer lines as a function of the number of absorbing atoms above the photosphere, according to Minnaert and Slob (*Proc. K. Akad. v. Wetens. Amsterdam*, **34**, no. 4). The parameter a of the curves is given by $a = \delta\lambda\sqrt{(2m/kT)}$, where $4\pi\delta$ is the damping constant as defined in § 33, λ is the wavelength, T temperature, and m the atomic mass. The theory on which the calculations were based is that of Pannekoek (cf. § 45). The curves show well the three regions of the intensity, where R is proportional to N, nearly independent of N, and proportional to $\sqrt{N}$. The importance of the damping constant is also well brought out.

As an illustration to this work we reproduce two diagrams taken from the papers in question.

53. Test by Observational Data

In the paper of Minnaert and Mulders observational material on the intensities of lines in the sun were given and compared with the theoretical curves. It appeared that the observational material illustrated the whole run of the curve, from the square root law for strong lines to the direct linear relation for weak lines. This latter part of the curve is not accessible in ordinary laboratory experiments, and this discloses an interesting case of superiority of astrophysical methods.

It was an interesting outcome of the work of Minnaert and Mulders that it yielded an observational value of the average damping constant of such lines as were strong enough for the square root law to hold. The value found by them was about nine times larger than that corresponding to the classical theory. There is no particular reason for expecting the damping constants to approximate very closely to the classical value corresponding to a harmonic oscillator,

and Minnaert and Mulders tried to show that the damping by transitions from both states of the lines would be sufficient to explain the observations. But some doubt was left over whether there might also be present some effect of collisions.

Pannekoek† made careful observations of the intensities of absorption lines in the spectrum of Deneb from spectrograms obtained at the Dominion Observatory at Victoria, B.C. Again the observations showed a good agreement with the theoretical curves. The damping constant obtained from this analysis came out even larger than for the lines discussed by Minnaert and Mulders. This fact militates against the idea that collisions can play any part in the damping, as Deneb is a typical super-giant, with a correspondingly vastly more tenuous atmosphere. Further progress in this problem would seem to be dependent upon actual calculations of the damping constants from the quantum theory.

54. Analysis of Line Intensities by Adams and Russell

It is clear that if it can be proved that the total intensity of a spectral line is a unique function of the concentration of the particular atom in the atmosphere, and various atomic constants, the analysis of total intensities will provide a new high road to the analysis of stellar atmospheres. As we shall see later, conditions are somewhat more involved, but as a first approach to the problem the assumption of a simple relation between the intensity and the relevant N, f, and δ may be permissible.

This idea was the starting-point of a well-conceived piece of work planned by Adams and Russell‡ before any thorough theoretical work on total intensities had been done. These authors were not aware of the role played by the damping constant, and assumed the intensity to be a unique function of the product Nf. As we have seen above, this assumption holds good only for lines which are so weak that their profile is essentially determined by the thermal Doppler effect. Otherwise this work corresponds to a certain mean value of the damping constant for all lines in use.

The work of Adams and Russell‖ was based on the estimates of solar line intensities given by Rowland. These estimates are on an arbitrary scale, which must be calibrated in terms of the energy units

† *Proc. K. Akad. v. Wetens. Amsterdam*, **34** (1931), no. 6.
‡ *Astrophys. J.* **68** (1928), 279.
‖ H. N. Russell, W. S. Adams, and G. E. Moore, ibid. **68** (1928), 271.

to be used before any conclusions can be based on them. This calibration was made by means of the theory of intensities in multiplets. In the case of a multiplet where all lines have the same initial state, we know that the N's are all alike, and the variation in intensity must be due to the accompanying factor f since the variation in damping constant was ignored. If the multiplet is of the narrow type, it does not matter much whether the initial state is the same for all lines or not, since the variations in N-values must then necessarily be small. Since the solar spectrum contains at least 12,000 lines, it is not difficult to find suitable multiplets in all parts of the spectrum.

This calibration being effected, Adams and Russell went on to a comparative study of line intensities in the sun and in the stars. For this purpose spectra of the sun and the stars were secured by the same instrument and on the same scale, so the spectra could be compared directly in a stereo-comparator. By means of the calibrated intensity scale of the solar lines it was then possible to find the relative abundance of the atomic states in question. Denoting the ratio of corresponding N's in the stars and in the sun by N_r, the excitation energy of the lower state of the line by X_r, and the temperatures of the star and the sun by T'' and T' respectively, we have by Boltzmann's principle

$$\log N_r = \text{const.} + \frac{X_r}{k}\left(\frac{1}{T'} - \frac{1}{T''}\right)$$

to a first approximation. In so far as it is permissible to assume the same value of T' and T'' for all lines in a spectrum this formula demands a linear variation of $\log N_r$ with X_r. This question has been discussed by Adams and Russell on the basis of an extensive material, and it appears that the requirement is in a fairly close correspondence with facts. In particular, since the coefficient of X_r/k is positive or negative according as the temperature of the star is larger or smaller than the temperature of the sun, we should expect $\log N_r$ to increase with increasing X_r in the first case and to decrease in the second. This requirement agrees with the observations.

On the other hand, it is not to be expected that the discussion of Adams and Russell can be trusted when it comes to the finer points of the theory, both because of the assumed universal relation between intensity and the product Nf only, and also because the discussion makes no distinction between principal and subordinate lines. But

the attempt in question is nevertheless very interesting as an important piece of pioneering work in this field.

Russell† has also attempted to use the calibration of the Rowland scale for a large scale determination of abundance in the solar atmosphere. For this purpose it is obviously necessary to know the values of the oscillatory strengths of all lines in question. As this is far from being the case, Russell had to satisfy himself by tabulating the product Nf itself. But even so the results might claim some interest, as the abundance of different elements arrived at in this way showed a distinct parallelism to the abundance of the elements in the earth's crust, apart from a few discordant cases. The main results will be apparent from the diagram given in Fig. 15.

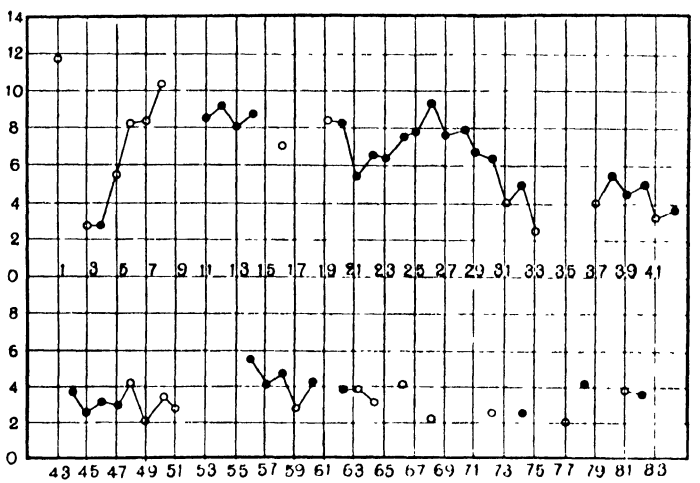


FIG. 15. Abundance of the elements in the solar atmosphere, according to Russell (*Astrophys. J.* **70**). Ordinates are logarithms of the abundance, and abscissae atomic numbers. Reliable results are marked by black dots, and the less trustworthy ones by open rings.

† *Astrophys. J.* **70** (1929), 11. Cf., however, also *Nature*, **135** (1935), 1047, where Russell expresses himself with reserve about conclusions based on the multiplet calibration of Rowland's intensity scale.

XII

THERMAL EXCITATION OF STELLAR ATMOSPHERES

55. Preliminary Considerations

For the interpretation of the main features of stellar spectra it is sufficient to base the discussion on the theory of total intensities developed above. On following the variation of intensity of a single line through different spectral classes it is then sufficient to know that the intensity is some function of N which increases with increasing N. The fact that at low intensities the temperature of the atmosphere will enter the problem also through the thermal Doppler effect can be considered as a second-order effect.

The theory was admittedly applicable only to resonance lines. But in the present case little hesitation may be felt with regard to an extension of it to lines in general, since the main point in question is the fact that a functional dependence of intensity on N exists, and that an increase in N will entail an increase in intensity. In order to obtain a general impression of the intensity of a definite line on passing through the spectral classes, it suffices to derive a theoretical expression of the corresponding N in its dependence on density and temperature.

Before proceeding farther with this problem it is necessary to consider some auxiliary hypotheses concerning the way in which different elements are kept mixed in the atmospheres. In the sun it has been found that the elements are partially separated out by gravity, although far from the extent to be expected from Dalton's law, and also with some conspicuous exceptions, notably calcium and strontium. We have provisionally no corresponding knowledge of the stellar atmospheres, and must consider whether this will introduce marked uncertainties in the result.

Assume first that all elements are evenly intermixed. Denote by N_i the number of free particles of the ith kind per cm.^{2} above the photosphere, and by N the total number in the same volume. A particle of the ith kind has a mass m_i and a relative abundance of r_i. If p_0 is the total pressure in the photospheric level and g the acceleration of gravity, we have

$$p_0 = Nmg, \qquad N = \sum_i N_i, \qquad m = \sum_i r_i m_i.$$

The relative abundance is a constant throughout the atmosphere, and we have

$$N_i = r_i N = \frac{r_i p_0}{mg}. \tag{93}$$

Consider next the case when each element is distributed independently of the other elements. This state is characterized by the fact that each element obeys the hydrostatic law

$$p_i = N_i m_i g,$$

where p_i is the partial pressure of the ith element. Hence follows

$$N_i = \frac{p_i}{m_i g} = \frac{r_i' p_0}{m_i g}, \tag{94}$$

where r_i' is the relative abundance of the element at the photospheric level. Comparing the formulae (93) and (94) we notice that the difference between the two cases consists in the fact that for an even intermixture of elements the average molecular weight m is substituted for the individual molecular weight m_i. Moreover, only the relative abundance in the atmosphere appears in the formulae, while in the case of gravitational separation of the elements the relative abundance in the photospheric level appears explicitly. In this latter case it is therefore possible, besides the average relative abundance given by the ratio of two N-values, also to determine the relative abundance in the photospheric level.

In the following survey of the spectral classes we shall assume the elements to be evenly intermixed. How far it is possible to detect deviations from this state of things will be discussed on several occasions later.

In the theory of solar lines the coefficient of continuous absorption was assumed proportional to density (constant *mass* absorption coefficient). This assumption was introduced for simplicity, mainly, and because it is known to hold under terrestrial conditions, and so forms some sort of transition to stellar cases. It has become clear, however, that the problem of continuous absorption is of far greater importance when it comes to the interpretation of stellar data, where minor changes in the assumed coefficient of general absorption may lead to very different results. The reason for this is clear. The coefficient of general absorption determines the depth of the stellar atmosphere which is accessible to observation from the earth. The smaller the general absorption, the farther we see into the star, and

the greater diversity the atmosphere may be expected to exhibit. The problem of the continuous absorption coefficient is therefore fundamental for stellar theory. To derive an expression for this coefficient on purely theoretical grounds is, however, such a difficult problem, that it is a better procedure to proceed in successive steps. We consider first the problem by assuming the simplest possible expression for the coefficient, and then afterwards try to find out from observations and theory in combination what may be said about the problem. We therefore proceed provisionally with the original formulae.

56. Electrically Neutral Atoms

With the limitations accepted the task of the theory consists only in the calculation of N-values for atoms of different atomic weights and in various stages of ionization and states of excitation. In order to be able to do this some further simplifications must be admitted. First of all, it is necessary to assume the temperature to be constant in the atmosphere. This approximation is the most serious one, since the temperature influences the ionization and the state of excitation through exponential factors. It will also be convenient to work with an average value for the electron density, and to assume the atmosphere to be in hydrostatic equilibrium.

Consider first lines belonging to electrically neutral atoms. It is then usually sufficient to consider the first stage of ionization only, since higher stages of ionization will have a small probability as long as the neutral atoms persist in appreciable quantities.

Let n_0 and n_1 denote the number of neutral and singly ionized atoms per unit volume at a certain level of the atmosphere, and let n_e denote the corresponding electron density. Using the simplified form of the partition function adopted in the corresponding applications to the sun we have

$$\frac{n_1 n_e}{n_0} = \varpi f_e e^{-\chi/kT} \qquad (f_e = 2(2\pi\mu kT)^{\frac{3}{2}}/h^3),$$

where χ denotes the first-stage ionization energy and ϖ the ratio of the weights of the lowest energy stage of the ionized atom to the lowest stage of the neutral atoms. The total density of the atoms is now

$$n = n_0 + n_1 = n_0\left(1 + \frac{f_e \varpi}{n_e} e^{-\chi/kT}\right),$$

or conversely
$$n_0 = \frac{n}{1+\frac{f_e \varpi}{n_e} e^{-\chi/kT}}.$$

Let χ_i denote the energy of ionization of the ith quantum state of the neutral atom and ϖ_i the ratio of the weight of the ith to the lowest state. The number of atoms in this state is then approximately
$$n_i = n_0 \varpi_i e^{-(\chi-\chi_i)/kT},$$
or eliminating n_0 by the preceding equation
$$n_i = \frac{n \varpi_i e^{-(\chi-\chi_i)/kT}}{1+\frac{f_e \varpi}{n_e} e^{-\chi/kT}}.$$

Let z denote the height of the level under consideration above the photospheric level. The N-value for the state in question is then expressed by the integral
$$N_i = \int_0^\infty \frac{\varpi_i n e^{-(\chi-\chi_i)/kT}}{1+\frac{f_e \varpi}{n_e} e^{-\chi/kT}} dz.$$

Let p denote the total pressure at the given level, and r the relative abundance of the given element. The equation of hydrostatic equilibrium is then
$$dp = -\frac{m}{r} g n \, dz. \tag{95}$$

Assuming the temperature constant through the atmosphere, and denoting by $\bar{n}_e$ a certain mean value of the electron density, the above expression of N_i assumes the form
$$N_i = \frac{r p_0 \varpi_i}{mg} e^{-(\chi-\chi_i)/kT} \left(1+\frac{f_e \varpi}{\bar{n}_e} e^{-\chi/kT}\right)^{-1}, \tag{96}$$
where p_0 as usual denotes the total pressure at the photospheric level.

The last problem to solve is now to determine the depth of the photospheric level, so that p_0 may be calculated. It is here that a knowledge of the coefficient of continuous absorption is essential. Adhering to the earlier simple hypothesis we assume this coefficient proportional to the density, which is the same as proportionality with n, since we have assumed all elements evenly intermixed.

We therefore write
$$\kappa_0 = \kappa_0' \frac{n}{r},$$

where κ_0' is a certain constant. Using the equation of hydrostatic equilibrium (95) we find for the optical depth of the photosphere

$$\tau_0 = -\int \kappa_0 \, dz = \kappa_0' p_0/mg, \tag{97}$$

which defines p_0. Let us pause for a moment at this relation. Since τ_0 is obviously the same for the photospheric levels of all stars, the above relation states that p_0 is proportional to g. This result is a direct consequence of the assumed proportionality of κ_0 to n. In the case of actual stars, however, the coefficient κ_0 will be a more complicated function of density and temperature, and the expression for τ_0 accordingly a complicated function of p_0, which also involves the temperature. Solving this equation for p_0 and introducing in (96), the dependence of N_i on temperature is affected, and, what may be of greater consequence, N_i becomes a function of the surface gravity g. In the above simple case g drops out, since N_i depends only on the ratio p_0/g, which also enters the expression of τ_0. Eliminating p_0/g from (96) and (97) we find in fact

$$N_i = \frac{r\tau_0}{\kappa_0'} \varpi_i e^{-(\chi-\chi_i)/kT} \left(1 + \frac{f_e}{\bar{n}_e} \varpi e^{-\chi/kT}\right)^{-1}, \tag{98}$$

which is independent of g. This then holds also for the intensity of the line starting from the ith level. This means that the intensity will be the same for giant and dwarf stars, provided they have the same temperature and relative abundance for the elements. The fact that many lines do exhibit marked differences on passing from dwarfs to giants indicates strongly that the dependence of the coefficient of general absorption on density is more complicated than assumed above. The fact that these differences are on the whole small indicates, on the other hand, that the simpler theory will suffice for a first survey of the spectral classes, when ignoring the splitting-up of the sequence into parallel series.

We shall therefore consider the variation of N_i when the temperature increases from zero to a very large value. In very cold stars the matter is complicated by the fact that many atoms will enter into molecular combinations, so that before an element can make itself felt by the presence of an atomic line, the temperature must have become sufficiently large for the dissociation of the molecular compounds. But for the moment we assume the element to remain in the atomic state throughout the range of temperature under consideration.

57. Intensity of Arc Lines

Take first the case of subordinate lines of neutral elements, like the Balmer lines of hydrogen, for example. From the expression (98) for N_i it follows that with increasing temperature the intensity will increase initially, but by and by more slowly, and finally reach a maximum, after which it will decrease uniformly for still higher temperatures.† The reason for this behaviour is clear enough. That the intensity will be zero in the beginning is due to the fact that there are not sufficient atoms in the excited state. By increasing temperature more and more atoms wander over into excited states, and an increasing intensity results. But finally ionization of the atoms sets in, and this tends to reduce the number of atoms in the right state for the absorption of the line to be possible. When this effect just balances the other, the maximum is reached. The final falling-off of the intensity is due to advancing ionization.

58. Intensity Maxima of the Lines

The preceding considerations were in principle first advanced by the Indian physicist Megh Nad Saha,‡ who in this way became the founder of a rational theory of the spectral classification of the stars. However, the theory of Saha was brought into a much more precise form by Fowler and Milne.‖ These authors contributed to the advance of the theory especially by developing a theory for the determination of stellar temperature by finding out in which spectral class a given line attains its maximum intensity. Saha had earlier tried to find a method of temperature determination by considering the marginal disappearance of the lines. But the conditions making a line disappear are not determined by temperature alone but are also dependent on relative abundance. The determination of the maximum intensity of the line, however, is independent of abundance, provided only it remains the same in all stars under consideration.

Before considering the maximum value of N_i it is necessary to decide how the electron density $\bar{n}_e$ is likely to vary with T. This brings another uncertainty into the calculations. It may be partially met by assuming $\bar{n}_e$ to be proportional to T^m, where m is a constant. It will appear from the final formula that, provided a power law for $\bar{n}_e$ is permissible, the exact value of m will matter but little. In the

† One must remember the dependence of f_e on T.

‡ *Proc. Roy. Soc.* A, **99** (1921), 135.

‖ *Monthly Not.* **83** (1923), 403; **85** (1924), 499.

work of Fowler and Milne the electron pressure was assumed to remain constant through the spectral classes, which corresponds to $m = -1$.

Proceeding in the usual way we find as a condition for the maximum value of N_i

$$\bar{n}_e = \frac{\{\chi_i + (\tfrac{3}{2} - m)kT\}2(2\pi\mu kT)^{\frac{3}{2}}\varpi}{(\chi - \chi_i)h^3} e^{-\chi/kT}.$$

This equation gives the temperature at maximum intensity as a function of ionization energy of the atom and the excitation energy of the line. But the exponential factor dominates to such a degree that the most characteristic feature of the equation is a displacement of the temperature at maximum nearly proportional to the ionization energy.

This is also in good agreement with the facts. At a temperature of about 3,000–4,000° the ionization potential of the elements dominating the spectra is of the order of 6 volts. At a temperature of about 10,000 degrees the ionization potential of the dominant elements is 14 volts (hydrogen), and at twice this temperature the spectrum is dominated by helium, which has an ionization potential of 24 volts.

TABLE 3

Resonance Lines in Stellar Spectra

El.	λ	*El.*	λ	*El.*	λ
Li	6707	V	4333	Zr	4496 ?
Na	5896	Mn	4034		4392
	5890		4033	Mo	3903
Al	3962		4031		3864
	3944	Co	3454		3798
K	7699		3405	Ag	3383
	7665	Ni	3415		3281
Ca	4227	Cu	3274	In	4511
Sc	4247		4348		4102
	3652	Ga	4172	Sn	3262
Ti	5065		4035	Ba	5535
	5040	Rb	7948	La	3949
	5014		7800	Pb	5058
V	4331	Sr	4607		3684

The simplest case is met with in resonance lines. In this case the excitation energy $\chi - \chi_i$ vanishes, and there is no maximum for the lines, which will decrease uniformly in intensity with increasing temperature. This fact forms a simple test of the theory, which corresponds completely to facts. The only notable exceptions

to the rule are due to the fact that the elements in question enter into chemical combinations at low temperatures, so that conditions are intrinsically different from the assumptions underlying the above formulae. All resonance lines known to appear in stellar spectra are collected in Table 3.

59. Intensity of Spark Lines

The run of the intensity of lines due to ionized atoms resembles in a general way that of subordinate lines of neutral elements in that the intensity increases to a certain maximum, and then declines. This holds both for resonance lines and subordinate lines, but the maxima of the latter lines will differ from the other by being flat and ill defined. That this time even the resonance lines must show a maximum is obviously due to the fact that the atoms must first reach the stage when ionization sets in before the lines can appear. The decline in intensity at high temperature is due to second-stage ionization. The same consideration will reappear for lines due to higher stages of ionization, so little further insight is gained by considering such lines in detail. We shall therefore limit the consideration to first-stage ionization only.

It is now necessary also to take into account the atoms having lost two electrons. The number of such atoms per unit volume will be denoted by n_2, and the corresponding partition function by f_2. This time there will be two equations for the equilibrium state, which may be taken to be

$$\frac{n_1 n_e}{n_0} = \frac{f_1 f_e}{f_0} \quad \text{and} \quad \frac{n_2 n_e}{n_1} = \frac{f_2 f_e}{f_1}.$$

Since higher stages of ionization will be neglected, the total particle-density of these atoms will be

$$n = n_0 + n_1 + n_2.$$

Eliminating n_0 and n_2 from these equations we find

$$n_1 = n\{1 + n_e f_0 f_1^{-1} f_e^{-1} + n_e^{-1} f_1^{-1} f_2 f_e\}^{-1}.$$

In exactly the same way as before it follows from this equation that the corresponding N-value for an excited state i of the ionized atom is given by

$$N_i = \frac{\tau_{\tau_0}}{\kappa_0'} \frac{w_i}{f_1} e^{-E_i/kT} (1 + \bar{n}_e f_0 f_1^{-1} f_e^{-1} + \bar{n}_e^{-1} f_1^{-1} f_2 f_e)^{-1},$$

where E_i denotes the energy of the state in question, and w_i the

weight. The two averages of the electron density which figure in this equation are strictly speaking different from each other. But this difference will be ignored here, since the whole theory is only intended to give orders of magnitude.

It would lead us too far afield to discuss a large number of different cases encountered in the theory of spark lines in stellar spectra. We shall here limit the consideration to the case of resonance lines which fall within the part of the spectrum accessible to observation in astronomy. The lines which belong to this class are

$$Ca^+ \ 3934, 3968; \quad Sr^+ \ 4216, 4078; \quad Ba^+ \ 4934, 5354.$$

The fact that these lines are resonance lines in the accessible part of the spectrum entails a separate treatment, because for this to be possible the first excited state of the atom must be situated abnormally near to the ground state, so near in fact that its representative term must be preserved in the corresponding partition function. This fact, which was first clearly pointed out by Fowler,† has the consequence that the rise of these lines to a maximum and their subsequent falling off is not primarily due to second-stage ionization, but much more to excitation to the first excited state. In so far as it is only the question of the calculation of the maximum intensity it is therefore legitimate to neglect the second-stage ionization, that is neglect n_2 and f_2 in our formulae. But it is necessary to amplify the earlier approximate expression of f_1 by adding a second term:

$$f_1 = w_1 e^{-E_1/kT} + w_1' e^{-E_1'/kT},$$

where the second term refers to the first excited state of the ionized atom. The expression of N_1 then assumes the form

$$N_1 = \frac{r\tau_0}{\kappa_0'} \frac{w_1 e^{-E_1/kT}}{(f_1 + \bar{n}_e f_0/f_e)},$$

and the corresponding maximum condition becomes

$$\bar{n}_e = \frac{(E_1' - E_1) w_1' \, 2(2\pi\mu kT)^{\frac{3}{2}}}{\{E_1 - E_0 - (m - \frac{3}{2})kT\} h^3 w_0} e^{-(E_1' - E_0)/kT},$$

on assuming as before that the electron-density varies proportionally to T^m, and using the usual approximate expression for f_0:

$$f_0 = w_0 e^{-E_0/kT}.$$

Provided the denominator of this expression does not become exceptionally small, this relation demands approximate proportionality

† *Monthly Not.* **85** (1925), 970.

of T with $E_1'-E_0$, as was to be expected. The numerical discussion of the formula shows, however, that using the value of $\bar{n}_e$ derived from maxima of ordinary arc lines, the temperatures at maximum are found to be smaller than those actually observed. This fact has been interpreted by Fowler† to mean that these lines originate in high strata of the atmosphere, where the temperature as well as the electron-density are smaller than in the deeper strata involved in the production of the arc lines. This is a special case of a general phenomenon, which manifests itself on several occasions in the analysis of stellar atmospheres, and particularly in the case of the solar chromosphere. This fact gives a first clue to the distribution of temperature and density in the stellar atmospheres.

60. The Rate of Fading-out of the Lines

The preceding considerations have been limited to regions in the vicinity of the intensity maximum. This is the sole reason why it has been possible to take into account only two or three stages of ionization. When considering the falling-off of intensity past maximum conditions are rather different, because then the intensity will depend intimately upon the ease with which the remaining electrons may be removed from the atom. This was first pointed out by Fowler.† The following simple example will serve to illustrate the point.

Assume that the surface group of the atom consists of q electrons. The electrons of the succeeding group we assume to be so firmly bound that ionization of this group plays no part. Denote as usual the number of un-ionized atoms in unit volume by n_0 and the number of atoms having lost all q surface electrons by n_q. The equations of dissociative equilibrium then reduce to

$$n_0 = n_q \frac{f_0}{f_q}\left(\frac{n_e}{f_e}\right)^q.$$

When most of the atoms have lost all q electrons n_q will be approximately proportional to total density or pressure of the substance. To a sufficient approximation the ratio f_0/f_q has the form

$$\frac{f_0}{f_q} = Ae^{\chi/kT},$$

where A is a certain constant and χ the total ionization energy of

† *Monthly Not.* **85** (1925), 977.

the electron group in question. Assuming as usual $\bar{n}_e$ proportional to T^m we have

$$n_0 = \frac{\text{const.} n_q e^{\chi/kT}}{T^{q(\frac{5}{2}-m)}}.$$

Comparing this expression with the corresponding expression for atoms containing a single electron of ionization energy χ it will be realized that the falling-off of intensity after the passage of maximum intensity will become all the more rapid the more electrons there are in the group. This argument holds only as long as $\frac{5}{2}-m$ is positive and does not become very small in comparison with unity.

61. The Electron Density $\bar{n}_e$

Thus far there has been no question of the absolute value of the electron density $\bar{n}_e$. It is of great interest, however, perhaps the essential interest in the case, that the condition of maximum intensity permits a calculation of this quantity when the temperature of the stars in which maxima are found is determined. In Table 4 there is collected a series of such calculated values of $\bar{n}_e$.

TABLE 4

The Electron Density

Element	*Spectrum*	T_e	χ	$\chi-\chi_i$	$\log \bar{n}_e$
Mg^+	*A* 3	9,000	14·97	8·84 volts	12·90
Ca	*Ma*	3,000	6·09	1·88 ,,	11.02
Ti	*K* 2	3,500	6·80	0·85 ,,	12·12
Cr	*K* 5	3,000	8·24	0·96 ,,	10·14
Mn	*K* 0	3,500	7·40	2·31 ,,	11·01
Zn	*G* 0	5,600	9·35	4·04 ,,	11·91
Ca^+	*K* 0	4,000	11·82	0·00 ,,	11·67
Sr^+	*K* 2?	3,500	10·89	0·00 ,,	10·62
Ba^+	*M* 0?	3,000	9·96	0·00 ,,	10·05
Mg	*K* 0?	4,000	7·61	2·69 ,,	11·51

The most impressive feature of this table is probably the fact that the electron-density is so small. Even the largest of the tabulated values is more than a million times smaller than the number of molecules in a cm.^{3} of air at ordinary temperature and pressure. This result was already derived for the case of the sun, but here it appears to be a general feature of all stars. There is some tendency towards a rise of the electron-density on passing towards hotter stars, which perhaps would justify a slightly positive value of m in our equations. But practically speaking the table suggests a constant average value of $\bar{n}_e$, as the fluctuations may just as well be due to

differences between level of arc lines and spark lines, or similar individual properties of the lines.

The order of magnitude derived above for the electron-density, by the method of Fowler and Milne, received independent confirmation from an investigation of Payne and Hogg† on the profiles of the *H*- and *K*-lines in stellar spectra. These lines are so wide that fairly trustworthy profiles may be obtained even from slitless spectra. From such profiles Payne and Hogg derived *N*-values for Ca and Ca^+ independently, and from the ratio of these values they calculated the electron-density in the manner shown in § 51. The result was in fair accord with the value derived from the *H*- and *K*-lines in the sun.

62. Giants and Dwarfs. Difference in Effective Temperature

The above expression of the electron-density at maximum may be used to illustrate the difference in effective temperature between giants and dwarfs of the same spectral class. This difference is in fact essentially due to the difference in atmospheric density. The spectral class of a star is essentially determined by the state of ionization in its atmosphere, and the ionization is again determined by temperature and density. If we pick out a typical element and keep its degree of ionization constant while lowering the temperature, the density must be reduced too. A simple way of bringing this idea to a quantitative form is to consider the relation between electron-density and temperature at maximum intensity of an arc line

$$\bar{n}_e = \frac{\chi_i + (\frac{3}{2} - m)kT}{\chi - \chi_i} f_e \varpi e^{-\chi/kT},$$

which was given earlier. Let the above relation hold for stars in the main sequence, and denote by $\bar{n}'_e, T'$ the electron-density and temperature at maximum for the same line in the giant branch. Since the difference in temperature is relatively small, we may neglect the difference caused by a transition to the giant in the factor linear in T, and only take account of the exponential factor. We then find

$$\frac{\bar{n}'_e}{\bar{n}_e} = e^{\chi(T'-T)/kTT'}.$$

As an example we may consider the neutral iron lines, which reach

† *Harvard Obs. Circ.* **334** (1928). Cf. also O. Struve and C. D. Higgs, *Astrophys. J.* **70** (1929), 131.

their maximum in class $K2$. The effective temperature T and T' may then be taken to be about 5,000° and 4,000° respectively. The ionization energy of iron is $1{\cdot}3.10^{-12}$ erg. With these data we find

$$\frac{\bar{n}_e'}{\bar{n}_e} = 10^{-3},$$

which gives a good idea of the difference in density between giants and dwarfs.

63. Determination of Temperature from Intensity Maxima

When the run of the electron density through the spectral classes has been determined, it is possible, as was first emphasized by Fowler and Milne, to reverse the procedure and to use the condition for maximum intensity as a means of determining stellar temperatures. This method is of special interest in the case of early type stars, where ordinary methods of temperature determination become largely illusory. We give in Table 5 as an illustration some such 'ionization temperatures' of O and B stars.

TABLE 5

Temperature from Intensity Maxima

Element	χ_0	$\chi_0 - \chi_1$	*Spectrum*	*Temperature*
He	24·48	20·88	$B3$	16,000
C^+	24·28	18·0	$B3$	16,000
He^+	54·2	48·2	O	35,000
Si^{++}	31·7	4·8	$B1$–$B2$	18,000
Si^{+++}	45·0	24·0	O	25,000

64. Abundance of Elements in the Stars

The preceding considerations were on the whole independent of the relative abundance of the elements, provided the abundance is the same for all stars, and all elements well mixed at all levels. However, considering the fading out of the line past the intensity maximum, the relative abundance will obviously play a part, and it should be possible from the relative persistence of the lines to determine the relative abundance. This idea was first advanced by Milne, and Payne† made an attempt to determine the relative abundance of the elements on this basis. The result is given in Table 6, together with corresponding figures derived by Russell for the sun.

† *Harvard Obs. Mono.* **1** (1925), 184.

TABLE 6

Relative Abundance of Elements in the Stars and in the Sun

Element	*Log N*		*Element*	*Log N*	
	Star	*Sun*		*Star*	*Sun*
H	12·9	11·5	Ca	6·7	6·7
He	10·2	..	Ti	6·0	5·2
Li	1·9	2·0	V	4·9	5·0
C	6·4	7·4	Cr	5·8	5·7
O	8·0	9·0	Mn	6·5	5·9
Na	7·1	7·2	Fe	6·7	7·2
Mg	7·5	7·8	Zn	6·1	4·9
Al	6·9	6·4	Sr	3·5	3·3
Si	7·5	7·3	Ba	3·0	3·3
K	5·3	6·8			

It appears from these figures that the relative abundance of elements in the different stars and in the sun show great similarities. This result does not preclude considerable differences in abundance between different stars. But the cases in point are rare, and are essentially found among exceptional stars of very high or very low temperatures.

Taking the table on its face value, it brings out forcefully the fact that the most abundant element in the cosmos is hydrogen, with helium coming in a belated second. All other elements are from one hundred thousand to several hundred million times less abundant. Even if very great errors be admitted in these figures, there is no escape from the conclusion that the stellar atmospheres are primarily made from hydrogen, with a moderate admixture of helium, and an insignificant amount of other elements. We shall find several occasions by and by which will serve to bring this fact out in still greater relief.

XIII

THE OPACITY

65. Various Causes of Opacity

THE low value of the electron-density obtained for the reversing layer raises the pertinent question as to the origin of the opacity of a stellar atmosphere. In many stars the number of free electrons in the atmosphere is of the same order of magnitude as the number of atoms, and a low electron-density means therefore a low density of matter. The opacity of a stellar atmosphere must therefore be much higher than that of the terrestrial atmosphere where the mass per cm.^{2} exceeds by many millions the corresponding mass above the photosphere of a star.

The answer to this question is to be sought in differences of chemical constitution and of temperature. A short survey of the absorption of radiation in our own atmosphere will indicate the scope of the problem also for other cases.

Beginning with very long waves in the broadcasting band we know that the main source of interference with the linear propagation of electromagnetic waves is due to the ionization in the Heaviside-Kennelly layer at about 100 km. height, and the successive layers at greater heights. The ionization in question is due to the ionizing action of ultra-violet solar radiation. The reflection of the waves at the lower boundary of the layer demands that the wave-length is large in comparison with the distance between neighbouring electrons. This condition is also necessary for a strong interference to take place, regardless of a reflection, because it conditions the addition of the *amplitudes* of the light scattered from all electrons in a volume of the order of the cube of the wave-length. When the wave-length is small in comparison with the distance between contiguous electrons, the squares of the amplitudes will be added, which demands a very great drop in the scattered intensity. The passage of waves through this layer is also strongly dependent on the density of the air, and the momentary height of the layer. In fact, the air-density determines whether the waves will be actually absorbed, or merely scattered without loss of total intensity. The magnetic field of the earth also plays a part here.

The effect of free electrons on the passage of electromagnetic waves

in the atmosphere appears to cease when the wave-length is shorter than a few metres, which then means that this is the order of magnitude of the distance of neighbouring free electrons in the ionized layer, and it so happens that in this region there will also be approximate equality of the time spent by an electron on a free path, and the period of the wave, so that true absorption is certainly insignificant for shorter wave-lengths. At any rate, there is no reason to suspect that the free electrons in the Heaviside layer will exert any perceptible influence at all on the passage of ordinary light. In the visible region the continuous absorption is essentially due to Rayleigh scattering by the air molecules which necessarily accompanies the dispersion of light. Averaged for all wave-lengths this extinction amounts to about a quarter of a magnitude,† and is thus comparable with the general absorption in a stellar atmosphere. However, the mass of gas per cm.^{2} of the terrestrial surface exceeds by thousands of times the mass overlying the photospheres of most of the stars. It is therefore not possible to explain the opacity of the stellar atmospheres as Rayleigh scattering by molecules or atoms with the same dispersive power as ordinary air, and it looks improbable that the substitution of other elements for air will change the order of magnitude of the result. But the possibility may be kept in mind.

Finally, there is the strong absorption to be considered which sets in at about 2,900 A., and after passing a strong minimum at about 2,100 A. sets in with new force at about 1,800 A. This absorption is due to oxygen, first in the form of O_3, which is responsible for the first band, and then to O_2, which absorbs from 1,800 A. and towards shorter wave-lengths. In both cases there is a mixture of line absorption and continuous absorption, and a dissociation of the molecule is involved. For still shorter wave-lengths a new absorption sets in, which is accompanied by ionization of the molecules, and which is responsible for the maintenance of the Heaviside-Kennelly layer.

Turning next to the analogous problem of the stellar atmospheres, we must expect the same agencies to be responsible for the opacity, due regard being paid to the effect of different chemical constitution, temperature, and density, and the consequent different state of excitation and ionization.

Beginning with the long wave-length end of the spectrum we must

† Cf. Pertner-Exner, *Meteorologische Optik*, 2nd ed., p. 747.

expect that the passage of radiation will be effectively blocked up to a certain limiting wave-length, where the opacity will fall off rapidly. Due to the strong ionization of the stellar atmospheres we must expect this limit to be displaced towards shorter wave-lengths as compared with conditions in the atmosphere of the earth; but it seems *a priori* clear that the difference cannot be large enough for this kind of absorption to dominate altogether. Provisionally we may therefore feel certain that the principal part of the opacity will have a different origin.

The effect of Rayleigh scattering has already been deemed improbable. Since molecular combinations are rare in most stars, there is no reason to expect absorption bands like those due to O_2 and O_3 to play any part in the majority of the stellar atmospheres.

There remains as the only probable source of opacity the absorption beyond the series limit of the atoms, which is usually termed the photoelectric effect.

In our atmosphere this absorption is confined to the far ultra-violet, but this is essentially a temperature effect. Raising the temperature of the gas, more and more atoms will be found in excited states, the absorption of subordinate lines will show up and the continuous absorption at the limits of the subordinate series must become of importance. A single instance of this sort of absorption is directly shown in stellar spectra by the series-limit absorption of hydrogen, which is prominent in early *A* stars.

The continuous absorption responsible for the opacity of the atmospheres in the visible region must be connected with series in the infra-red, like the Paschen series and the Brackett series of hydrogen. The strength of this absorption will depend very much on the abundance and the state of ionization of the elements, and will require separate consideration.

We shall take it for granted that series-limit absorption is the main cause of general absorption, and follow up some consequences of this hypothesis.

On following the intensity of a particular line through the spectral classes, the dependence of the absorption coefficient on frequency does not come in, as the frequency remains the same. The important thing is the dependence on density and temperature. This latter will turn out very different according to the state of ionization and excitation of the active constituent.

The following cases are of special interest:

1. **Absorption by the main constituent in its normal state.** Degree of ionization very small. Then practically every atom present is able to absorb, and the mass absorption coefficient of the medium is practically constant. This is, for instance, the case of the oxygen absorption in our atmosphere in the ultra-violet.

2. **The ionization is so far advanced that the neutral atoms, which are responsible for the absorption, are in minority.** Denote as usual by n_0, n_1, and n_e the numbers of neutral and singly ionized atoms and free electrons per unit volume. The absorption coefficient per atom of the substance is then proportional to $n_0/(n_0+n_1)$. Writing the equation of dissociative equilibrium in question in the form

$$n_0 = \frac{n_1 n_e}{K},$$

where K is a function of the temperature, it appears that the atomic absorption coefficient is proportional to

$$\frac{n_1 n_e}{K(n_0+n_1)}.$$

If the ionization is so far advanced that $n_1/(n_0+n_1)$ does not differ much from unity, it appears further that the atomic absorption coefficient is proportional to the electron density.

The fact that hydrogen is by far the most abundant element in the stellar atmospheres might suggest that this element is responsible for the general absorption. If that is the case, the above considerations must apply, and the mass absorption coefficient must be independent of density for stars of types later than A say, while in A, B, and O stars it must be proportional to the electron density.

3. **As in 2, the ionization is so far advanced that the neutral atoms, which are responsible for the absorption, are in a minority, but the absorbing element has the relative abundance r, where r is small in comparison with unity.** If r is constant through the atmosphere and from star to star, there will be no difference from the preceding case: the mass absorption coefficient will be independent of density for low temperatures and proportional to electron density at high temperatures.

4. **The absorbing atoms are singly ionized, and the element has the**

relative abundance r. The mass absorption coefficient is then proportional to

$$\frac{n_1}{n_0+n_1} = \frac{n_0}{n_0+n_1}\frac{K}{n_e}.$$

For low temperatures, when the neutral state is the most abundant, the absorption coefficient is inversely proportional to the electron density. As the temperature increases, n_1 becomes the most abundant constituent, and the coefficient tends to become independent of temperature. The conditions are thus reversed from the earlier case.

According to this analysis we should expect to find absorption bands in the spectra of the stars, corresponding first of all to the head of the Balmer series, at 3643·39, at the head of the Paschen series at 8205·82, as well as at the heads of numerous series due to metallic atoms. The alkalis, in particular, have well-known series limits in the accessible spectral region.

It has long been known that the absorption at the head of the Balmer series is apparent in the spectra of *A* stars, but this is about the only indication of series-limit *absorption* we know in astrophysics by direct observation. Several investigators have looked for absorption limits in the solar spectrum without success. It is a remarkable fact that an absorption coefficient independent of the wave-length appears to give a good representation of facts for the sun, and the inconsistencies present are at least not such as would be relieved by introducing absorption edges. Minnaert† has searched carefully for traces of absorption edges corresponding to alkali series in the solar spectrum, but without success. The question is not quite easy to decide, as no sharp discontinuity is to be expected but an increasing crowding of the series lines, which then finally merge into a continuum in the region of the theoretical position of the series limit. But still this does not explain the absence of any trace of the limits in the sun.

In order to save the assumption that series-limit absorption plays the dominant role Woolley‡ has suggested that its effect in the continuous spectrum of the sun may be masked by cyclic processes, which in the outermost layers produce a continuous emission tending to fill out the expected drop in intensity.

Although we have seen earlier that cycles of this kind are likely to play a considerable role in the formation of spectra, one may feel

† *Bull. Astron. Inst. Netherlands*, **7** (1934), 234.
‡ *Monthly Not.* **95** (1934), 101.

some hesitation in admitting this as a probable solution of the present difficulty. But Woolley's case is not so weak as it seems. In the nebulae the emission beyond the head of the Balmer series is a prominent feature, and many investigators have called attention to 'ultra-violet appendages' of early type spectra, which indicate that in certain cases there is an emission at the Balmer series limit which more than counterbalances the absorption in the emergent flux.†

66. Theory of the Photoelectric Effect

Although the evidence is a little conflicting we may thus still be entitled to expect that series-limit absorption will prove to be the main cause of continuous absorption in the stars. It is therefore of interest here to consider the principal points of the theory of series-limit absorption, which, moreover, is of interest also in other connexions, especially in the theory of nebulae, where conditions are more easily analysed. It is necessary to limit the theory to hydrogen, or hydrogen-like atoms. But this is no grave limitation, as hydrogen is the most important element in the stellar atmospheres, and most atoms will behave rather similarly in respect to continuous absorption.

Returning to the theory of the hydrogen atom developed earlier, we recall that the wave function of a quantized state of principal quantum number n, subsidiary quantum number l, and magnetic quantum number m may, in polar coordinates, be written down in the form

$$\psi = \frac{e^{-rk_n}\, r^l}{N_r N_\theta N_\phi} F(-n'_r, 2l+2, 2rk_n) P_l^m(\cos\theta) e^{im\phi}.$$

Here F is defined by

$$F(a, b, z) = \frac{\Gamma(b)}{\Gamma(a)} \sum_{\nu=0}^{\infty} \frac{\Gamma(a+\nu)\, z^\nu}{\Gamma(b+\nu)\, \nu!},$$

Γ denoting the gamma function. The quantities N_r, N_θ, N_ϕ are constants of normalization, and k_n depends on the energy E by the relation

$$k_n = \frac{2\pi}{h}\sqrt{(-2\mu E_n)} = \frac{1}{an},$$

where

$$a = \frac{h^2}{4\pi^2 \mu e^2} \frac{1}{Z} = \frac{a_0}{Z},$$

that is, a_0 is the radius of the electronic orbit in the normal state of

† C. S. Yu, *Lick Obs. Bull.* **12** (1926), 104, 155; P. B. Gerasimovic, *Harvard Obs. Circ.* No. 339; H. Kienle, *Monthly Not.* **88** (1928), 700.

hydrogen in Bohr's original theory. Further, n_r' is given by

$$n_r' = n-l-1,$$

and the nuclear charge is assumed to be Ze.

In the present case, when we are considering transitions to and from the continuum, the solutions corresponding to positive values of the energy are also needed. These solutions are derived by exactly the same methods as for the quantized states, and we therefore give the result directly:

$$\psi = \frac{e^{i\kappa r} r^l}{N_r' N_\theta N_\phi} F(-n_r, 2l+2, -2i\kappa r) P_l^m(\cos\theta) e^{im\phi},$$

where

$$\kappa = \frac{2\pi}{h}\sqrt{(2\mu E)}, \qquad n_r = \frac{i}{a\kappa} - l - 1.$$

For the discrete states the factors of normalization are given by

$$N_\phi = \sqrt{(2\pi)}, \qquad N_\theta = \sqrt{\left(\frac{(l+m)!}{(l+\tfrac{1}{2})(l-m)!}\right)},$$

$$N_r = \frac{(2l+1)!\, n^2}{2^{l+1} k_n^l} \sqrt{\left(\frac{a^3(n-l-1)!}{(n+l)!}\right)}.$$

For the continuous wave functions N_ϕ and N_θ are the same as for the discrete, but the radial part will differ. This is not only because of the slightly different expression of the radial wave function in this case, but is due to the process by which the normalization must be carried out when the characteristic functions no longer form a discrete set. The correct procedure for this case was found by Weyl,† and may be summarized as follows:

Assume that a function ϕ may be expanded as a series of characteristic functions ψ_k of the form

$$\phi = \sum_k a_k \psi_k + \int \gamma(E) \psi_E \, dE.$$

We have here assumed the characteristic functions to form a discrete sequence ψ_k and a continuous sequence ψ_E. The continuous parameter E plays the same part as the discrete indices k, and the function $\gamma(E)$ plays the part of the totality of the coefficients a_k. The integral in the above formula may evidently be written as the limit of a sum

$$\int \gamma(E)\psi_E \, dE = \lim_{\Delta E_n = 0} \sum_n \gamma(E_n) \int \psi_E \, dE.$$

The integrals

$$\Delta_n \psi = \int_{\Delta E_n} \psi_E \, dE \tag{99}$$

† *Math. Ann.* **68** (1910), 220; *Gött. Nachr.*, 1910. Cf. also E. Fues, *Ann. d. Physik*, **81** (1926), 281, and M. Stobbe, ibid. **7** (1930), 669.

are called characteristic differentials, and may be directly compared to the characteristic functions ψ_k. They will, in particular, be orthogonal to each other and to the ψ_k, so that the function $\gamma(E)$ is given by

$$\gamma(E) = \frac{1}{\sigma_E} \int \phi \left(\int_{\Delta_n E} \psi_E^* \, dE \right) d\tau,$$

where

$$\sigma_E = \int \left| \int_{\Delta E} \psi_E \, dE \right|^2 d\tau,$$

both integrals being taken over the given ranges of the variable involved.

Since $\gamma(E)\sigma_E$ must be of the order ΔE it is natural to normalize the characteristic differentials by putting $\sigma_E = \Delta E$, or

$$\lim_{\Delta E = 0} \frac{1}{\Delta E} \int \left| \int_{E}^{E+\Delta E} \psi_E \, dE \right|^2 d\tau = 1.$$

It is an important fact that the characteristic differentials *cannot* be written in the form

$$\Delta_n \psi = \psi_E \Delta E$$

as one would be inclined to do. This expression is in fact not equivalent to (99), because $\Delta_n \psi$ is not a uniformly continuous function of ΔE. The case of hydrogen-like atoms, which we are considering, shows this clearly. For very large distances from the nucleus our continuous radial wave function assumes the approximate form

$$\psi_E \sim \frac{1}{r} e^{2\pi i \sqrt{(2\mu E)} r/h},$$

where E is the energy of the electron. It is not necessary to identify the parameter called E above with the energy, but it must be a function of this quantity. The simplest procedure in this case, however, is to use the energy itself. Carrying out the integration involved, it will be seen that $\Delta\psi$ contains a term of the form

$$e^{2\pi i \sqrt{(2\mu E)} \Delta E r/h}.$$

It is here evidently not permissible to expand the exponential in a power series, and to retain only the first power in ΔE, because however small ΔE is taken, it is always possible to find a value of r which is so large that the exponent becomes a finite quantity. This is a general phenomenon affecting all characteristic functions forming a continuous sequence.

The calculations leading up to the explicit expressions of the transition probabilities from discrete to continuous states—or from one continuous state to another—are so complex that we shall refrain from giving details, and only reproduce typical formulae and the necessary reference to the literature.

The normalization factor for the radial wave function in the continuous region, when $\kappa = 2\pi\sqrt{(2\mu E)}/h$ is adopted as a continuous index, is given by†

$$N_r^{-1} = \frac{2^l \kappa^{l+1} e^{\pi/2a\kappa}}{(2l+1)!} \sqrt{\left(\frac{2a\kappa \prod_{s=0}^{l} \left(s^2 + \frac{1}{a^2\kappa^2} \right)}{\sinh(\pi/a\kappa)} \right)}.$$

After very long calculations it is found that the intensity of the radiation emitted per unit time in spontaneous jumps of an electron from a continuous to a discrete state, and corresponding to an interval $\Delta\kappa$ of κ, is given by

$$\frac{dI}{d\kappa}\Delta\kappa = \frac{64\pi^4\nu^4 e^2}{3c^3} l_1 \{C_{n'_r l'}^{n_r l}\}^2 \Delta\kappa.$$

The transition is from a continuous state in the interval $\Delta\kappa$ to a discrete state with quantum numbers n' and l'. The transition takes place only if $l' = l \pm 1$, and the number denoted by l_1 in the above formula is the larger of l or l'. The quantities C are very complicated functions of κ and the quantum numbers. But since they may be useful for reference we give them in full. For transitions $l \to l-1$:

$$C_{n'_r l-1}^{n_r l} = \frac{(-1)^{n'_r} i}{8\kappa^4 (2l-1)!} e^{\frac{1}{a\kappa}\{\frac{1}{2}\pi - 2\arctan\frac{\kappa}{k_n}\}} \left\{ \frac{4\kappa k_n}{k_n^2+\kappa^2} \right\}^{l+1} \times$$

$$\times \sqrt{\left(\frac{(n'+l-1)!\, 2 \prod_{s=1}^{l} \left(s^2 + \frac{1}{a^2\kappa^2} \right)}{(n'-l)! \sinh(\pi/a\kappa)} \right)} u^{n'_r - 1} \times$$

$$\times \{F(-n_r, -n'_r, 2l, 1-1/u^2) - u^2 F(-n_r-2, n'_r, 2l, 1-1/u^2)\};$$

and for transition $l \to l+1$:

$$C_{n'_r l+1}^{n_r l} = \frac{(-1)^{n'_r+1} a}{8\kappa^4 (2l+1)!} e^{\frac{1}{a\kappa}\{\frac{1}{2}\pi - 2\arctan\frac{\kappa}{k_n}\}} \left\{ \frac{4\kappa k_n}{k_n^2+\kappa^2} \right\}^{l+2} \times$$

$$\times \sqrt{\left(\frac{(n'+l+1)!\, 2 \prod_{s=0}^{l} \left(s^2 + \frac{1}{a^2\kappa^2} \right)}{(n'-l-2)! \sinh(\pi/a\kappa)} \right)} u^{n'_r} \times$$

$$\times \{F(-n_r, -n'_r, 2l+2, 1-1/u^2) - u^2 F(-n_r, -n'_r-2, 2l+2, 1-1/u^2)\},$$

† Cf. W. Gordon, *Ann. d. Physik*, **2** (1929), 1031; M. Stobbe, ibid. **7** (1930), 669; H. Bethe, *Handb. d. Physik*, **24**/i (1933), 475; J. A. Gaunt, *Phil. Trans. Roy. Soc.* A, **229** (1930), 163.

where $u = e^{2i\arctan(\kappa/k_n)}$,

and $F(a, b, c, d)$ is the usual hypergeometric function. Proceeding in the usual way, and calculating the absorption coefficients from the emission coefficients via Einstein's relations,† Stobbe finds for the atomic absorption coefficient for absorptions from the ground state

$$\sigma_\nu = \frac{2^7\pi e^2\nu_1^3}{3c\mu\nu^4}\,\frac{e^{-4q\,\mathrm{arccot}\,q}}{1-e^{-2\pi q}},$$

where $$q = \sqrt{\left(\frac{\nu_1}{\nu-\nu_1}\right)},$$

and ν_1 is the frequency of the absorption limit, which for hydrogen corresponds to a wave-length about 912 A. This absorption thus lies far outside the region of stellar spectra.

The absorption coefficient for the first excited state is similarly found to be given by

$$\sigma'_\nu = \frac{2^{10}\pi e^2\nu_2^3}{3\mu c\nu^4}\left(1+3\frac{\nu_2}{\nu}\right)\frac{e^{-8q'\,\mathrm{arccot}\,q'}}{1-e^{-4\pi q'}}$$

for the two sharp terms, and

$$\sigma_\nu = \frac{2^{10}\pi e^2\nu_2^4}{9\mu c\nu^5}\left(3+8\frac{\nu_2}{\nu}\right)\frac{e^{-8q'\,\mathrm{arccot}\,q'}}{1-e^{-4\pi q'}}$$

for the principal terms. Multiplying these expressions by the respective weights 2 and 6, adding and dividing by the total weight of the state, 8, we find for the final absorption coefficient

$$\sigma_\nu = \frac{2^8\pi e^2\nu_2^3}{3\mu c\nu^4}\left(1+6\frac{\nu_2}{\nu}+8\frac{\nu_2^2}{\nu^2}\right)\frac{e^{-8q'\,\mathrm{arccot}\,q'}}{1-e^{-4\pi q'}}.$$

In the above expressions ν_2 stands for the absorption limit of the first excited state, and

$$q' = \sqrt{\left(\frac{\nu_2}{\nu-\nu_2}\right)}.$$

The dependence of the absorption coefficient on frequency is rather complicated. In the vicinity of the absorption limit it is nearly proportional to ν^{-3}; but for increasing values of the frequency it will decrease faster, and finally be proportional to ν^{-7}.

The above series-limit absorption makes itself felt for wave-lengths shorter than 3646 A. and is, as mentioned before, directly observed in *A* and *B* stars. But there is no direct evidence that this absorption is of importance in colder stars.

† Cf. § 8.

The next series-limit absorption in hydrogen is that beyond the Paschen series. The formulae for this case have not been worked out explicitly, however. But it is clear that the absorption coefficient will also in that case fall off nearly proportionally to some negative power of the frequency, which is of the same order as in the two previous cases.

In approximate calculations it is frequently convenient to assume the continuous absorption coefficient to be proportional to ν^{-3}, which gives a sort of average value of the theoretical prediction for hydrogen. The very approximate character of this assumption should be kept in mind, however, especially when applied to atoms other than hydrogen.

An approximate formula for the series-limit absorption in any hydrogen series had been given by Kramers† before the discovery of the new quantum mechanics. Kramers gave the equivalent of the formula

$$\sigma_\nu = g \frac{16e^2}{3\sqrt{(3)}\mu c n} \frac{\nu_n^2}{\nu^3},$$

where n is the principal quantum number of the state, ν_n the frequency of the corresponding series limit, and g a factor of the order unity. This formula has been employed extensively in astrophysical work both in the optical region dealt with in spectroscopic theory, as well as in the X-ray region which is of interest for the theory of the stellar interior. One sometimes has the feeling that the applicability of this formula has been stressed beyond the breaking limit.

67. Differential Effects of Opacity

It was Milne‡ who first showed how the dependence of opacity on density will reflect itself in the intensities of stellar absorption lines. The number of effects to which this test might be applied is legion, as is abundantly clear from Russell's paper on the subject.‡ But it is doubtful whether it pays in the present state of knowledge to go beyond the main effects, since we do not even know definitely which elements are responsible for the general opacity. In this section we shall therefore only discuss the main points of the problem.

From several concurrent sources we know that hydrogen is so

† *Phil. Mag.* **46** (1923), 836.

‡ *Monthly Not.* **89** (1928), 17, 157. See further W. H. McCrea, ibid. **91** (1931), 836; E. A. Milne and S. Chandrasekhar, ibid. **92** (1932), 150; H. N. Russell, *Astrophys. J.* **78** (1933), 239.

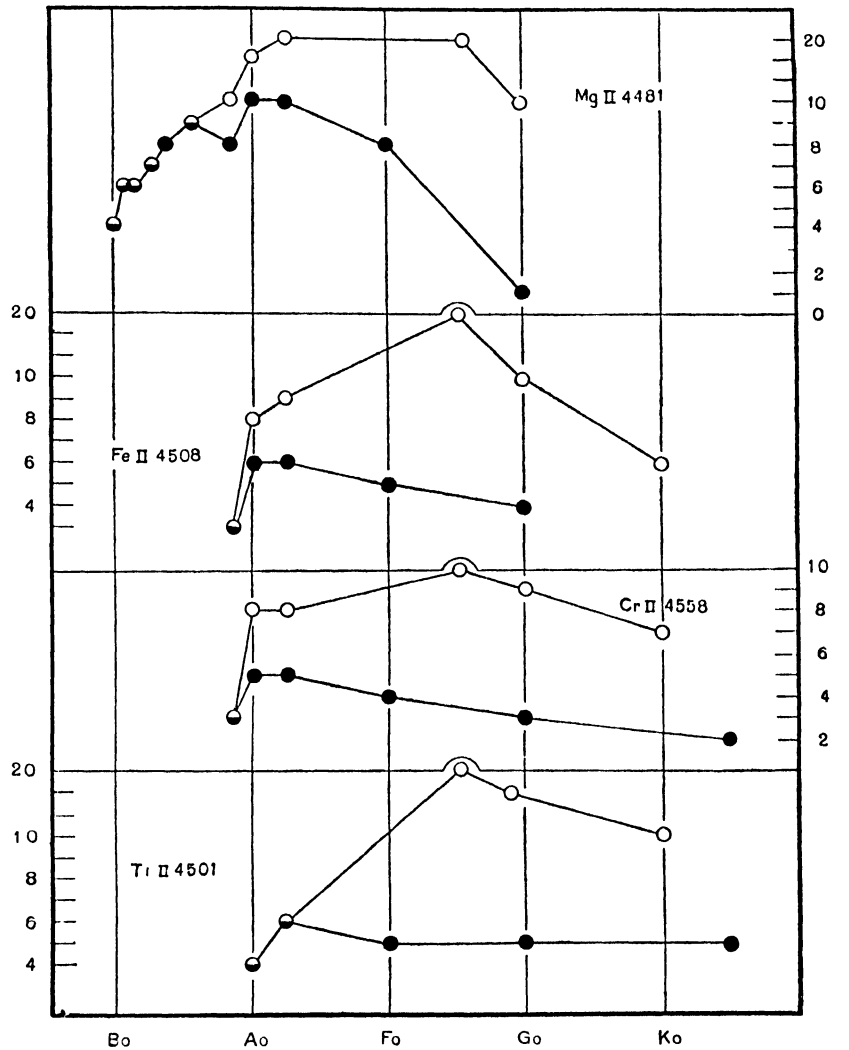


FIG. 16. Intensities in some of the common elements in giants and dwarfs (Morgan, *Astrophys. J.* 77). Ordinates are intensities on the Rowland solar scale. Open circles giants, closed circles dwarfs. The diagram illustrates how lines are affected by change in atmospheric density. Note the change of vertical scale in the case of the largest intensities of the lines (except Cr II, 4558), the highest points corresponding to Rowland intensities 20. These lines, which are absorbed by ionized atoms, behave in a 'normal' way, in that they are strengthened in giants.

abundant that it dominates completely stellar atmospheres. This fact makes it possible to formulate a few simple problems which are likely to show some similarity to actual facts. The closer analysis of

such problems may reveal something about the finer details in the constitution of the atmospheres. But it is scarcely worth while to be very confident in the matter, as it may well be that the observa-

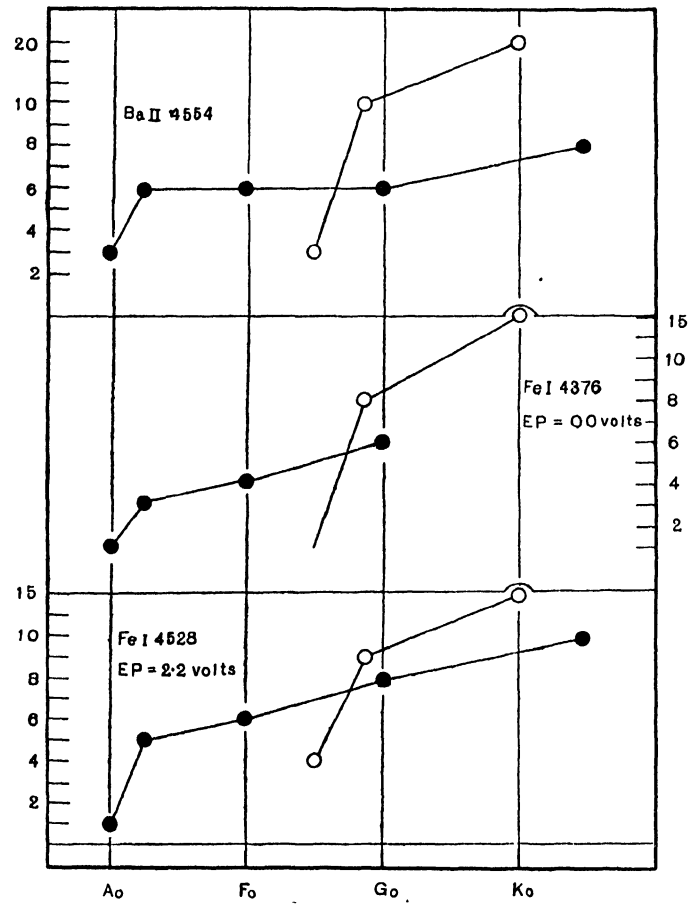


FIG. 17. Line intensities in giants and dwarfs (Morgan, *Astrophys. J.* **77**). The intensities are here seen to behave rather differently from those in Fig. 16, which is due to the fact that their excitation potential (*EP*) is very low.

tional details are due to quite different causes, being connected with the formation of extensive chromospheres, for example, as was mentioned on an earlier occasion.

It is natural to limit the consideration to the following simple problem: An atmosphere consisting essentially of an inert gas (that

is, one showing no perceptible ionization in the range to be contemplated) is mixed up with small fractions of other gases which are more easily ionizable and which have lines in the accessible stellar region.

In most stellar atmospheres there will be a large number of such elements to consider. But since we are mostly interested in the maximum intensities of the lines, and these maxima occur in different temperature ranges for different groups of elements, it is permissible to make some simplifying assumptions. Elements of nearly the same ionization potential may be grouped together, fairly inert elements may be included in the hydrogen content, strongly ionized elements may be collected into a single group. In this way we understand that a tolerable representation of facts may be had by considering only a small number of active elements to be embedded in the inert hydrogen atmosphere.

We assume the atmosphere to be in hydrostatic equilibrium and the elements to be evenly intermixed, so that the hydrostatic equation

$$nmg\,dz = \epsilon\,dp$$

holds for an element of relative abundance ϵ and particle density n. Further, p is the total pressure, g the acceleration of gravity, and m the average atomic mass.

Consider a certain element for which the density of neutral atoms is n_0, of singly ionized atoms n_1, etc., and the corresponding numbers of atoms per cm.^{2} of the photosphere are $N_0, N_1 \ldots$. We have then directly

$$N_0 = \int n_0\,dz = \frac{\epsilon}{mg}\int \frac{n_0}{n}\,dp,$$

$$N_1 = \frac{\epsilon}{mg}\int \frac{n_1}{n}\,dp, \text{ etc.},$$

and the total sum for all stages of ionization becomes

$$N = \epsilon p_0/mg,$$

where p_0 is the pressure at the photospheric level. The optical depth of the photospheric level is

$$\tau_0 = \int \kappa\,dz,$$

and the problem consists in finding appropriate solutions of these equations, together with the equations of dissociative equilibrium for certain typical mixtures of elements and states of ionization of the atmosphere.

Let ϵ and ϵ' be the abundances of the two active elements X and X' and consider first-stage ionization only. We denote the electron pressure by P and write the equation of ionization in the form

$$n_1 = n_0 K/P,$$

where K is a known function of temperature. There may be more elements contributing to the total pressure than the two we are considering, so that $\epsilon+\epsilon'$ need not be unity. But these elements are supposed not to be ionized, so as not to produce free electrons for the electron pressure P. For each of the constituents we have its equation of state:

$$(n_0+n_1)kT = \epsilon(p-P), \qquad (n_0'+n_1')kT = \epsilon'(p-P).$$

Owing to our assumption that only X and X' can produce free electrons contributing to the electron pressure, we have

$$P = kT(n_1+n_1').$$

Finally we have the two equations of dissociative equilibrium

$$n_0 = Pn_1/K, \qquad n_0' = Pn_1'/K'.$$

Eliminating the four quantities n_0, n_0', n_1', n_1 from these five equations, we find the following equation between the total pressure and the electron pressure

$$P = (p-P)\left\{\frac{\epsilon}{1+P/K}+\frac{\epsilon'}{1+P/K'}\right\}.$$

By a simple rearrangement of the terms this equation may be written in the form

$$\frac{p}{P} = 1+\frac{(P+K)(P+K')}{Q(P+R)},$$

where $$Q = \epsilon K+\epsilon' K', \qquad QR = (\epsilon'+\epsilon)KK'.$$

Having thus expressed p by the electron pressure P, we may proceed to perform the integrations giving the N's. The integrations all involve elementary functions only, and we find

$$mgN_0 = \int \frac{n_0}{n}\,dp = \epsilon \int \frac{dp}{1+K/P} = \frac{\epsilon P}{Q}\left\{\frac{P(P+K')}{P+R}+Q-K\right\}+$$
$$+(1-\epsilon)K\log\left(1+\frac{P}{K}\right)-\frac{\epsilon'}{\epsilon+\epsilon'}\frac{R^2}{K}\log\left(1+\frac{P}{R}\right),$$

and similarly

$$mgN_1 = \frac{K\epsilon P}{Q}\left(1+\frac{P+K'}{P+R}\right)-(1-\epsilon)K\log\left(1+\frac{P}{K}\right)+\frac{\epsilon'}{\epsilon+\epsilon'}\frac{R^2}{K}\log\left(1+\frac{P}{R}\right).$$

In both these formulae the P occurring in the right-hand side expression means the electron pressure at the photospheric level; moreover the temperature has been assumed to be constant in the atmosphere.

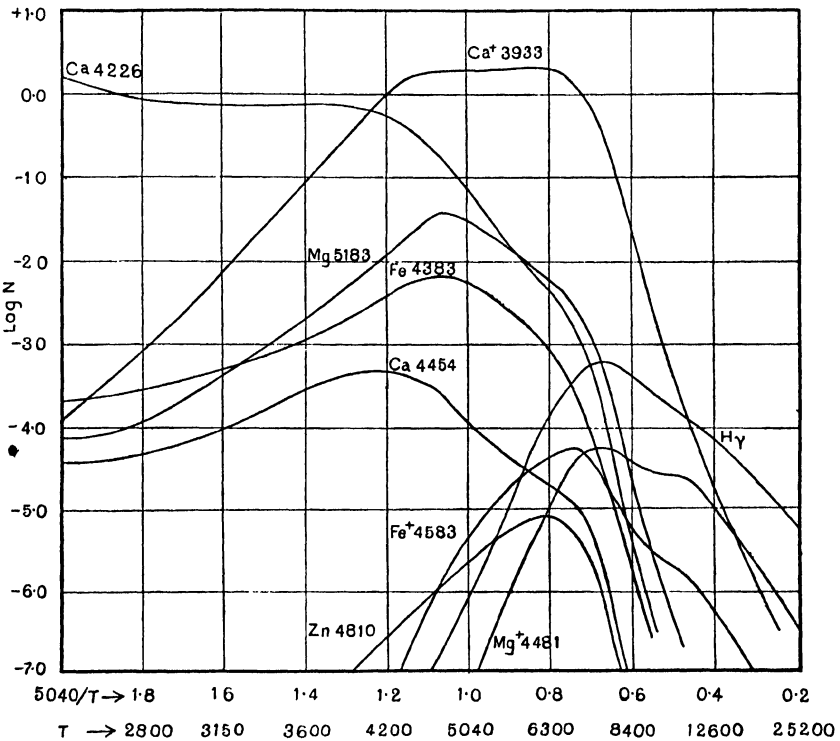


FIG. 18. This diagram gives theoretical line intensities in main sequence stars for some common metals and hydrogen, according to Russell's theory (*Astrophys. J.* **78**). The details in the curves are influenced by the particular assumptions regarding general absorption underlying Russell's work, but the main features would come out similarly on any reasonable thermodynamic theory.

The values of N_0' and N_1' are obtained from the above formulae by interchanging primed and unprimed letters. The extension of the calculations to higher stages of ionization is obvious.

There are here two essentially different possibilities. First we may imagine the inert gas to be responsible for the absorption. The number of inert atoms above the photosphere being

$$N(1-\epsilon-\epsilon') = \frac{p_0}{mg}(1-\epsilon-\epsilon'),$$

the optical depth will then be given by

$$\tau_0 = \kappa(1-\epsilon-\epsilon')\frac{p_0}{mg},$$

where κ is the atomic absorption coefficient.

The other extreme is met with when the inert gas is an inefficient absorber, and the active gases X and X' are responsible for the absorption.

Denoting the absorption coefficient per atom by κ_0 for the neutral and κ_1 for the ionized element, we have, on neglecting second-stage ionization

$$\tau_0 = \kappa_0 N_0 + \kappa_1 N_1 + \kappa_0' N_0' + \kappa_1' N_1'.$$

The above five equations are sufficient for the determination of the five quantities N_0, N_1, N_0', N_1', and P, the other quantities being given functions of the temperature. The above case is really the most fundamental in this theory, and all other cases may be considered as mere variants of it.

Much work has been done by Milne, Chandrasekhar, Payne, and especially Russell in order to analyse the spectroscopic data on the basis of the above equations, or nearly equivalent equations. It appears to be certain that observation favours an absorption coefficient increasing faster with increasing pressure than in the case when the inert constituent is responsible for the absorption. It looks, therefore, as though hydrogen cannot be responsible for the general absorption, at least not in the later types of the spectral sequence. This looks fairly reasonable too, as the energy of excitation of the states in which absorption of the Paschen series and higher series is possible, is so large that a high temperature is necessary to make that type of absorption become effective.

It looks, therefore, as though we are led to assume that the absorption is due to the scarcer constituents, especially the metals. To carry the analysis farther and to say definitely which elements are in each case responsible for the absorption appears to surpass the realm of possibilities at present. Those interested in the subject may consult Russell's voluminous memoir. An attempt by Milne and Chandrasekhar† to test a particular hypothesis about the temperature dependence of the absorption coefficient appears not to be conclusive, because the formula used was smoothed over all wave-lengths. It

† *Monthly Not.* **92** (1932), 150.

is an essential feature of the problem in hand, however, that the intensity of a definite line, of given wave-length, is studied, and this

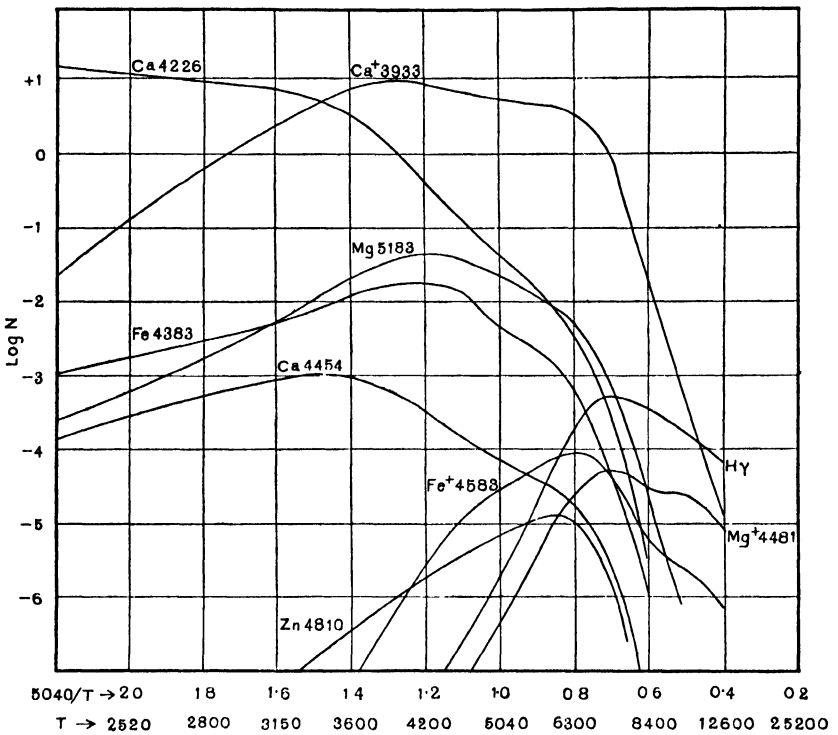


FIG. 19. Same as Fig. 18 except that the calculations are carried out for typical giants, so that the two figures should be studied together. In both cases there is general agreement with observations except in case of hydrogen. The predicted low intensity of H_γ will be apparent in both figures, and also that the maximum intensity falls at a too low temperature. This anomalous behaviour of hydrogen remains unexplained so far.

wave-length is kept constant in all spectral classes. To replace this wave-length by a mean value depending on temperature is therefore not permissible.

XIV

ROTATING STARS

68. Widening of Line Profiles by Rotation

The rotation period of the sun has been measured by spectroscopic methods, particularly by Adams,† from the displacement of absorption lines in the spectrum of the limb. This is the most sensitive method in existence for the measurement of solar rotation.

When applying this same method to the stars‡ conditions become somewhat more complicated, since a stellar spectrum is integrated over the whole disk. In consequence of this all the various displacements of one and the same line appear in a certain sense in juxtaposition. The line profile nevertheless affords some valuable information about the state of rotation of the star, if it rotates sufficiently fast. In the following we shall develop a theory of rotationally broadened line profiles which is suitable for this purpose.

Consider an isolated star in a state of steady rotation, so that the angular velocity **W** is independent of latitude. If the vector **W** is inclined at an angle i with the line of sight, we may decompose it into a component $W\cos i$, parallel to the line of sight, and a second component $W\sin i$, perpendicular to the line of sight. It is only this latter component which manifests itself in the spectrum. We may therefore limit the consideration to stars with axes of rotation perpendicular to the line of sight.

The linear velocity of rotation is the vector product of **W** and the radius vector **R**. The component of the velocity in the line of sight is consequently given by the product of W and the component of the radius vector which is perpendicular to W and to the line of sight. If we determine position on the stellar disk by rectangular coordinates x and y, oriented as in Fig. 20, with **W** along the positive y-axis, the radial velocity at the point x,y is Wx.

Consider a surface element $dxdy$ on the disk, and the intensity of radiation in the frequency interval ν to $\nu+d\nu$ received from this element. In order to find out how this radiation is distributed in

† *Astrophys. J.* **29** (1909), 110.

‡ The first to call attention to the fact that the axial rotation of the stars can be determined from line widths was Captain W. de W. Abney, *Monthly Not.* **37** (1877), 278.

frequencies relative to an observer at rest on the corresponding point of the star, we must transform the spectrum from ν to $\nu' = \nu(1+Wx/c)$ as an argument. Denoting the intensity of frequency ν' at the point x, y by $I(x, y, \nu')$ we obtain the *observed* intensity of frequency ν

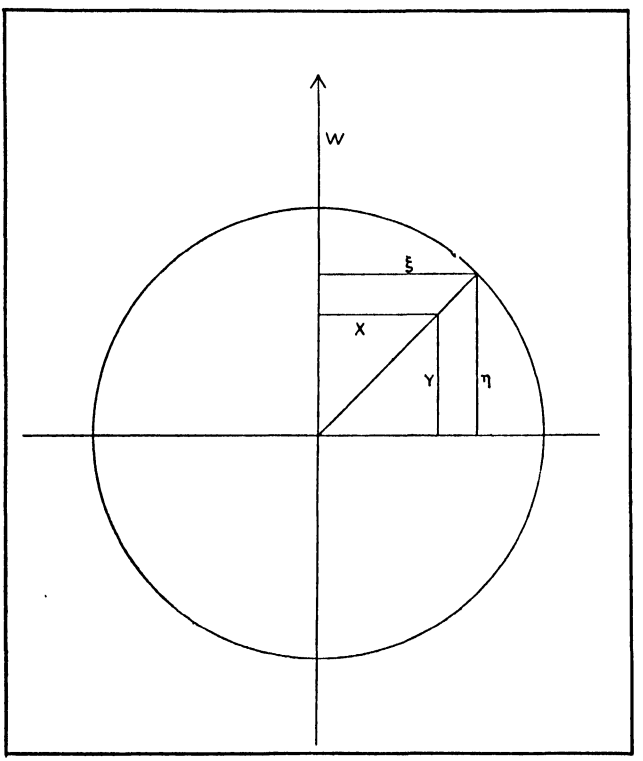


FIG. 20. Diagram showing orientation of coordinates used in the theory of the rotational broadening of spectral lines.

represented by the integral

$$O(\nu) = \iint I(x, y, \nu')\, dxdy,$$

the integral being extended to the whole disk of the star.

69. Intrinsically Narrow Lines

It is of interest to consider first the case of pronounced rotation, in which case the intrinsic width of the line is small compared with that due to Doppler broadening. To bring out the rotation effect we may further neglect the darkening towards the limb, and assume the

angular velocity to be constant and the star to be spherical. Since the intensity I is now independent of y, the integration with respect to y may be performed. We denote the ordinate of the limb corresponding to abscissa x by η, and the intensity $O(\nu)$ assumes the form

$$O(\nu) = 2\int_{-a}^{a} I\left(\nu+\nu\frac{Wx}{c}\right)\eta\, dx,$$

a being the radius of the star.

If the line is sufficiently narrow, the intensity I will vanish except when $\nu+\nu Wx/c$ is in the close vicinity of the proper frequency of the line ν_0. It is then permissible to take η outside the integral sign and integrate, on the express assumption that η will correspond to an abscissa ξ, given by

$$\nu+\nu\frac{W\xi}{c} = \nu_0 \quad \text{or} \quad \xi = c\frac{\nu_0-\nu}{\nu_0 W}.$$

We have written ν_0 in the denominator in the last expression, since $\nu_0-\nu$ is always very small in comparison with ν_0, and this practice will be followed generally. The ordinate η is now given by

$$\eta = \sqrt{(a^2-\xi^2)} = a\sqrt{\left(1-c^2\frac{(\nu_0-\nu)^2}{a^2\nu_0^2 W^2}\right)}.$$

Denoting the intensity of the line by I_0,

$$I_0 = \int_0^{\infty} I\, d\nu,$$

the integral being extended over the line with centre at ν_0, we find

$$O(\nu) = \frac{2c}{\nu_0 W}\eta I_0 = \frac{2cI_0}{\nu_0 W}\sqrt{(a^2-\xi^2)}.$$

The line profile will thus have the form of a half ellipse, and the relative half-width of the line

$$\frac{\nu_0-\nu}{\nu_0} = \frac{aW}{c}$$

gives the equatorial velocity of the star in units of the velocity of light. The result is the same for emission lines and absorption lines.

Let us now assume the star to be darkened towards the limb by the usual factor $1-u+u\cos\theta$, where u is the coefficient of darkening,

and θ the angular distance from the centre of the disk, as viewed from the stellar centre. In our rectangular coordinates we have

$$\cos\theta = \sqrt{\left(1-\frac{x^2+y^2}{a^2}\right)}.$$

In an exact treatment of this case it would be necessary to take account of the fact that the spectral lines also change in character from the centre to the limb. But this may be neglected in the first approximation, and the intensity assumed to be darkened by cutting down the intensity by a factor common to all wave-lengths. Then

$$O(\nu) = \int\int \left\{1-u+u\sqrt{\left(1-\frac{x^2+y^2}{a^2}\right)}\right\} I(\nu')\, dxdy.$$

Just as in the previous case $I(\nu')$ will remain constant during the integration with respect to y. We find

$$O(\nu) = 2\int_{-a}^{a}\left[(1-u)\eta+\frac{\pi}{4}u\frac{\eta^2}{a}\right] I\left(\nu+\nu\frac{Wx}{c}\right) dx, \qquad (100)$$

η denoting as before the ordinate at the limb with x as abscissa.

If the line is sufficiently narrow we may again assume the integrand to vanish except in the close vicinity of $x = \xi$ and we find

$$O(\nu) = \frac{2cI_0}{\nu_0 W}\left\{1-u+\frac{\pi}{4}u\frac{\eta}{a}\right\}\eta,$$

where

$$\eta = \sqrt{(a^2-\xi^2)} = \sqrt{\left\{a^2-\frac{c^2}{W^2}\left(\frac{\nu-\nu_0}{\nu_0}\right)^2\right\}}.$$

With the usual value of $\frac{3}{5}$ for u, the correction may cut down the intensity near the edges of the line by as much as 60 per cent., so that it must always be taken into account in accurate work.

70. Incipient Rotation Effect, and the Determination of the Intrinsic Line Width

Very large rotational velocities are necessarily rare, and it is necessary to consider more closely the case when the rotational broadening is of the same order as the intrinsic width. On the other hand, this intrinsic width cannot be considered as a datum given in all details, and it is of interest to consider also the converse problem of determining the intrinsic line profile from the observed profile

on assuming the shape to be affected by rotation. It is an interesting fact that a formal solution of this problem can be found, which is of great generality, and which has also been put to a practical test. This theory is due to Carroll.†

The problem of finding the intrinsic profile is identical with the problem of solving (100) for $I(\nu')$. In order to bring this integral equation to a more convenient form we write

$$x = at, \qquad \beta = \nu_0 Wa/c,$$

and

$$g(t) = 2a^2\{(1-u)\sqrt{(1-t^2)}+\tfrac{1}{4}\pi u(1-t^2)\}.$$

The equation for $O(\nu)$ then assumes the form

$$O(\nu) = \int_{-1}^{1} I(\nu+\beta t)g(t)\,dt. \tag{101}$$

It will be noticed that this equation bears a certain resemblance to the integral equation of Abel (cf. § 99). The theory of a generalized form of Abel's equation was developed by Mellin,‡ and it happens that the above equation may be transformed to Mellin's form.

Mellin proved that the solution of the integral equation

$$\frac{1}{2\pi i}\int_{C-i\infty}^{C+i\infty} e^{-zx}u(z)\,dz = v(x) \tag{102}$$

is given by

$$u(z) = \int_{0}^{\infty} e^{-zx}v(x)\,dx. \tag{103}$$

The path of integration in (102) is a straight line parallel to the imaginary axis at a distance C from this axis, and C is chosen in such a way that all poles of the functions u and v are situated to the left of the line. The functions u and v are, moreover, supposed to vanish sufficiently fast on approaching infinity.

The transformation of equation (101) into one of the type (102) may be effected as follows: First multiply (101) by $e^{-\nu x}$ and integrate with respect to ν over the contour C. This gives

$$\frac{1}{2\pi i}\int_{C} e^{-\nu x}O(\nu)\,d\nu = \frac{1}{2\pi i}\int_{C}\left[\int_{-1}^{1} I(\nu+\beta t)g(t)\,dt\right]e^{-\nu x}\,d\nu. \tag{104}$$

† *Monthly Not.* **93** (1933), 478.

‡ *Acta Math.* **25** (1902), 139. Cf. particularly § 7.

We here write

$$\frac{1}{2\pi i}\int_C e^{-\nu x}O(\nu)\,d\nu = \phi(x), \tag{105}$$

and

$$\frac{1}{2\pi i}\int_C e^{-\nu x}I(\nu)\,d\nu = \psi(x). \tag{106}$$

From this latter expression it follows that

$$\frac{1}{2\pi i}\int_C e^{-\nu x}I(\nu+\beta t)\,d\nu = e^{\beta x t}\psi(x). \tag{107}$$

Interchanging the order of integrations in (104) we then have

$$\phi(x) = \psi(x)\int_{-1}^{1} e^{\beta x t}g(t)\,dt = \psi(x)G(\beta x), \tag{108}$$

where

$$G(\beta x) = \int_{-1}^{1} e^{\beta x t}g(t)\,dt.$$

Equation (108) defines $\psi(x)$ as a known function of x. It is therefore now possible to employ in (106) the known solution of (102) which, according to (103), is

$$I(\nu) = \int_0^\infty e^{-\nu x}\frac{\phi(x)\,dx}{G(\beta x)}.$$

The right-hand side of this expression contains only known terms, and represents thus the desired solution.

In order to use this solution for the practical analysis of line profiles it is necessary to fit the observed profile on to an analytical expression so as to be able to carry out the integration in the complex plane. Carroll has worked out several ways of doing this in practice.

71. Harmonic Analysis of Observed Profiles

Another way of analysing the observed profile, which also was suggested by Carroll, is to represent $I(x)$ and $O(x)$ by Fourier integrals:

$$F(x) = \frac{1}{\pi}\int_0^\infty d\nu \int_{-\infty}^{\infty} F(t)\cos\nu(t-x)\,dt.$$

For the case in hand it is sufficient to assume both $I(x)$ and $O(x)$ to be symmetric, so that $\cos\nu(t-x)$ may be replaced by $\cos\nu t\cos\nu x$ in the above expression.

Consider for an illustration the case when $O(x)$ is given by equation

(101). We then find for the harmonic amplitude of frequency ν of the observed profile:

$$\int_{-\infty}^{\infty} O(t)\cos\nu t\,dt$$
$$= 2a^2\int_{-\infty}^{\infty} dt\int_{-1}^{1} dx\; I(t+\beta x)[(1-u)\sqrt{(1-x^2)}+\tfrac{1}{4}\pi u(1-x^2)]\cos\nu t.$$

Here we may first write

$$\int_{-\infty}^{\infty} I(t+\beta x)\cos\nu t\,dt = \int_{-\infty}^{\infty} I(t)\cos\nu(t-\beta x)\,dt,$$
$$= \cos\nu\beta x\int_{-\infty}^{\infty} I(t)\cos\nu t\,dt,$$

since $I(t)$ is symmetric. The integral over $O(t)$ is therefore split up into a product of an integral over t and another over x:

$$\int_{-\infty}^{\infty} O(t)\cos\nu t\,dt$$
$$= 2a^2\int_{-\infty}^{\infty} I(t)\cos\nu t\,dt\int_{-1}^{1}\cos\nu\beta x[(1-u)\sqrt{(1-x^2)}+\tfrac{1}{4}\pi u(1-x^2)]\,dx.$$

The latter integral may be evaluated in terms of trigonometric and Bessel functions as follows:

$$\int_{-1}^{1}\cos\nu\beta x[(1-u)\sqrt{(1-x^2)}+\tfrac{1}{4}\pi u(1-x^2)]\,dx$$
$$= \frac{\pi(1-u)}{\nu\beta}J_1(\nu\beta)-\pi u\left[\frac{\cos\nu\beta}{(\nu\beta)^2}-\frac{\sin\nu\beta}{(\nu\beta)^3}\right],$$

since the Bessel function of order unity is defined by

$$J_1(z) = \frac{z}{\pi}\int_{-1}^{1}\cos zx\sqrt{(1-x^2)}\,dx,$$

as is easily seen by a simple transformation of the expression

$$J_n(z) = \frac{2}{\sqrt{\pi}\Gamma(n+\frac{1}{2})}\left(\frac{z}{2}\right)^n\int_0^{\frac{1}{2}\pi}\cos(x\sin\phi)(\cos^{2n}\phi)\,d\phi$$

of the Bessel function of the nth order.†

† Gray and Matthew, *A Treatise on Bessel Functions*, 2nd ed., 1922, p. 45.

We have therefore finally

$$\int_{-\infty}^{\infty} O(t)\cos\nu t\,dt$$

$$= \frac{2\pi a^2}{\nu\beta}\left\{(1-u)J_1(\nu\beta) - \frac{u}{\nu\beta}\cos\nu\beta + \frac{u}{(\nu\beta)^2}\sin\nu\beta\right\}\int_{-\infty}^{\infty} I(t)\cos\nu t\,dt. \tag{109}$$

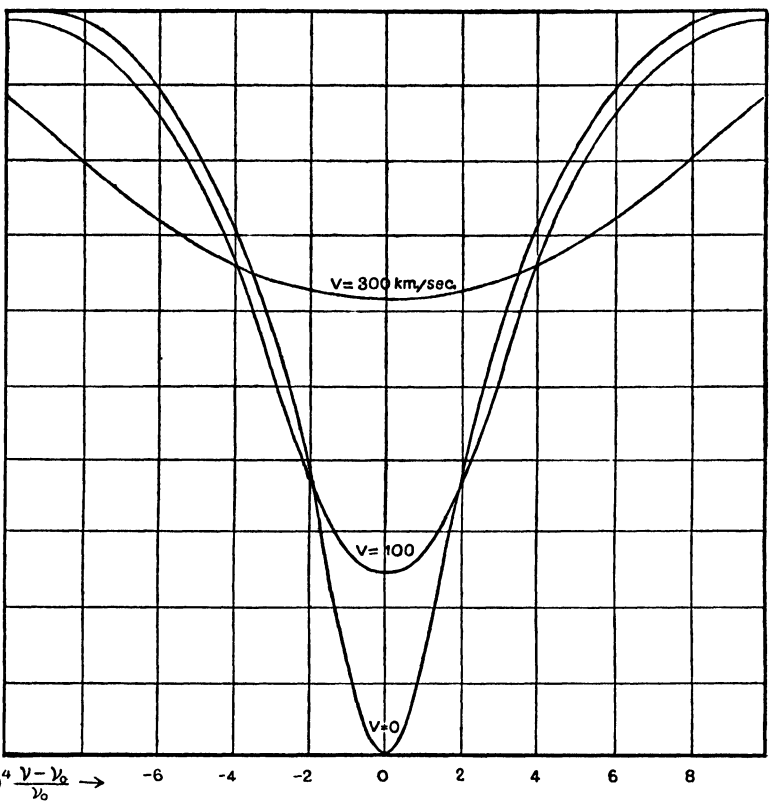


FIG. 21. The original line profile and its distorted form corresponding to rotational speeds of 100 and 300 km./sec. (Carroll, *Monthly Not.* **93**).

The point is here that it has been possible to separate out the part of the right-hand side depending on β, i.e. on the rotational speed of the star. Putting $u = \frac{3}{5}$ this part has a series of zeros, the first four of which are, according to Carroll,

$$\nu\beta = \begin{array}{r} 4{\cdot}147 \\ 7{\cdot}303 \\ 10{\cdot}437 \\ 13{\cdot}569. \end{array}$$

The left-hand side of the equation must have the same series of zeros. These may be found from the observed profile, and the value of β characteristic of the profile, and therefore also the equatorial speed of the star, follows directly. Solving (109) for the harmonic amplitude of the intrinsic profile, a single integration gives the expression of the intrinsic profile itself. The same method is obviously applicable to other laws of darkening than the one assumed here.

In Fig. 21 we reproduce a diagram of profiles computed by Carroll for different rotational speeds. It will be seen that the effect of rotation is to fill out the centre of the line and to widen the wings, the total intensity remaining unchanged.

In the given case the undisturbed central residual intensity was zero. In order to have the same case with a finite initial central residual intensity, one has only to displace the curves vertically by the desired amount.

72. Eclipsing Binaries

The stars in which the rotation effect is to be expected to be most pronounced are the short-period spectroscopic binaries. Among these, again, the eclipsing binaries are the most interesting, because it is then possible to study various parts of the stellar disk during the progression of the eclipse. The configuration of the stars at a given moment of the eclipse is, as a matter of fact, reflected in the form of the spectral lines of the partially eclipsed star. The theory of such lines is slightly different from that given above, the difference consisting in a change in the limits of integration. Little progress has been made by analytic methods in this problem as yet. In Fig. 22 we reproduce a series of theoretical line profiles corresponding to different phases of Y Cygni, which were computed by Shajn and Struve.† In Fig. 24 we give further the observed profile of Mg^+ 4,481 in the spectrum of Algol during partial eclipse, as observed by Struve and Elvey,‡ and also the corresponding theoretical profiles.

Further work in this field may lead to interesting results. Thus Shajn‖ has pointed out that short-period spectroscopic variables may be expected to show a periodic widening of the lines, as a consequence of the ellipticity of the component stars in the direction of the central lines. A theoretical profile for the case of RR Cen given by Shajn is

† *Monthly Not.* **89** (1929), 222.
‡ *Monthly Not.* **91** (1930), 663.
‖ *Poulkovo Obs. Circ.* **7** (1933).

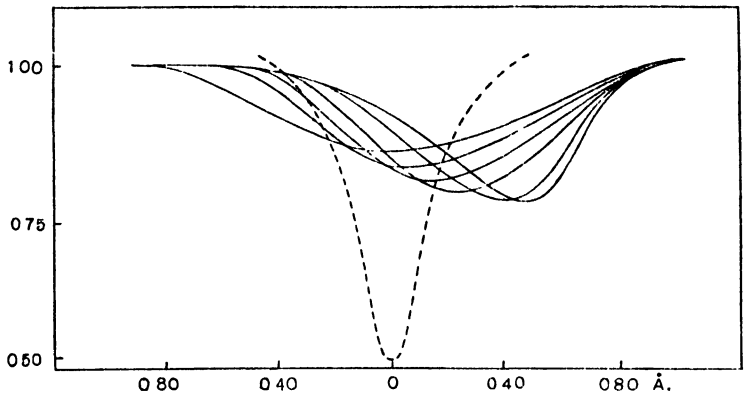


FIG. 22. Line contours at different phases during eclipse, for Y Cygni, computed by Shajn and Struve (*Monthly Not.* **89**, 222). This diagram has not been, and cannot well be, tested observationally, since the spectrum of the fainter component of Y Cygni is not sufficiently weak. But the diagram illustrates well the progressive asymmetry of the lines in the spectrum of a partially eclipsed star.

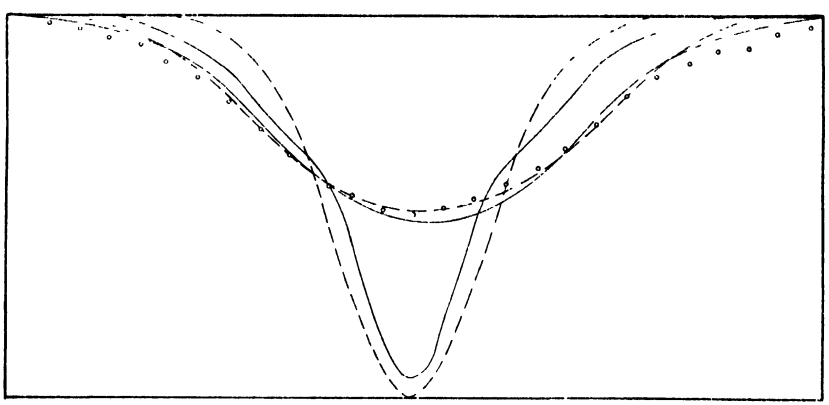


FIG. 23. Profile of the magnesium line Mg + 4,481 in Algol. Open circles represent observed intensities by Struve and Elvey (*Monthly Not.* **91**). Dotted lines represent the intrinsic profile assumed by Struve and Elvey, and the corresponding profile when distorted by a rotation with equatorial velocity $V = 60$ km./sec. Full lines represent an independent analysis by Carroll and Ingram (*Monthly Not.* **93**). First a rotational velocity of $V = 51$ km./sec. of the star was derived by harmonic analysis of the observed profile, and then the original intrinsic profile is calculated from the Fourier integrals. Darkening towards the limb is neglected in this work.

reproduced in Fig. 25. But thus far this effect does not seem to have been established with certainty by observations.

73. Observational Results

The first observations indicating a rotation of a star were made by Schlesinger,† who also measured the displacement of the lines in the

† *Pub. Allegheny Obs.* **1** (1909), 134; *Monthly Not.* **71** (1911), 719.

eclipsing variable δ Lib and λ Tau during the partial eclipse. This so-called limb-effect has been studied by a number of observers.†

The first extensive study of rotationally broadened lines in general is that of Shajn and Struve, which has been mentioned on several occasions above.

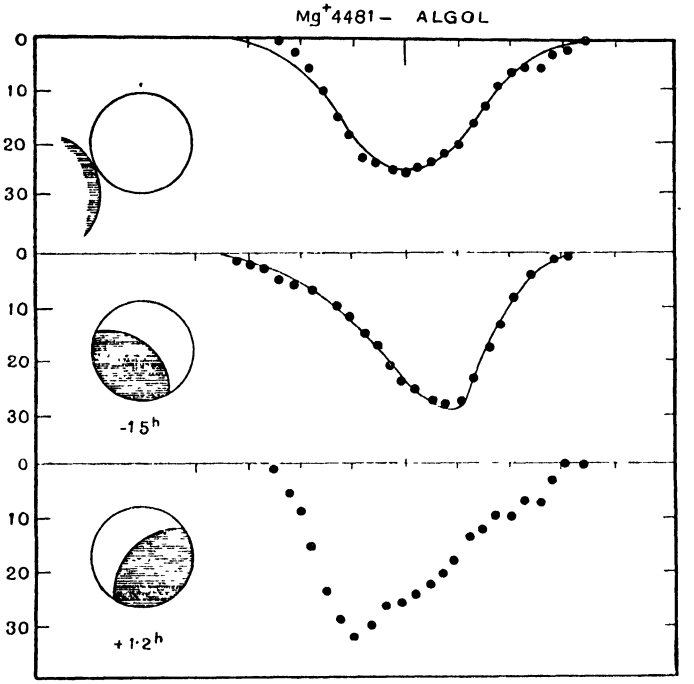


FIG. 24. Profiles of Mg + 4,481 in Algol during eclipse. Filled circles are observations, full lines theoretical contours (Struve and Elvey, *Monthly Not.* **91**).

In 1930 Elvey‡ published an extensive study of the Mg II line 4,481 with the view to detection of rotation. The theoretical profiles for different rotational speeds were determined graphically, and the result compared with observed profiles in 59 stars of spectral classes *O*, *B*, *A*, *F*. Elvey derived for these stars an average equatorial speed of 60 km./sec. In this figure account is taken of a random

† J. Hellerich, *Astron. Nach.* **216** (1922), 277. Rotational effects in β Lyr and Algol were studied by R. A. Rossiter, *Astrophys. J.* **60** (1924), 15, and by Dean B. McLaughlin, ibid. **60** (1924), 22. See further J. S. Plaskett, on the rotation of 21 Cas, *Pub. Dom. Astrophys. Obs. Victoria*, **3** (1926), 247, and Dean B. McLaughlin, on λ Tau, *Pop. Astron.* **34** (1926), 624.

‡ *Astrophys. J.* **71** (1926), 221.

inclination of the axes of rotation, but no account is taken of darkening towards the limb. Taking this into account, the result would probably have been higher speed.

Observed and computed profiles agreed generally well, especially when the rotation was rapid, so that all profiles assumed the elliptical ('dish-shaped') form demanded by the theory. In most stars it is to be expected that *all* lines of an element will be widened, and the width proportional to wave-length.

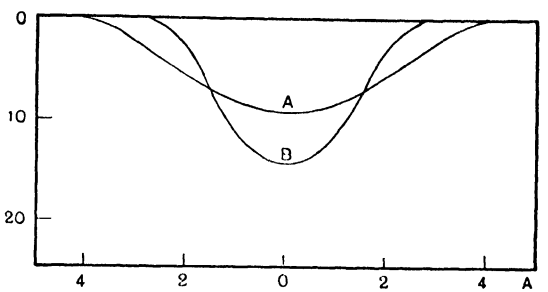


FIG. 25. Hypothetical change in line profiles of RR Cen, as a consequence of the ellipticity of the primary component (Shajn, *Poulkovo Obs. Circ.* 7). Profile *B* corresponds to an end-on view of the star, while *A* corresponds to a broadside view.

A certain amount of latitude must be allowed when comparing lines of different elements. Narrow lines may occur in rotationally widened spectra, as a result of absorption in a surrounding nebulosity, or in interstellar gas.

The theory of Carroll was applied to some of Elvey's profiles by Carroll and Miss Ingram.† The stars studied were 44 ι Ori, 44 ξ Per, 39 λ Ori, and Algol. Their study of Algol is interesting, as it may be checked by the value for the rotational speed which follows from the orbit and the dimensions of the stars, or from the measurement of the limb-effect, or from the detailed study of the rotation effect by Struve and Elvey.‡ There is in all cases a fairly close agreement, which fixes the rotational speed of the bright component of Algol at about 55 km./sec. An inspection of Fig. 23 will show the broadened and intrinsic profiles as given by Struve on one hand, and Carroll and Miss Ingram on the other.

A general survey of stellar spectra shows, according to Struve‖

† *Monthly Not.* **93** (1933), 608.
‡ Ibid. **91** (1931), 663.
‖ *Astrophys. J.* **72** (1930), 1.

and Miss Westgate,† that the *A* stars have the greatest velocities of rotation. The *B* and *F* stars show rotation quite frequently, but appreciably less than *A* stars. There is, according to Miss Westgate, no definite correlation between absolute magnitude and rotation, and

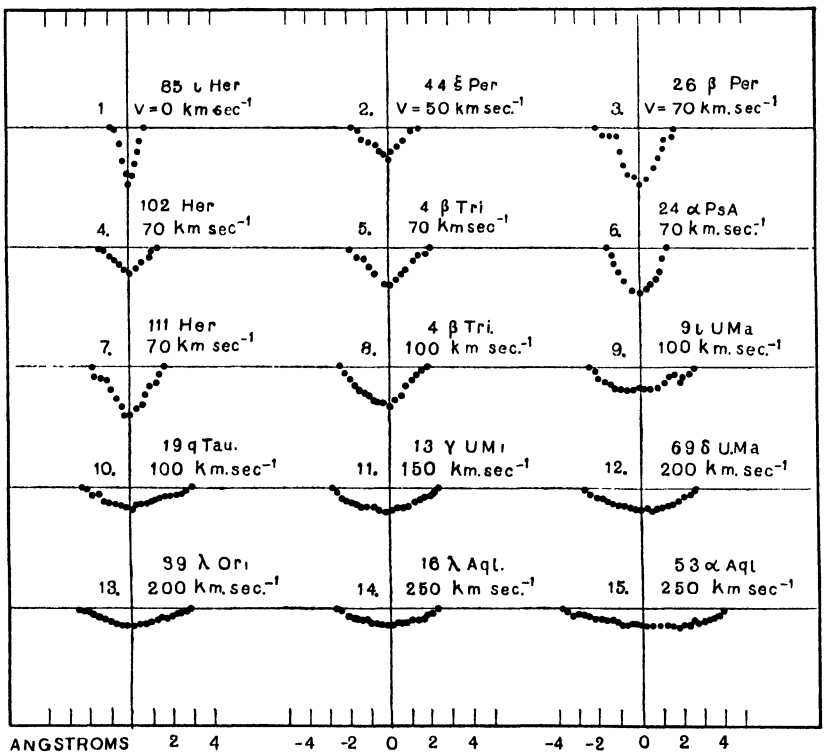


FIG. 26. Profiles of the Mg^+ line 4,481 in various stars, showing broadening by rotation. The equatorial velocity given for each line is not corrected for the inclination of the axis of rotation to the plane of the sky (Elvey, *Astrophys. J.* **71**).

the average rotational speed appears to be independent of galactic latitude, which indicates a random distribution of the direction of the axes.

For the later classes, *G*, *K*, and *M*, no case of rapid rotation appears to have been observed in single stars.

Some cases of very high rotational velocities may be noted. Thus in α Vir the brighter component rotates with an equatorial velocity of 200 km./sec. The fainter component appears to have an equatorial velocity of but 50 km./sec. Since in a close pair like this there is

† *Astrophys. J.* **77** (1933), 141; **79** (1934), 355.

every reason to believe that both components rotate with about the same angular velocity, we must conclude that the brighter component is four times larger than the fainter one.

The lines in the spectrum of Benetnasch (η U Ma) are wide and dish-shaped, typical of the kind to be expected for rapidly rotating stars. They indicate according to Struve an equatorial velocity of 200 km./sec.

Although these velocities are a hundred times larger than those of the sun, the stars are still within the limits of stability for a homogeneous mass rotating with a constant angular velocity. How much this means is difficult to say at present, as there is reason to believe that stars in general will rotate faster in the interior than on the outside, which indicates less stability than this for actual stars. On the other hand, the increase of density inwards in the star must tend to make the stars more stable.

XV

THE EFFECT OF MAGNETIC AND ELECTRIC FIELDS ON ABSORPTION LINES

74. The Zeeman Effect

In the course of the development of solar physics the influence of a magnetic field on spectral lines, the Zeeman effect, has become of importance. In 1896 Zeeman discovered that lines in series spectra are affected when the light source is situated in a magnetic field, or, in the case of absorption lines, when the absorbing substance is placed in a magnetic field. By longitudinal observations, i.e. by looking in a direction parallel to the magnetic field, the lines appear in the simplest case resolved into a doublet. By transverse observations the lines appear triple, the middle component appearing in the position of the original line, and the two satellites coinciding with the longitudinal doublet, and at a distance from the middle component of

$$\Delta\nu = \pm\frac{e}{\mu}\frac{H}{4\pi c} = 4{\cdot}70\,.\,10^{-5}H\ \text{cm.}^{-1},$$

where e and μ are charge and mass of the electron, and H the magnetic field-strength. This is the so-called normal Zeeman effect. In most cases, however, a much more complicated type of magnetic multiplicity appears, which then is termed the 'anomalous' Zeeman effect. The terminology is exactly the opposite of what it ought to have been, but has now gathered too much historical momentum to be discarded lightly.

The magnetic components of a line have a definite type of polarization. The undisplaced component is linearly polarized in the direction of the magnetic field, and the satellites are circularly polarized in opposite directions perpendicularly to the field. It is thus sufficient to observe the Zeeman pattern of a line in one direction, and it is then possible to predict the pattern for any other direction.

Inspired by Zeeman's discovery Lorentz gave a theoretical interpretation of the normal Zeeman effect, on the basis of his electron theory. A formal quantitative theory of the anomalous Zeeman effect was first possible after the discovery of wave mechanics, and the introduction of the electronic spin.

Consider a single electron moving in a field of axial symmetry,

on which is superposed a magnetic field parallel to the axis of symmetry and of strength H. Introduce cylindrical coordinates z, r, ϕ, where z is parallel to the axis of symmetry and r perpendicular to this axis, while ϕ determines the position of a plane through r and z. The vector potential of this field is simply

$$A = \tfrac{1}{2}Hr$$

in the direction of ϕ (i.e. perpendicularly to z and r), so that the operator $\mathbf{A}\nabla$ reduces to $\frac{1}{2}H(\partial/\partial\phi)$. The wave equation of the problem becomes consequently

$$\frac{h^2}{8\pi^2\mu}\nabla^2\psi - \left(\frac{e^2H^2r^2}{8\mu c^2} + V\right)\psi + \frac{h}{2\pi i}\left(\frac{e}{2\mu c}H\frac{\partial\psi}{\partial\phi} - \frac{\partial\psi}{\partial t}\right) = 0.$$

Neglecting the square of H in this equation, it is possible to prove that the problem with magnetic field is equivalent to the corresponding problem without a field by the following transformation of coordinates†

$$\phi' = \phi + \frac{eH}{2\mu c}t; \quad t' = t; \quad r' = r; \quad z' = z.$$

This transformation means simply that ϕ' is substituted for ϕ in the Laplacean operator, that r and z are left unaltered, and that the last parenthesis in the equation reduces to $-\partial\psi/\partial t'$. If the term H^2 may be neglected, the equation has been reduced to the case of no field by the transformation.

Now, since the Laplacean operator in cylindrical coordinates has the form

$$\nabla^2 = \frac{\partial^2}{\partial z^2} + \frac{1}{r}\frac{\partial}{\partial r}\left(r\frac{\partial}{\partial r}\right) + \frac{1}{r^2}\frac{\partial^2}{\partial\phi^2},$$

and the potential function V is independent of ϕ (because the field has axial symmetry), it is clear that the solution of the field-free equation will have the form

$$\psi_0 = u(r', z')e^{-2\pi iEt'/h - mi\phi'},$$

where m is a positive or negative integer, the magnetic quantum number encountered earlier, and E the energy of the system. Transforming back to our original coordinates we find

$$\psi = u(r, z)e^{-\frac{2\pi i}{h}\left(E + \frac{mheH}{4\pi\mu c}\right)t - mi\phi}.$$

† O. Klein, *Zeits. f. Physik*, **41** (1927), 407.

The magnetic contribution to the energy is thus

$$\Delta E = mh\frac{eH}{4\pi\mu c}.$$

The appearance of m in this place justifies the appellation 'magnetic' quantum number. By the selection principle for m proved earlier ($m' = m \pm 1, m$) it follows that each line splits up into three components, one in the place of the undisturbed line, with a satellite on each side, displaced a distance $\Delta\nu = \pm eH/4\pi\mu c$ from the middle component, in agreement with Lorentz' theory. That the intensities of the components also come out in agreement with the observed polarization rules is easily verified from the expressions of the harmonic amplitudes given earlier.

In the theory of the anomalous Zeeman effect the electronic spin plays a principal part. The formula for the additional magnetic energy was given earlier (p. 57). Its derivation involves calculations too lengthy to be reproduced here.

75. Zeeman Effect in the Sun

The discovery of magnetic fields in the sun was made by Hale,† on June 25th 1908 at the Mount Wilson Observatory. It had already been noted by Young in 1892 that absorption lines in sun-spots are widened and sometimes distinctly double. Hale succeeded in showing that the components of such doublets showed the right sort of polarization in order to be due to the Zeeman effect. This effect then provides a means of measuring the intensity of the magnetic field on the sun. In large sun-spots the intensity runs up to 4,000 gauss. The proper interpretation of these fields, which is still an open question, must be sought in the interior motion of the spot.

The discovery of magnetic fields in sun-spots naturally led to a search for a general magnetic field on the sun, analogous to the magnetic field on earth. This search was crowned with success in 1913,‡ when it was announced from the Mount Wilson Observatory that the sun had been found to have a general magnetic field, with an axis inclined about 5° to the axis of rotation, and with the poles oriented in the same sense as those of the earth's magnetic field, the magnetic field intensity being directed from north to south. The field strength appeared to decrease rapidly with height.

† *Astrophys. J.* **28** (1908), 315. ‡ G. E. Hale, ibid. **38** (1913), 27.

This latter fact shows† conclusively that the solar magnetic field should not be compared to the field of an elementary magnet in all respects, as is freely done in the literature. It is inadmissible to calculate the magnetic field strength *parallel* to the axis from the observed component perpendicular to this axis by using the formulae for an elementary magnet. In this way it has been stated that the general magnetic field strength in the lower solar strata is of the order 50 gauss. This, however, is not the observed value, but only one calculated from observations on using erroneous formulae. The correct field strength must be about three times larger, as we shall see.

The observations are all concerned with the component of the magnetic field intensity in the line of sight,‡ which in the case of the sun means practically the component perpendicular to the magnetic axis. Assume then that the observations are right about the component perpendicular to the axis, so that we may write for the rectangular components of H parallel to the plane of the magnetic equator:

$$H_x = xzA(r); \quad H_y = yzA(r); \quad r^2 = x^2+y^2+z^2.$$

We have further the equation

$$\operatorname{div} \mathbf{H} = 0,$$

from which H_z may be calculated:

$$H_z = \int_z^\infty \left(\frac{\partial H_x}{\partial x} + \frac{\partial H_y}{\partial y}\right) dz;$$

or introducing for H_x and H_y,

$$H_z = -(x^2+y^2)A(r)+2\int_z^\infty A(r)z\, dz.$$

The last term is negligible. For assuming an exponential drop of field strength with height, say

$$A(r) = A_0 e^{-(r-a)/h},$$

a being the radius of the sun and h a constant of the order of the height of the solar atmosphere, say $h = a.10^{-3}$, we find

$$H_z = -(x^2+y^2)A(r)+2(hr+h^2)A(r).$$

The last term is clearly negligible, except in the close vicinity of the poles. Comparing this expression with that of an elementary magnet,

$$H_z = -\tfrac{1}{3}(x^2+y^2)A(r)+\tfrac{2}{3}z^2A(r),$$

† S. Rosseland, *Astrophys. J.* **62** (1925), 387.

‡ Cf. for instance F. H. Seares, *The Observatory* **43** (1920), 311, equation (3).

it follows that, as regards order of magnitude, the true z-component must be about three times larger than that corresponding to an elementary magnet, the H_x and H_y components of which are given the observed values. The value given at Mount Wilson for the horizontal component of the field in the equatorial regions, about 50 gauss, should therefore be increased three times to 150 gauss. Moreover, the resultant field strength is given by

$$H = r^2 \cos\theta\, A(r),$$

where θ is the heliocentric latitude.

A field of this type requires a steady current system in order to be maintained. From Maxwell's first equation we find the current density to be

$$\mathbf{J} = \frac{c}{4\pi} \operatorname{curl} \mathbf{H} = \frac{\mathbf{V}}{\omega} \frac{c}{4\pi r^4} \frac{\partial}{\partial r}(r^5 A),$$

where $\mathbf{V}$ and ω are the linear and angular velocities of the atmosphere at the point in question. The conditions are thus as though the current were produced by a uniform rotation of a spherically symmetrical distribution of electricity. The dominant term in the expression of the current is negative, and it follows that the electron current flows in the direction of rotation of the sun.

76. The Stark Effect

In 1913 the German physicist Stark found that spectral lines emitted in a homogeneous electric field are resolved into a system of components. Hydrogen, and atoms like He II which are hydrogen-like, are much more sensitive to the Stark effect than most other atoms. The separation of the Stark components of a line is in this case proportional to the applied electric field. The separation increases rapidly with increasing principal number of the line.

The quantum theory of the Stark effect was first given by Schwarzschild and Epstein in 1916. The refinements in this theory introduced by wave mechanics concern only second-order effects so far as hydrogen-like atoms are concerned. The expression for the additional energy of a state of principal quantum number n is, according to these theories,

$$\Delta E = \frac{3h^2 F}{8\pi^2 \mu e Z} nn' \qquad (n' = 0, 1, \ldots, n-1),$$

where n' is a new quantum number, which distinguishes the different sub-levels of the state, F the electric field intensity, and Ze the

charge of the nucleus. The Stark components are polarized according to definite rules, which all come out correctly from theory.

While the magnetic splitting of the spectral lines in sun-spot spectra is quite plain, the corresponding splitting of the lines by electric fields has thus far not been observed in the sun or the stars. The reason for this lies probably in the strong ionization of stellar material. A static electric field cannot well exist without the presence of insulating material. A rapidly changing magnetic field would, on the other hand, produce an electric field by induction. But thus far no such field has been discovered.

The strong ionization of stellar material introduces another kind of electric field—the interatomic field produced by ions and free electrons. This field may attain quite large values. But as it changes rapidly in time and space it cannot produce the same effects as a static field. It is perhaps permissible to think of this field as producing a Stark effect of the atomic lines in any given very short time interval. But owing to the rapid variation of the field the observed lines will be a superposition of a continuous set of Stark effect patterns, the net result being a line broadened in a definite manner. This type of broadening has been studied in the laboratory, and its theory has been considered by various authors, notably Holtsmark.† According to Holtsmark's theory the half-width of an originally very narrow line broadened widely out by ionic forces is given by

$$\Delta\nu = 3{\cdot}26Ae n^{\frac{2}{3}},$$

where n is the density of broadening ions or electrons of charge e, and A is the half-width of the ordinary Stark effect in a field of unit intensity.

It has become customary in astrophysical literature to reserve the name 'Stark effect' for this type of line broadening, which is not unreasonable since there is little chance that a true Stark effect will ever be observed in the stars.

It is clear that the effect, if at all observable, must be looked for first of all in the hydrogen lines, and much work has been spent on this problem, first of all by Struve and Elvey,‡ and by Unsöld.|| It seems to be a settled matter that an interionic broadening of the higher members of the Balmer series is definitely present in the sun and the stars.|| Struve†† also believes that he has found the Stark

† *Ann. d. Physik*, **58** (1919), 577.
‡ *Astrophys. J.* **72** (1930), 277.
|| *Zeits. f. Astrophys.* **2** (1931), 199.
†† *Astrophys. J.* **70** (1929), 85.

effect for the diffuse series in helium. This was done by selecting a series of spectra in which lines of C II and Si III, which are known to be insensitive to Stark effect, were perfectly sharp. Struve found then that in several of these spectra the diffuse helium lines were perceptibly broader and more diffuse than the other helium lines, and also that the forbidden helium line 4,470 A. appeared in these spectra. There was a definite correlation between the width of the diffuse helium lines, and the intensity of 4,470 A. and the width of the Balmer lines, indicating that the same cause was at work.

Since dwarfs have denser atmospheres than giants, it is clear that Stark effect broadening must show up first in dwarfs, and thus form a criterion of absolute magnitude. The white dwarfs are particularly interesting in this connexion, as was recently emphasized by Pannekoek and Verwey.† A single instance mentioned by these authors may serve to show the importance of this point. Some years ago Öhman‡ published observations of the spectrum of the white dwarf star 'van Maanen 1166' in the *h* Persei region. On Öhman's tracings the hydrogen lines have widths of the order of 100 to 200 A. for H_γ. For 40 Eridani B, another recognized white dwarf, the widths are less, but still of the order of 100 A. These widths alone are sufficient to establish the white-dwarf character of these stars. It is to be expected that this criterion will find interesting applications in the search for white dwarfs in the future.

† *Proc. K. Akad. v. Wetens. Amsterdam*, **38** (1935), No. 5.
‡ *Monthly Not.* **92** (1922), 71.

XVI

MOLECULAR COMBINATIONS IN STARS

77. Spectra of Molecular Compounds

THE tendency to form molecules is only met with in the atmospheres of the cooler stars, especially in the late-type spectral classes. The difficulty with which molecular compounds are formed is evidenced by the fact that in all recognized cases it is only a question of diatomic molecules. The identification of more complex compounds has been announced from time to time, but such identifications have in all cases later been recognized as erroneous.

There are several interesting problems associated with the occurrence of molecular compounds in stellar atmospheres, which are of general importance in astronomy. First of all the bands offer an alternative method of determining the temperature of the reversing layers of the stars, and by two independent methods, since we may either study the distribution of intensity within a single band, or the relative intensity of different bands belonging to the same sequence. These two alternatives will then involve different Boltzmann factors, but refer to the same temperature. Next there is the problem of the chemical equilibrium of a mixture of elements capable of forming chemical compounds in stellar atmospheres. Since this problem involves temperature and density, it gives another clue to these quantities. Thirdly it may be mentioned that the opinion has been advanced from time to time that the band absorption exerts a marked influence on the variation in luminosity of long-period variables.† How far this is so still remains to be decided, but it adds, at any rate, additional interest to the problem of chemical combinations in stars.

In the following we shall give a brief account of the theory of the spectra of diatomic molecules in so far as this is of interest for astrophysics. Next we shall consider in due course the various astrophysical problems in which the theory of molecules plays a definite part.

The motion of a molecule may usually with a fair approximation be split up into three parts, which are only slightly interdependent. Thus the motion of the electrons in the molecule will in the main be

† A. Joy, *Astrophys. J.* **63** (1926), 281; *The Internal Constitution of the Stars*, by Sir Arthur Eddington, p. 207.

the same as if the nuclei were held rigidly apart and at rest. Superposed on this electronic motion, as a kind of perturbation, comes the oscillations of the nuclei with respect to one another. Finally, there is the rotation of the molecule as a whole to be considered. This takes place like that of a rigid body, very nearly, the axis of rotation being either fixed in space or performing a slow precession about an invariable axis.

Each of these component motions has a definite counterpart in the structure of molecular spectra. Rotation alone produces narrow bands which in the case of absence of vibrational or electronic excitation are situated in the far infra-red end of the spectrum, one for each substance. Setting a rotating molecule into vibrations has the effect of shifting the bands a good way towards the visible region of the spectrum, and sometimes also of splitting a single band into a sequence of bands. Thirdly, excitation also of electronic motions produces a shift of bands into the visible or ultra-violet part of the spectrum, and the splitting up of the bands into long sequences, due to the interference of the electronic motion with the oscillations and the rotation. At the same time the band structure is definitely deformed by the formation of band heads. These are due to the change of moment of inertia during an electronic transition.

78. Rotatory Oscillation Band

Adopting a simple rotator of moment of inertia A as a model of the molecule, the energy in different rotational states is given (p. 43) by

$$E = \frac{h^2}{8\pi^2 A} j(j+1) \qquad (j = 0, 1, 2, \ldots),$$

while the wave functions will be associated Legendre functions of order m and degree j, m being the magnetic quantum number, which is always smaller than j and larger than $-j$, or equal to these limits.

The possible transitions are limited by the selection principle that j can only change by ± 1 in a jump, while m may change by ± 1 or zero. The characteristic frequencies of the radiation emitted or absorbed by the molecule will therefore be given by the formula

$$\nu = \frac{h}{4\pi^2 A}(j+1) \qquad (j = 0, 1, 2, \ldots).$$

The spectrum will thus consist of a series of equidistant lines, the separation of adjacent lines in frequency units being $h/4\pi^2 A$. By

measuring this separation the moment of inertia of the molecule may be found, and hence, for diatomic molecules, the distance between the two atoms. Spectra of this kind have been found in water vapour (Rubens, Eva v. Bahr) and hydrogen chloride (Czerny), both in the infra-red part of the spectrum.

The next simplest case is met with when the molecule besides a rotation also performs oscillations, while the electrons still remain in their normal state. For small vibrations the motion may be considered purely harmonic and independent of the rotation, so that the characteristic frequencies of the radiation emitted or absorbed during a simultaneous jump in rotational as well as vibrational energy is given by

$$\nu = \nu_0(l-l') \pm \frac{h}{4\pi^2 A}(j+1) \qquad (l, l' = 0, 1, 2, \ldots), \qquad (110)$$

where ν_0 is the characteristic frequency of vibration. This is always found to be large in comparison with the term $h/4\pi^2 A$. For this reason a spontaneous jump of a molecule must always take place from a higher to a lower state of vibrational energy, while the rotational energy may either increase or decrease in the jump according to circumstances. Hence the double sign in (110). This duplicity has the effect that a spectrum of the rotatory-vibrational type will consist of two wings of lines with nearly constant frequency differences, and symmetrical with respect to the frequency ν_0. For a strictly harmonic vibrator all jumps go to adjacent states, so that $l-l'$ should always equal unity. The intra-molecular oscillations, however, are never strictly harmonic, so that jumps in several units of the vibrational quantum number are not excluded. Mostly only a single vibratory band is observed for each substance, but this means only that under ordinary conditions of excitation of these spectra the molecules are practically all in the lowest vibrational state.

79. Electronic Bands

As a last step we consider simultaneous jumps in rotational, vibratory, and electronic excitation energy of the molecule. For this case the simple formulae used above are markedly inadequate, and it becomes essential to take account of the fact that the moment of inertia of the molecule and the frequencies of vibration vary systematically from one state to another. This must not be confounded with the change in A with varying state of vibration and rotation which turns up in vibratory-rotation bands. While this

latter effect may be present even for a molecule containing an excited electron, it is masked by the much bigger effect of the excited electrons themselves. In fact, the conditions are such that a molecule containing an excited electron has a bigger or smaller moment of inertia than the same molecule with the electron unexcited, *no matter which are the states of rotation or vibration involved*, at least within very wide limits. We shall therefore distinguish between A_i, the moment of inertia of the molecule in the initial state, and A_f, the moment of inertia in the final state, the state being assumed to correspond to different electronic excitations.

The change in rotational energy in a jump will now be

$$\Delta E_{\text{rot}} = \frac{\hbar^2}{8\pi^2}\left\{\left(\frac{1}{A_i}-\frac{1}{A_f}\right)j^2+\left(\frac{1}{A_i}-\frac{1\pm 2}{A_f}\right)j-\frac{1\pm 1}{A_f}\right\} \quad \begin{pmatrix}+: j\to j+1\\ -: j\to j-1\end{pmatrix},$$

it being assumed that j changes by $+1$ or -1 in the jump. The coefficient of j^2 preserves the same sign for all lines of the band, while the coefficient of j may either be positive or negative, giving rise to the two branches of a band.

Let us introduce two new quantities A and δ instead of A_i and A_f, defined by

$$A = \sqrt{(A_i A_f)}, \quad \text{and} \quad \delta = \frac{A_f - A_i}{\sqrt{(A_i A_f)}}.$$

The coefficient of j^2 in the above formula then assumes the form

$$\frac{1}{A_i}-\frac{1}{A_f} = \frac{\delta}{A},$$

and the coefficient of j becomes, correct to the first order of δ,

$$\frac{1}{A_i}-\frac{1\pm 2}{A_f} = \begin{cases}\frac{2}{A}(\delta-1) & (j\to j+1)\\ \frac{2}{A} & (j\to j-1).\end{cases}$$

Finally, we have for the additional term:

$$-\frac{1\pm 1}{A_f} = \begin{cases}-\frac{1}{A}(2-\delta)+\ldots & (j\to j+1)\\ 0 & (j\to j-1).\end{cases}$$

The quantity δ measures the relative increase of nuclear distance during the electronic transition in question, and is to a first approximation just twice this value. In the linear and constant terms of ΔE_{rot}, δ only plays the part of a small correction, and may be left out of consideration. In the quadratic term, on the other hand, it

is of predominant importance. We therefore write ΔE_{rot} in the simplified form

$$\Delta E_{\text{rot}} = \frac{h^2}{8\pi^2 A}\{\delta j^2 \mp 2j - (1 \pm 1)\} \quad (j \to j \pm 1).$$

Due to the term proportional to j^2 in this expression, the two wings of the band will no longer be symmetrical with respect to a given frequency, as was the case with the infra-red bands of oscillatory-rotatory character. According as δ is positive or negative, the negative or positive wing of the band will double back upon itself and run in the direction of the other wing. This is the physical interpretation of the *band head*, which is the most characteristic feature of visible bands. Assume, for instance, that δ is positive. The frequencies of the negative branch will then be given by the formula

$$\nu = \nu_r + \frac{h}{8\pi^2 A}(\delta j^2 - 2j - 2), \tag{111}$$

where ν_r summarizes the terms which are independent of j. The frequency difference between consecutive lines will be

$$\Delta\nu = -\frac{h}{4\pi^2 A}\{1 - \delta(j + \tfrac{1}{2})\}, \tag{112}$$

which shows that the head is reached when j is of the order δ^{-1}. It is important to realize that the head thus marks nothing but the region of maximum proximity of adjacent lines in the branch which doubles back upon itself, and its exact position gives no definite clue to the structure of the band. For the theoretical disentangling of observed bands the important question is the determination of the frequency ν_r, since this gives the zero point for the reckoning of the j-numbers, which further enables us to determine the moment of inertia of the molecule.

In addition to the positive and negative branches of a band, there occurs also, in the case of visible bands, a so-called zero branch, corresponding to no change in j during the transition, the change in rotational energy being solely due to the change in moment of inertia. Hence for this case

$$\Delta E_{\text{rot}} = \frac{\delta h^2}{8\pi^2 A} j(j+1).$$

Zero branches might also be expected in the case of pure rotation or vibratory rotation spectra, but they seem to be absent, although methane (CH_4) possibly provides an exception.

A decrease in the moment of inertia means evidently a tightening-up of the molecule and an increased strength of the chemical bond between the atoms. This means again greater stability, and the vibrational frequencies should increase in the process. This inference is confirmed by the empirical rule that for electronic bands the oscillatory frequencies depend to a marked degree upon the vibrational quantum numbers, and in such a way that the sequence of bands is degraded in the same direction as the individual bands of the sequence.

80. Multiplet Structure of Bands

In the preceding section we have assumed the molecule to behave rotationally like a spinning top without precession of its axis of rotation. This amounts to the assumption that there is no appreciable coupling between the motion of the electrons of the molecule and the rotation of the nuclei, which will hold, approximately, if the electrons have no resultant spin or orbital angular momentum, particularly in the direction of the line joining the two nuclei. In the consideration of cases when this condition no longer holds it is advantageous to distinguish between two typical cases, and to consider other cases as belonging to either of these types.

Case (*a*) The first type we define as follows: The influence of the resultant spin of the molecule on the spectral terms is large in comparison with the influence of the rotation. This means that the coupling between the spin and the orbital angular momentum of the electrons is so strong that it is not appreciably influenced by the rotation. The component of the total angular momentum of the electrons (orbital+spin) along the central line of the nuclei will now be a constant of motion and assume the quantized values $h\Omega/2\pi$, where Ω may assume the integral values $0, 1, 2, \ldots$. The total angular momentum will be similarly quantized, its square assuming the value $h^2J(J+1)/4\pi^2$, where J may assume the values $\Omega, \Omega+1, \Omega+2, \ldots$. The angular momentum of the rotation appears as the vectorial difference between these quantities, and the wave-mechanical theory† gives for the square the expression

$$|M|^2 = \frac{h^2}{4\pi^2}\{J(J+1)-\Omega^2\}.$$

† Cf. R. de L. Kronig, *Band Spectra and Molecular Structure*, Cambridge 1930.

Different values of Ω will distinguish between different fine-structure components of the band. The rotational energy becomes

$$E_{\text{rot}} = \frac{h^2}{8\pi^2 A}\{J(J+1)-\Omega^2\}.$$

Case (*b*) It may happen, on the other hand, that the influence of the spin is small in comparison with that of the rotation, but that the motion of the electrons is still not very much influenced by the rotation. Neglecting the spin provisionally, it will be the component of the orbital motion of the electrons alone along the central line of the nuclei which will assume quantized values, which we this time denote by $h\Lambda/2\pi$. This component will combine vectorially with the angular momentum of rotation to a resultant K, the square of which has the value $h^2K(K+1)/4\pi^2$, and the rotational energy becomes

$$E_{\text{rot}} = \frac{h^2}{8\pi^2 A}\{K(K+1)-\Lambda^2\},$$

the allowed values of K being $\Lambda, \Lambda+1, \Lambda+2\ldots$. Denote the spin component of the molecule by S, and the total angular momentum by J as before. The vectors K and S will then combine vectorially to form J.

The quantum number Λ plays a similar part for molecules as the subsidiary quantum number for atoms, and in order to bring out this parallelism farther it is customary to denote the spectral terms corresponding to the values $0, 1, 2,\ldots$, etc., of Λ by the Greek capital letters $\Sigma, \Pi, \Delta,\ldots$, corresponding to the $S, P, D,\ldots$ terms of atomic spectra. For transitions $\Sigma \to \Sigma$ the selection rule $K \mp 1 \to K'$ holds. The consideration of the spin gives finally a splitting-up of the rotational terms according to the scheme

$$J = |K+S|, |K+S-1|, |K+S-2|,\ldots, |K-S|,$$

the total multiplicity being $2S+1$, which usually is written as a superscript on the upper left side of the Greek letter in question. Thus ${}^2\Sigma$ means that $S = \frac{1}{2}$ and ${}^3\Sigma$ that $S = 1$.

Inasmuch as all relevant molecular bands in stellar spectra are resonance bands, the fine structure due to the angular momentum of the electrons is not of much importance in astrophysical cases.

81. Occurrence of Bands in the Spectral Sequence

So far the following compounds have been identified with certainty in stellar spectra:

BO*	AlO*	TiO	ZrO	
H_2*	NH**	OH**	CaH*	MgH*
C_2	CH	CN		
LiF.				

The molecules marked with one asterisk have been found only in sunspots, while those marked with two asterisks have also been found in the solar atmosphere outside spots. The other molecules have been found in stars. They are all resonance bands—that is, they are due to absorption from the lowest energy level. No stellar bands have been found corresponding to ionized molecules. A few such bands are known (CO^+, N_2^+, O_2^+, MgH^+, and some others), but their absence from stellar spectra is a feature demanded by theory, since they cannot be excited thermally at all, their ionization energies being much greater than their energies of dissociation.

Titanium Oxide Bands. Most prominent among stellar bands are those due to TiO, which are so characteristic of stars of the *M*-class. They also occur, though more faintly, in stars of class *K* 2–*K* 8. They were first attributed to TiO_2, but later on Birge and Christy† gave the correct identification. These bands belong to two different electronic transitions, $^3\Pi \rightarrow {}^3\Pi$ and $^3\Pi \rightarrow {}^3\Sigma$ respectively, the former being situated in the green and the latter in the red part of the spectrum. These bands have been studied most thoroughly in the spectrum of Mira.

Zirconium Oxide Bands. The zirconium oxide bands are characteristic of class *S* spectra, which were separated out as a class by Merrill‡ for this reason. Merrill was also the first to identify the bands. The star χ Cygni is remarkable in being the only star known to exhibit bands of TiO and of ZrO at the same time.

Certain bands in the spectrum of Mira remained long unidentified until Baxandall recognized them as being due to AlO.

The hydrides NH, OH, CaH, and MgH have only been found in the sun, and the two latter only in the spots. Joy‖ reported that he

† *Phys. Rev.* **33** (1919), 701. ‡ *Astrophys. J.* **56** (1922), 457.
‖ Cf. P. Swings, *Actualités sci. et indust.* **50** (1932), 14.

had found MgH in the spectrum of Mira and claimed further also to have found a faint band of AlH in the spectrum of this star. These observations have not thus far been confirmed. On the other hand, Richardson† reports that AlH may exist in sun-spots.

The Swan Bands. The bands dominating the spectral classes *R* and *N*, the so-called Swan bands, are due to C_2. These bands belong to an electronic transition $^3\Pi \rightarrow {}^3\Pi$. It appears that the Swan bands and the TiO bands are mutually exclusive, a fact which may be interpreted on the dissociation theory, as will be shown later.

The *G* Band. The bands due to CH are also of considerable astrophysical significance. Two band systems due to CH occur in the stars, one situated near 4314 A. and the other near 3900 A. The first constitutes the so-called *G* band of Fraunhofer, and is particularly prominent in the sun. This band belongs to the electronic transition $^2\Pi \rightarrow {}^2\Sigma$, $^2\Pi$ being probably the ground-level of the molecule. The identification of the CH bands in the stars was made by Newall, Baxandall, and Butler.‡ They are observed in both branches of the spectral sequence. Passing along the sequence from *O*, *B*, *A* downwards, the *G* band is first definitely seen in some spectra of class *F*. It increases steadily in intensity until it reaches a maximum somewhere in the *K* class in the *M* branch, and in the *R* class in the *N* branch. Thus Rufus|| places the maximum of the *G* band in *R* 0 and *K* 0 respectively, while Shane†† places the maximum in the *G*–*R*–*N* branch at *R* 3. There is also some uncertainty with regard to the spectral class in which the bands first become visible.

It is difficult to obtain a clear estimate of the intensity of this band group, because there is a troublesome overlapping of lines due to calcium, titanium, and other metallic elements. The number of bands in the 4314 group is very large, and they are distributed with increasing mutual distance from 4314 to about 4270 A. The bands turn their heads towards longer wave-lengths, in conformity with the general rule that the bands within a sequence degrade in the direction of the sequence itself.

The Cyanogen Bands. These bands, which are due to CN, form two groups extending from about 4215 and 3885 A. towards shorter wave-lengths. They are resonance bands, and their intensity in-

† *Astrophys. J.* **73** (1931), 216.
‡ *Monthly Not.* **76** (1916), 640.
|| *Publ. Obs. Univ. Mich.*, **3** (1923), 258.
†† *Lick Obs. Bull.* **10** (1914), 79.

creases from their beginning in the latest F classes (F 8) to a maximum in late R classes. From then on through the N classes the intensity shows a steady decline. But there is a considerable uncertainty with regard to this maximum, and also the class in which the bands first become visible.

82. Isotope Effects in Band Spectra

Since the vibrational as well as the rotational energy of a molecule depends directly on its moment of inertia, it is clear that bands from molecules made from different isotopes will not coincide. The bands originating in the heavier molecules will be displaced towards the red with respect to bands belonging to the lighter molecule.

This fact provides a method for the detection of isotopes, which was first introduced by Kratzer† and Loomis.‡ The method is particularly well suited for astrophysical purposes, where practically all other methods are out of the question.

There are two cases of astrophysical interest, the discovery of isotopes of atmospheric oxygen by W. F. Giauque and H. L. Johnston,‖ and of the carbon isotopes in N stars by Sanford,†† both in 1929. The new oxygen isotopes have atomic masses 17 and 18 respectively, the heavier component being the most abundant of the two. Various determinations of the relative abundances give the result that O^{16} is ten thousand times more abundant than O^{17}, and about a thousand times more abundant than O^{18}. It is interesting that this discovery was first made from the terrestrial bands in the solar spectrum, being thus a strictly astrophysical discovery.

Conditions were a little different in the case of the carbon isotopes. Here the discovery was first made by King and Birge from laboratory spectra obtained by King. Shortly afterwards Sanford then pointed out that the existence of these isotopes would serve to explain the presence of three absorption bands in the vicinity of 4270 A. which have been known to exist in N type spectra for many years.

The table given on p. 233 will make the matter clear. Sanford suggested that the third band head might indicate the existence of a third isotope of mass 14 and ascribed the band accordingly to the molecule $C^{12}—C^{14}$. The suggestion that this band is due to $C^{13}—C^{13}$ was made by Menzel.‡‡

† *Zeits. f. Physik*, **3** (1920), 460.
‡ *Astrophys. J.* **52** (1920), 248.
‖ *Nature*, **123** (1929), 318, 831.
†† *Pub. Astron. Soc. Pacific*, **41** (1929), 291.
‡‡ Ibid. **42** (1930), 34.

TABLE 7

	Wave-lengths of Band Heads		
Hale	4737·5	4745·3	4752·8
Shane	4737·8	4745·2	4752·8
Sanford . . .	4737·6	4745·2	4752·4
Mean (obs.) . . .	4737·6	4745·2	4752·7
Computed . . .	4737·9	4744·5	4751·0
Furnace . . .	4736·9	4744·5	47...
Isotopes . . .	$C^{12}-C^{12}$	$C^{12}-C^{13}$	$C^{13}-C^{13}$

While the isotopes of oxygen are rather an isolated fact of no particular interest to astrophysics, the isotopes of carbon seem to play a much more important part. It has, for example, been suggested by Wurm† that the close overlapping of lines due to different isotopes may be responsible for the very strong intensity of the Swan bands in late-type *N* stars.

† *Zeits. f. Astrophys.* **5** (1932), 268.

XVII

APPLICATION OF THE THEORY OF MOLECULES TO THE DETERMINATION OF EFFECTIVE TEMPERATURE

83. Intensity in Bands and the Determination of Temperature

After this short survey of the main observational data we turn to the theoretical analysis of the problem exhibited by molecular compounds in stellar spectra. We shall proceed in successive steps corresponding to the real complexity of the problem, beginning with the intensity distribution in bands, then in band sequences, and finally the intensity of all bands considered as a whole, and the relative intensity of bands belonging to different molecules. This may not be the usual procedure, but it seems to be most natural from the theoretical point of view. As was shown earlier the total intensity of a line in a band considered as being generated by a simple rotator will be proportional to the *a priori* probability of the rotational state in question, that is, to $2j+1$, if j is the rotational quantum number of the state. If the gas is in thermal equilibrium, the number of atoms in the j-state will be affected by a Boltzmann factor

$$e^{-j(j+1)h^2/8\pi^2AkT},$$

where T is the temperature. We may thus put the intensity of a single component line of the band equal to

$$I_j = \text{const.}(2j+1)e^{-j(j+1)h^2/8\pi^2AkT}, \tag{113}$$

it being assumed that j corresponds to the initial state of the line in question. This expression shows that the intensity will rise to a maximum for a certain value of j, which may be determined with a sufficient approximation by considering j as a continuous variable. By differentiation of I_j with respect to j we then find

$$\frac{(j+\frac{1}{2})}{I_j}\frac{dI_j}{dj} = 1-\frac{h^2(2j+1)^2}{16\pi^2AkT}. \tag{114}$$

For the maximum intensity the differential quotient vanishes. Denoting the corresponding j-value by j_m, and solving (114) for the temperature, we find

$$T = \frac{h^2(2j_m+1)^2}{16\pi^2Ak}. \tag{115}$$

This method of temperature determination was first proposed by Heurlinger.†

These considerations are only directly applicable to emission bands, the total intensity of the source being so small that reabsorption is negligible. It is not *a priori* clear that the maximum intensity in an absorption band should be found in the same position as in the case of emission bands, and the direct application of formula (113) to stellar spectra is therefore somewhat conjectural, especially as we must take into account the merging of the lines into a continuous band. But these conditions have not thus far been discussed in detail, all applications being based directly on (113).

84. Temperature of the Sun from Bands

Birge‡ and later Richardson‖ tried to apply this method to the sun. Birge studied the intensity distribution in one of the solar cyanogen bands (3883 A.), as well as in the same band photographed by King for different temperatures in the electric furnace. Birge arrived at a temperature for the reversing layer of

$$T = 4{,}300 \pm 500° \text{K}.$$

This value is considerably lower than the temperatures derived by other methods, and it is likely that it has been affected by some systematic errors.

More recently Richardson attacked the same problem, but this time not only for the reversing layer at the undimmed solar surface, but also in sun-spots. Richardson did not use the cyanogen bands, but the Swan bands of carbon at 5165 A. A direct comparison of the observed and theoretical intensities of lines in the 5165 band yielded a temperature of $4{,}900 \pm 600°$ K. for the sun-spot umbra and $6{,}000 \pm 700°$ K. for the reversing layer. From the position of the maximum intensity in the band he derived a temperature of $4{,}950 \pm 750°$ K. for the sun-spot umbra and $5{,}700 \pm 1{,}600°$ K. for the reversing layer. A third variation of the method yielded similar results.

It will be noticed that the temperatures derived by Richardson are, on the whole, higher than those found by Birge, and they accord on the whole with the temperatures derived from the distribution of energy in the continuous spectrum of the sun. The probable errors of the determinations, however, are so large that it is difficult to say

† *Dissertation*, Lund, 1918, p. 65. ‡ *Astrophys. J.* **55** (1922), 273.
‖ Ibid. **73** (1931), 216.

whether Richardson's temperatures are inherently better than those found by Birge.

The essential difficulty of this method in its application to the sun and similar stars consists in the blending of lines of the bands with foreign lines of the spectrum. Although the method affords an interesting alternative for the determination of solar temperatures, it seems still to be less reliable than the methods depending upon the distribution of radiation in the continuous spectrum.

85. Temperature of the Ionosphere

A corresponding application of the theory of band spectra was made by Vegard† in an attempt to determine the temperature of the upper terrestrial atmosphere. As is well known the aurora borealis originates in high atmospheric layers extending from a lower limit of about 80 km. to about 1000 km., the average being about 100 km. The spectra of aurorae show mostly narrow bands, which correspond so closely to the position of negative and positive bands of nitrogen that the identification with these bands is certain. A large number of these bands are found in the ultra-violet part of the spectrum, extending right down to the limit of atmospheric absorption; there are also numerous bands in the red part of the spectrum, which also seem mostly to belong to nitrogen, although some of the bands may have a different origin.

The intensity of auroral radiation is so weak that it has been impossible so far to photograph the spectrum with anything but very low dispersion, so that the bands are never resolved into individual lines. For this reason the formula (113) for the intensity distribution among the lines cannot be applied directly to the intensity in blended bands without a correction. For this formula only takes account of the intensity of the individual lines, and does not take account of the variation in the number of lines which fall within the slit-width, and which are thus blended together. This number will differ from place to place in the spectrum, and would produce a variation of intensity even if all individual lines had the same intensity.

Let $\Delta\nu_j$ be the frequency difference between adjacent lines at the place of the line with initial quantum state j. The number of lines blended by the slit may then be put proportional to $(\Delta\nu_j)^{-1}$, which must be introduced as an additional factor to I_j before it will give the

† *Geofys. Pub. Oslo*, **9** (1932), No. 11.

intensity in the blended band. As a sufficiently accurate expression of $\Delta\nu_j$ we may write for a positive branch

$$\Delta\nu_j = \tfrac{1}{2}(\nu_{j+1}-\nu_j)+\tfrac{1}{2}(\nu_j-\nu_{j-1}) = \tfrac{1}{2}(\nu_{j+1}-\nu_{j-1}) = \frac{h}{4\pi^2 A}(\delta j+1),$$

so that the final expression of I_j becomes

$$I_j = \text{const.}\frac{2j+1}{\delta j+1} e^{-j(j+1)h^2/8\pi^2 AkT}.$$

Proceeding in the usual way we find the following condition for the maximum of this expression

$$T = \frac{h^2(2j+1)^2(\delta j+1)}{8\pi^2 AkT(2-\delta)}.$$

Introducing into this formula the data given by Vegard for the negative nitrogen band 4278, the auroral temperature comes out of the order 70° C.

Just how much this means is not easy to say, because the difficulties of obtaining auroral spectra may introduce systematic errors which are very difficult to control. But it is interesting in so far as it shows that there exists at present no direct evidence of any low temperature of the upper atmosphere.

It must be remembered, however, that since the phenomena of auroral display certainly are produced by agencies very far from thermal equilibrium, such as very fast electrical particles, the temperatures derived from band spectra in the above way may well differ from the temperature determining the pressure of the gas at the place in question. We shall see an interesting case of this sort, when we turn to the temperatures derived from the distribution of molecules among different vibrational states.

86. Temperatures from the Relative Intensities of Bands in a Sequence

In stellar spectra it will perhaps never be possible to resolve the bands into individual lines, so that any application of the above-mentioned method will have to adopt the correction due to varying spacing of the lines. Thus far no one has made any attempt at temperature determinations of this sort. Probably the theory would need further amplification before being applicable to absorption bands. However, there is another possibility, which is likely to have a wider range of application, namely, to derive the relative number of molecules in different *vibrational* states, and then to apply Boltzmann's

principle to these numbers. Differential measurements of intensities within a single band would then be superfluous, as all that is needed is the total intensity of each band. This method was applied in a preliminary way by Wurm to the spectra of comets as well as to the spectra of late-type stars, where his results were in qualitative agreement with other temperature determinations.

One of the difficulties encountered here is the fact that a fairly accurate knowledge of the probabilities of transition between different vibrational levels is a necessary corollary for the method. It is therefore to be regretted that the theory of these transition probabilities has only been developed in a few simple cases, and even there further amplifications are greatly to be desired. All the molecules in question are of the non-polar kind, which have no electric moment in their normal state with no electronic states excited. For this reason the theory of the harmonic oscillator, which gives a good approximation when it is a question of calculating the lowest vibrational energy levels, cannot be used directly even for these lowest levels. It was showed by Hutchisson, however, that the theory of the harmonic oscillator when properly interpreted, will serve as a starting-point even for the theory of vibrational transition probabilities of non-polar molecules. We shall therefore consider the essential points of the Hutchisson theory.

87. Hutchisson's Theory of Vibrational Transition Probabilities

We assume that the coupling between the electronic motion and the rotation of the molecule is so slight that we may write the wave function of a molecular state in the form

$$\psi = \phi_E \Gamma_V,$$

where ϕ_E includes all terms pertaining to the rotation of the molecule and the electronic motion, while Γ_V is the wave function of the vibrational motion. We divide the electrical moment of the molecule in a given configuration into two parts: M_E which is associated with rotation and electronic configuration, and M_V which is associated with the vibrations. The harmonic amplitude of the molecule associated with a transition from the given state to one distinguished by a double prime will be given by the formula

$$J = \int (M_V + M_E)\psi'\psi''^* \, d\tau,$$

where an asterisk denotes the complex conjugate, and $d\tau$ is the elementary volume of the configuration space of the molecule, the integration extending over all configuration space. This formula is so far quite general. For non-polar molecules, however, the vibrational electrical momentum vanishes, so that M_V may be dropped from the above formula. Next M_E will be sensibly independent of the states of vibration, that is, it will depend on the electronic and rotational coordinates only. The above multiple integral will therefore split up into a product of two integrals

$$J = J_E J_V,$$

where $$J_E = \int M_E \phi'_E \phi''^*_E \, d\tau_E; \quad J_V = \int \Gamma'_V \Gamma''^*_V \, d\tau_V.$$

Now J_E must be expected to be independent of the states of vibration, at least in a region of low vibrational quantum numbers, as a consequence of which the *relative* probability of different vibrational transitions corresponding to the same electronic transition will be given by J_V alone. These are the fundamental assumptions of Hutchisson's theory.†

It is easily proved that to a first approximation the vibrations may be considered similar to those of a linear harmonic vibrator. However, if we were to use unperturbed wave functions for a harmonic vibrator in calculating J_V, all vibrational transition probabilities would vanish. We know, however, that in an electronic transition the molecule suffers a perceptible deformation, resulting in a change of moment of inertia, which again means a change in the relative separation of the atoms as well as a change in the proper frequency of vibration. In virtue of these changes the Hermitian wave functions will no longer be necessarily orthogonal, and so will allow of a non-vanishing value for the integral J_V. Since this result is due expressly to the excited electron, it is in harmony with the fact that homopolar molecules rarely show pure rotation bands or bands of the vibration-rotation type.

The wave function of a harmonic oscillator was earlier found to be given by (see p. 34)

$$\Gamma_V = \frac{1}{N_s} e^{-\frac{1}{2}\eta^2} H_s(\eta) \qquad (N_s = \sqrt{\{2^s s! \sqrt{\pi} k^{-1}\}}),$$

where H_n is a Hermitian polynomial, while the independent variable

† *Phys. Rev.* **36** (1930), 410.

η is related to the radial separation r of the atoms by the formula

$$\eta = 2\pi\sqrt{\left(\frac{\nu_0 m}{h}\right)}(r-r_0) = 2\pi\sqrt{\left(\frac{\nu_0 A}{h}\right)}\left(\frac{r}{r_0}-1\right).$$

In this formula r_0 is the equilibrium value of the separation of the two nuclei, m is the reduced mass of the molecule, and A the moment of inertia. The change in frequency ν_0 and average nuclear separation of the atoms in other states may be brought into the formulae by assuming the independent variable η to be changed into $\alpha\eta+\delta$, where α and δ are constants characteristic for each state. We have then

$$\alpha\eta+\delta = 2\pi\sqrt{\left(\frac{\nu_0' A'}{h}\right)}\left(\frac{r}{r_0'}-1\right) = 2\pi\sqrt{\left(\frac{\nu_0' A}{h r_0^2}\right)}\{(r-r_0)+(r_0-r_0')\},$$

since A is proportional to r_0^2, which makes A/r_0^2 independent of the change of state. It follows therefore simply from the above equation that

$$\alpha = \sqrt{\left(\frac{\nu_0'}{\nu_0}\right)},$$

and

$$\delta = 2\pi\sqrt{\left(\frac{\nu_0' m}{h}\right)}(r_0-r_0') = 0{\cdot}1221\sqrt{(\nu_0' M)}(r_0-r_0').$$

Here ν_0' is expressed in cm.$^{-1}$. M is the atomic weight of the atoms, which here are supposed identical. It is assumed that (r_0-r_0') is measured in Ångström's units. In order to carry out the integrations involved it is most convenient to use the generating function of the Hermitian polynomials

$$e^{-x^2+2x\eta} = \sum_{s=0}^{\infty}\frac{H_s(\eta)}{s!}x^s.$$

Similarly for the varied state

$$e^{-y^2+2y(\alpha\eta+\delta)} = \sum_{r=0}^{\infty} H_r(\alpha\eta+\delta)\frac{y^r}{r!}.$$

Multiplying these expressions together, multiplying the product by the proper exponentials, and integrating over all values of η we find

$$\sum_{s=0}^{\infty}\sum_{r=0}^{\infty}\frac{x^s y^r}{s!\,r!}\int_{-\infty}^{\infty} H_s(\eta)H_r(\alpha\eta+\delta)e^{-\frac{1}{2}\eta^2-\frac{1}{2}(\alpha\eta+\delta)^2}\,d\eta$$

$$= \sqrt{\left(\frac{2\pi}{1+\alpha^2}\right)}\exp\left(-\frac{\delta^2}{2(1+\alpha^2)}+\frac{(1-\alpha^2)(x^2-y^2)-2\alpha\delta x+2\delta y+4\alpha xy}{1+\alpha^2}\right).$$

By expanding the last exponential factor in powers of x and y, and

then equating coefficients of the same powers on both sides, we obtain just those integrals which are needed for the calculation of the harmonic amplitudes. The result is

$$J_\nu = \int_{-\infty}^{\infty} H_s(\eta) H_r(\alpha\eta+\delta) e^{-\frac{1}{2}\eta^2-\frac{1}{2}(\alpha\eta+\delta)^2} \frac{d\eta}{kN_sN_r}$$

$$= K\sqrt{\left(\frac{s!\,r!}{2^{s+r}}\right)} \sum_{l=0}^{s \text{ or } r} \sum_{i=0}^{\frac{1}{2}(s-l)} \sum_{j=0}^{\frac{1}{2}(r-l)} a_{2l} b_{2i} c_{2j} d_{s-2i-l} e_{r-2j-l}. \qquad (116)$$

The various letters introduced here have the following meanings:

$$K = \left(\frac{2\alpha}{\alpha^2+1}\right)^{\frac{1}{2}} e^{-\frac{1}{2}\delta^2/(\alpha^2+1)}; \qquad a_{2l} = \frac{1}{l!}\left(\frac{4\alpha}{1+\alpha^2}\right)^l;$$

$$b_{2i} = \frac{1}{i!}\left(\frac{1-\alpha^2}{1+\alpha^2}\right)^i; \qquad c_{2j} = \frac{1}{j!}\left(\frac{\alpha^2-1}{\alpha^2+1}\right)^j;$$

$$d_{s-2i-l} = \frac{1}{(s-2i-l)!}\left(\frac{-2\alpha\delta}{1+\alpha^2}\right)^{s-2i-l}; \qquad e_{r-2j-l} = \frac{1}{(r-2j-l)!}\left(\frac{2\delta}{1+\alpha^2}\right)^{r-2j-l}.$$

TABLE 8

Values of J_ν/K, equation (116)

r	s	
0	0	1
0	1	$\sqrt{\frac{1}{2}}e_1$
0	2	$\sqrt{\frac{1}{2}}(e_2+c_2)$
1	1	$\frac{1}{2}(d_1e_1+a_2)$
0	3	$\frac{1}{2}\sqrt{3}(e_3+e_1c_2)$
1	2	$\frac{1}{2}(d_1e_2+c_2d_1+a_2e_1)$
0	4	$\frac{1}{2}\sqrt{6}(e_4+c_2e_2+c_4)$
1	3	$\frac{1}{4}\sqrt{6}(d_1e_3+c_2d_1e_1+a_2e_2+a_2c_2)$
2	2	$\frac{1}{2}(d_2e_2+c_2d_2+b_2e_2+a_2d_1e_1+a_1)$
0	5	$\frac{1}{2}\sqrt{15}(e_5+c_2c_3+c_4e_1)$
1	4	$\frac{1}{4}\sqrt{3}(d_1e_4+c_2d_1e_3+c_4d_1+a_2e_3+a_2c_2e_1)$
2	3	$\frac{1}{4}\sqrt{6}(d_2e_3+c_2d_2e_1+b_2c_3+b_2c_2e_1+a_2d_1c_2+a_2c_2d_1+a_4e_1)$
1	5	$\frac{1}{4}\sqrt{30}(d_1e_5+c_2d_1e_3+c_4d_1e_1+a_2e_4+a_2c_2e_2+a_2c_4)$
2	4	$\frac{1}{4}\sqrt{3}(d_2e_4+c_2d_2e_2+c_4d_2+b_2e_4+b_2c_2e_2+b_2c_4+a_2d_1e_3+a_2c_2d_1e_1+a_4c_2+a_4c_2)$
3	3	$\frac{3}{4}(d_3e_3+c_2d_3e_1+b_2d_1e_3+b_2c_2d_1e_1+a_2d_2e_2+a_2c_2d_2+a_2b_2e_2+a_2b_2c_2+$ $+a_4d_1e_1+a_6)$
2	5	$\frac{1}{4}\sqrt{30}(d_2e_5+c_2d_2e_3+c_1d_2e_1+b_2e_5+b_2c_2e_3+b_2c_4e_1+a_2d_1e_4+a_2c_2d_1e_2+$ $+a_2c_4d_1+a_4e_3+a_4c_2e_1)$
3	4	$\frac{1}{4}\sqrt{6}(d_3e_4+c_2d_3e_2+c_4d_3+b_2d_1e_4+b_2c_2d_1e_2+b_2c_4d_1+a_2d_2e_3+a_2c_2d_2e_1+$ $+a_2b_2e_3+a_2b_2c_2e_1+a_4d_1e_2+a_4c_2d_1+a_6e_1)$
4	4	$\frac{3}{2}(d_4e_4+c_2d_4e_2+b_2d_2e_4+b_2c_2d_2e_2+b_4e_4+c_4d_4+b_4c_4+a_2d_2e_3+$ $+a_2b_2d_1e_3+a_2c_2d_2e_1+a_2c_2b_2d_1e_1+a_4d_2e_2+a_4c_2d_2+$ $+a_4b_2e_2+a_4b_2c_2+a_6d_1e_1+a_8)$

Moreover, the double limits of the sums in (116) are to be understood in the manner that the *smaller* value of s or r in the first sum is the upper limit, while the upper limits in the second and third sums are either the upper or lower figure depending upon whether $s-l$ or $r-l$ is *even* or *odd*. The table on p. 241 for the integral (116) will be found useful.

In a later paper† Hutchisson extended this theory so as to take account of the tendency towards anharmonic vibrations which soon becomes perceptible for higher vibrational quantum numbers. The calculations then soon become very complicated. The above expressions will suffice for a survey of stellar spectra.

88. *N*-type Spectra

The above theory was applied by Wurm‡ to the intensity of the Swan bands and the cyanogen bands in the stars. In the case of cyanogen the molecule is not quite symmetrical, as assumed in the theory, but the difference in atomic weight between carbon (12) and nitrogen (14) is so small that there is no reason to anticipate any difficulties on this account. A practical difficulty of the theory is the fact that the change in nuclear separation during an electronic transition is not known with sufficient accuracy in the present case. This change is usually of the order a few per cent. of the separation itself. For C_2 the analysis of the bands yields the values

$$r_0 = 1{\cdot}265\,\text{A.}; \qquad r_0' = 1{\cdot}311\,\text{A.}; \qquad r_0 - r_0' = -0{\cdot}046\,\text{A.}$$

In order to bring out the influence of any uncertainty in $\Delta r_0 = r_0 - r_0'$ Wurm carried through the calculations for various values of Δr_0. Having derived the necessary constants, all there is to do is to calculate the square of the harmonic amplitudes and multiply them by the fourth power of the corresponding frequencies. This latter factor was omitted in the analysis of individual bands, because the width of a band is always very small in comparison with the frequency of its zero line. But when comparing different bands in a sequence the difference is no longer small, as will be realized by an inspection of the corresponding tables.

Although the theory is only developed for emission bands of small intensity, Wurm applied it directly to the absorption bands in late-type stars. It is clear then that the results of a comparison between theory and observations must be interpreted *cum grano salis*, and that

† *Phys. Rev.* **37** (1931), 45. ‡ *Zeits. f. Astrophys.* **5** (1932), 260.

it will have no meaning to stress details very much. Since the intensity of a band cannot well be expressed by a simple formula, we cannot determine the temperature by a simple determination of maximum intensities. It becomes necessary to compare the empirical intensities of a whole sequence of bands with theory for different assumed temperatures, in order to single out the temperature which best fits the facts. Another difficulty which prevents an accurate temperature-determination for the moment is the fact that there are no accurate photometric measurements of the intensities of the bands in late-type spectra. Wurm therefore had to be content with tentative estimates of intensities from the spectra published round the beginning of the century by Hale, Ellerman, and Parkhurst.† This is, of course, only a temporary difficulty, which will probably soon be overcome by some one interested in late-type star problems.

Wurm found that the group 6,191 A. and 5,635 A. of C_2 were most suited for the purpose, because they appear with moderate intensities in the spectra, while the other bands very soon grow so strong that estimates of intensities have very little meaning in this respect. Wurm studied the spectrum of the star 152 Schjellerup, which is of spectral class *N* 3, and of 19 Piscium of type *N* 0. The fact that the first band of the group 6,191 A. is slightly stronger than the second band indicates at once that the temperature of these stars is lower than 2,000° K. This estimate is not so arbitrary as might be imagined, because a drop in temperature in this region will affect the intensities very seriously. It is more sensitive to a lowering of the temperature than to an increase. This result was further confirmed by reference to the group at 5,635 A., so that a 'band' temperature of stars in the range *N* 0–*N* 3 of about 1,700° K. should not be far from the truth.

The comparison between theory and observation for the cyanogen bands was not carried out by Wurm on account of lack of suitable material.

89. Bands in the Spectra of Comets

The spectra of comets have been carefully studied during the whole reign of stellar spectroscopy, valuable observations being made as far back as seventy-five years ago.

The comet which has been studied most thoroughly spectroscopically is that of Halley. In particular, the observational facts as regards

† *Pub. Yerkes Obs.* **2** (1904).

spectrograms obtained of this comet during its appearance in 1910 were discussed by Bobrovnikoff.† Most typical for molecular bands in cometary spectra are the cyanogen bands of the type $^2\Sigma \rightarrow {}^2\Sigma$. The relative intensities of these bands are peculiar in that bands starting from excited vibrational states decrease very rapidly in intensity with increasing quantum number. According to Wurm‡ this indicates a temperature of the order of 300° K. for the radiating gases in the comets. But this estimate does not mean very much apart from emphasizing that the temperature evidenced by the spectrograms is much lower than that of the sun.

90. Intensity of Auroral Bands

A similar application of the band-spectrum theory was made by Rosseland and Steensholt|| to the spectrum of polar aurorae. From the Hutchisson theory they calculated the theoretical intensities of the ultra-violet negative and positive nitrogen bands, and compared the results with Vegard's estimate of relative intensities of these bands on auroral spectrograms. The temperature which would correspond to the observed data came out to be very large, something between 2,000° and 3,000° K. The fact that this result differs very much from the result obtained from the width of the band 4,276 A. discussed earlier indicates that the conditions under which the bands are emitted must differ radically from those of thermal equilibrium. That band spectra are sensitive to such deviations has been demonstrated experimentally,†† so that this result is not surprising. When the density is so low that ordinary collisions may be neglected, excitation by electron impacts will not change the distribution of the molecules very much as regards the rotations, while the electronic transitions suffered by the molecule will profoundly alter the oscillations. The high oscillation temperature shown by auroral bands is therefore restricted to the oscillations, and indicates that aurorae are excited by the impact of swift particles.

† *Lick Obs. Bull.* **17** (1931), 204. See also W. H. Wright, ibid. **209** (1912–13), 13 (comet Brooks). W. W. Campbell, ibid. **5** (1908), 31 (comet Daniel); V. Slipher, *Lowell Obs. Bull.* **52** (1911).

‡ *Zeits. f. Astrophys.* **5** (1932), 10. Cf. also H. Zanstra, *Monthly Not.* **89** (1928), 178, and Baldet, *Actualités sci. et. indust.* **16** (1930).

|| *Pub. Oslo Univ. Obs.* **7** (1933).

†† Cf. for example O. Oldenberg, *Phys. Rev.* **46** (1934), 210.

XVIII

DISSOCIATIVE EQUILIBRIUM OF STELLAR COMPOUNDS

91. The Reaction Isochore

HAVING discussed various problems related to the intensities of bands due to a given molecule, we shall now approach the more difficult problem of the relative intensities of bands belonging to *different* molecules. Only so can we hope to develop a theory of the role of molecular spectra in the stellar sequence, or in the change of type of a long-period star during its course of variation.

The first quantitative study of this problem was due to Atkinson.† Later improvements were given by Piccardi,‡ Christy,‖ Wildt,†† Swings,‡‡ Yvonne Cambresier and Rosenfeld,‖‖ and Russell.†††

The general problem of discussing all possible chemical reactions which may take place in a stellar atmosphere is, of course, too complicated to be attacked in full generality. We must therefore proceed in successive approximations, taking the observed facts as a guide. As a first approach to the problem we therefore study the equilibrium between different molecular combinations in an atmosphere under the same simplifying assumptions as were used for the ionization problem. We thus assume the state to correspond to local thermodynamic equilibrium, so that the state of dissociation is governed by the equations derived earlier (Chap. II):

$$n_k = f_k e^{\sum_i m_i^k \lambda_i + u_k},$$

n_k being the relative number of molecules of the kth type, f_k its partition function, m_i^k the number of atoms of the ith kind in the molecule, while u_k gives the deviation from the ideal gas laws. In the following we shall assume the ideal gas laws to hold, so that u_k may be neglected. We shall further, for the most part, only be concerned with diatomic compounds, which are the only ones directly observed in stellar spectra. In some cases we know, however, that more complex compounds will be present and play an important

† *Monthly Not.* **82** (1922), 396.
‡ *Mem. d. Soc. Astron. Italia*, **4** (1929), 379.
‖ *Astrophys. J.* **70** (1929), 1.
†† *Zeits. f. Physik*, **54** (1929), 856; *Zeits. f. Astrophys.* **9** (1934), 176.
‡‡ *Monthly Not.* **92** (1931), 140; *Actualités sci. et indust.* **50** (1932).
‖‖ *Monthly Not.* **93** (1933), 710.
††† *Astrophys. J.* **79** (1934), 317.

part. This is, for instance, likely to be the case with carbon dioxide in the carbon stars, as we shall see later. We further also know that the atmospheres of the outer planets contain large amounts of ammonia (NH_3) and methane (CH_4), so that the study of more complex molecular compounds cannot be avoided altogether in astrophysics.

Consider then a mixture of two elements A and B and the formation of a compound AB, while disregarding for the moment the possibility of the formation of compounds AA and BB. Let n_A and n_B denote the number of free atoms A and B per unit volume, and n_{AB} the corresponding number of molecules AB. The above system of equations of dissociative equilibrium then reduces to the following three:

$$n_A = f_A e^{\lambda_A}; \qquad n_B = f_B e^{\lambda_B}; \qquad n_{AB} = f_{AB} e^{\lambda_A + \lambda_B}.$$

Eliminating λ_A and λ_B we find the well-known equation

$$\frac{n_A n_B}{n_{AB}} = \frac{f_A f_B}{f_{AB}}.$$

For the range of temperatures met with in cool stars it is permissible to neglect the electronic excitation of the atoms, so that the atomic partition function reduces to that of translational motion multiplied by the statistical weight of the normal quantum state, which may be denoted by G. That is,

$$f_A = (2\pi M_A kT)^{\frac{3}{2}} G_A h^{-3}; \qquad f_B = (2\pi M_B kT)^{\frac{3}{2}} G_B h^{-3},$$

where M_A and M_B are the masses of the atoms in question.

The partition function of the molecule is a more complicated affair. For the present purpose it is sufficient to consider the molecular rotations and vibrations as independent, and quantized as simple rotators and harmonic oscillators. The partition function then splits up into a quadruple product

$$f_{AB} = G_{AB} f' f'' f'''.$$

Here f' may be taken to represent the translational motion of the molecule, so that

$$f' = \{2\pi(M_A + M_B)kT\}^{\frac{3}{2}} h^{-3}.$$

Next f'' may represent the rotations, so that

$$f'' = \sum_{j=0}^{\infty} (2j+1) e^{-j(j+1)h^2/8\pi^2 AkT}.$$

We have here used the expression $j(j+1)\hbar^2/8\pi^2A$ for the energy of a rotator, which was derived earlier (p. 43), assuming the moment of inertia constant. Under stellar conditions the rotations will be so strongly excited that it is permissible to treat j as a continuous variable and change the sum into an integral. We then have

$$f'' = \int_0^\infty e^{-(j+1)j\hbar^2/8\pi^2AkT}\, d[(j+1)j] = 8\pi^2AkT/\hbar^2.$$

Denoting the dissociation energy of the molecule by χ we have thirdly, for the vibrational partition function,

$$f''' = \sum_{n=0}^{\infty} e^{(\chi - n\hbar w)/kT}.$$

Here w is the fundamental vibration frequency of the molecule. The zero energy of the vibrations, $\frac{1}{2}hw$, has been included in χ. The summation of the geometrical series involved in this formula is easily performed, giving

$$f''' = \frac{e^{\chi/kT}}{1-e^{-\hbar w/kT}}.$$

Finally G_{AB} is the statistical weight of the ground state of the electronic configuration of the molecule, the nuclei being regarded as fixed.

We are now in the position to write down the final expressions of $n_A n_B/n_{AB}$. We find

$$\frac{n_A n_B}{n_{AB}} = \frac{G_A G_B}{G_{AB}} \frac{(2\pi M_{AB} kT)^{\frac{3}{2}}}{8\pi^2 kThA} e^{-\chi/kT}(1-e^{-hw/kT}),$$

where we have introduced the reduced mass of the molecule

$$M_{AB} = \frac{M_A M_B}{M_A + M_B},$$

which also is the molecular mass to be used in the calculation of the moment of inertia. If we prefer doing so we may therefore achieve some further simplification by introducing the equilibrium distance between the atoms r_0 by the formula

$$A = r_0^2 M_{AB}.$$

The dissociation formula may then be written in the equivalent form

$$\frac{n_A n_B}{n_{AB}} = \frac{G_A G_B}{G_{AB} h r_0^2} \left(\frac{M_{AB} kT}{8\pi}\right)^{\frac{1}{2}} e^{-\chi/kT}(1-e^{-hw/kT}).$$

We have here assumed that the two atoms A and B are intrinsically different. If the two atoms are identical, the above formula suffers a minor change, as it is necessary to multiply the right-hand side by 2. This point was discussed by Gibson and Heitler.† The above equation is usually called the reaction isochore.

92. Expressions for the Relative Abundance of Molecules in an Atmosphere

The next problem is to find out how varying dissociation at different levels will influence the resulting band intensity. This problem, which is completely analogous to the consideration of varying ionization with change of level in case of atoms, was first considered by Cambresier and Rosenfeld.‡

We suppose that the intensity of a definite component line of a band is essentially determined by the number of relevant molecules in the right quantum state, overlying the photosphere, just as for atomic lines. Denoting this number by N, and assuming it to refer to molecules AB, we have

$$N = \int n_{AB}\, dz,$$

the integral being extended from the top of the atmosphere and down to the photospheric level; z denotes the depth below a chosen outer limit of the atmosphere.

Further progress will now depend essentially on the functional form adopted for the coefficient of continuous absorption. The study of atomic lines has indicated that the mass absorption coefficient probably increases with pressure, so we may adopt as a rough expression

$$\kappa = \kappa_0 p\rho,$$

p being the total pressure, and κ_0 a quantity depending on temperature and the wave-length of light. We have discussed earlier the limitations involved in such a procedure, so that it need not detain us here. The optical depth of the photospheric level is thus

$$\tau_0 = \int \kappa_0 p\rho\, dz = \frac{\kappa_0}{g} \int p\, dp = \frac{1}{2}\frac{\kappa_0}{g} P^2,$$

where P is the total pressure in the photosphere. The atmospheric temperature is then assumed constant during the integration.

† *Zeits. f. Physik*, **49** (1928), 465.
‡ *Monthly Not.* **93** (1933), 710.

A case of wide applicability is met with when we assume that the element A is so excessively abundant that the fraction of it bound in molecules AB is negligibly small compared to the unbound fraction. The second element B may be present with a small absolute abundance, and an appreciable fraction of it may be in the form of molecules AB. We are then entitled to put n_A proportional to the total pressure. For the sake of generality we may assume that there exist a whole series of elements of type A and a series of elements of type B without introducing much more complication, provided only that we can neglect intercombinations of the type $A_1A_2, \ldots$, or $B_1B_2, \ldots$. This case is general enough to be of interest for the stellar problem.

Denote the relative abundance of the elements by μ_A and μ_B respectively, it being assumed that

$$\mu_A \gg \mu_B \quad \text{as well as} \quad \sum_A \mu_A \gg \sum_B \mu_B,$$

the sum being extended to all elements of the respective groups. Let further y_B denote the total number of B atoms in unit volume, and k Boltzmann's constant, so that

$$n_A = \mu_A \frac{p}{kT}, \qquad y_B = \mu_B \frac{p}{kT}. \tag{117}$$

The numbers n_{AB} and n_B will now satisfy the equations

$$\begin{aligned} n_{AB} + n_B &= y_B \\ \text{and} \qquad n_A n_B &= K n_{AB}. \end{aligned} \tag{118}$$

Eliminating n_B from these equations we find the following expression for n_{AB}:

$$n_{AB} = \frac{y_B}{1 + K/n_A}.$$

On the right-hand side we shall introduce the total pressure instead of n_A and y_B on using the above expressions (117). We then find

$$n_{AB} = \frac{\mu_B}{kT} \frac{p^2}{p + KkT/\mu_A}. \tag{119}$$

The equation of hydrostatic equilibrium is, R denoting the gas constant of the material,

$$dp = g \frac{p}{RT} dz,$$

so that the expression for N may be brought to the form

$$N = \frac{\mu_B}{g}\frac{R}{k}\int_0^P \frac{p\,dp}{p+KkT/\mu_A},$$

P being the pressure at the photospheric level.

Carrying out the integration we find

$$N = \frac{\mu_B P R}{gk} - \frac{\mu_B}{\mu_A}\frac{KRT}{g}\log\left(1+\frac{P\mu_A}{KkT}\right).$$

We have finally to express P in terms of the optical thickness of the atmosphere:

$$P = \sqrt{\left(\frac{2g\tau_0}{\kappa_0}\right)}.$$

The final expression of N is therefore

$$N = \mu_B\frac{R}{k}\sqrt{\left(\frac{2\tau_0}{g\kappa_0}\right)} - \frac{\mu_B}{\mu_A}\frac{KRT}{g}\log\left\{1+\frac{\mu_A}{KkT}\sqrt{\left(\frac{2g\tau_0}{\kappa_0}\right)}\right\}. \qquad (120)$$

It is not difficult to see that an increase in g corresponds to a decrease in N and vice versa. Hence the bands in question are all strengthened by an increase in the giant characteristics, i.e. decreasing density.

It goes without saying that the above considerations do not apply to molecules like B_2 where both atoms of the molecule are the same, and of small absolute abundance. However, limiting the consideration to the case when the number of molecules of the type B_2 is small in comparison with the number of molecular types like AB, it is not difficult to find the proper formulae. Writing the corresponding equation of dissociative equilibrium in the form

$$n_B^2 = K' n_{2B},$$

we may assume n_B to be given by equations (118) and (119), since we have assumed the earlier equilibrium to be practically undisturbed. Then

$$n_B = y_B\left(1-\frac{1}{1+(K/n_A)}\right) = \frac{\mu_B}{\mu_A}K\frac{p}{p+(KkT/\mu_A)},$$

and

$$n_{2B} = \left(\frac{\mu_B}{\mu_A}\right)^2\frac{K^2}{K'}\frac{p^2}{\{p+(KkT/\mu_A)\}^2}.$$

Hence it follows by integration that

$$N' = \int_{p=0}^{p=P} n_{2B}\,dz = \frac{RT}{g}\left(\frac{\mu_B}{\mu_A}\right)^2\frac{K^2}{K'}\left\{\log\left(1+\frac{\mu_A P}{KkT}\right) - \frac{1}{1+(KkT/\mu_A P)}\right\}. \qquad (121)$$

Apart from minor differences of notation the formulae (120) and (121) are the same as those given by Cambresier and Rosenfeld. It will be noticed that assuming $\mu_A P/KkT$ to be a small quantity in these formulae, and expanding the expressions in the bracket to the second power of P, N and N' become proportional to P^2/g, so that g drops out when P is expressed as a function of the optical thickness of the atmosphere.

The formulae we have been developing contain a mixture of quantities which either may be determined from laboratory experiments or are typical stellar data. The former quantities are included in the dissociation coefficient K, and are given in general in Table 14 in the Appendix, in so far as the molecular compound has an astrophysical interest. We may therefore consider K known when the temperature is given. The absorption coefficient is essentially a stellar datum, along with the quantities μ_A, μ_B, R, and g. In order to take the dependence of the absorption coefficient on temperature roughly into account, we write

$$\kappa_0 = \sigma T^{-n},$$

where n may be assumed to be known from the study of atomic lines. The stellar quantities then occur in the above formulae in the combinations

$$\alpha = \mu_A \sqrt{\left(\frac{g}{\sigma}\right)} \quad \text{and} \quad \beta = \mu_B R \sqrt{\left(\frac{1}{g\sigma}\right)}.$$

Both α and β are determinable in principle for an individual star, on assuming that the number N may be found from the intensity of the bands. If σ, R, and g were known beforehand, μ_A and μ_B could be determined. Conversely, knowing the μ's, $R^2/g\sigma$ and $g\sigma$ can be found. So much information about these quantities has been accumulating lately, that the day may not be far off when the theory may yield accurate values of the relative abundance of the elements forming chemical compounds in the stars, when further theoretical refinements may become desirable.

93. Abundance of H and N in Late-type Stars

The preceding theory may be used to determine the relative abundance of hydrogen and nitrogen in late-type stars, since hydrogen is known to be sufficiently abundant to play the part of the element A. It was pointed out by Swings and Struve† that it is possible to

† *Phys. Rev.* **40** (1932), 142.

estimate the relative abundance of two elements when they occur in combination with one and the same third element—for example, hydrogen and nitrogen both in combination with carbon in cyanogen and carbohydrogen (CH). This can, of course, only be done when CN and CH bands are found in the same spectrum. Swings and Struve's considerations were based on the equations of dissociative equilibrium directly, no account being taken of the effect of level considered above. The situation is the same in principle in the more refined theory, however, since expression (120) has the form of a product of μ_B into a factor depending on μ_A only, say

$$N = \mu_B F(\mu_A).$$

Hence it follows that for the ratio of the N's which correspond to the same B but two different A's, A_1 and A_2,

$$\frac{N_{A_1 B}}{N_{A_2 B}} = \frac{F(\mu_{A_1})}{F(\mu_{A_2})}. \tag{122}$$

If it is possible to measure the ratio of the N's occurring on the left-hand side, this equation gives for any selected value μ_{A_1} a definite value of μ_{A_2}. Expanding the F-expressions to the first power of μ_A, which is permissible for small band intensities, the ratio of the N's becomes directly proportional to the relative abundance μ_{A_1}/μ_{A_2}.

If a trustworthy theory connecting the N's with the band intensities were at hand, the above relation would afford a good way of attacking the problem of abundance. Neglecting the difference in absorption coefficient in the region of the two bands, some approach to the problem may be made even if this functional relation is not known. It is then sufficient in fact to look out for spectra in which the intensities of the two different band systems are equal. The corresponding N's must then be equal too, and relation (122) reduces to

$$F(\mu_A) = F(\mu_B).$$

After a careful survey of the material Swings and Struve found that the CH and CN bands both fade out in the same spectral class, F 8. The diverging opinion entertained by several earlier authors on this point is probably due to an erroneous interpretation of a batch of blended lines of atomic origin in the region of the G band. Assuming F 8 to correspond to a temperature of 6,500° K., it follows from the above formula in its simplified form that the abundance ratio of H and N is of the order 10^3. This is in general agreement with the

analysis of atomic line intensities. But Swings and Struve pointed out at the same time that the absorption coefficient in the region of

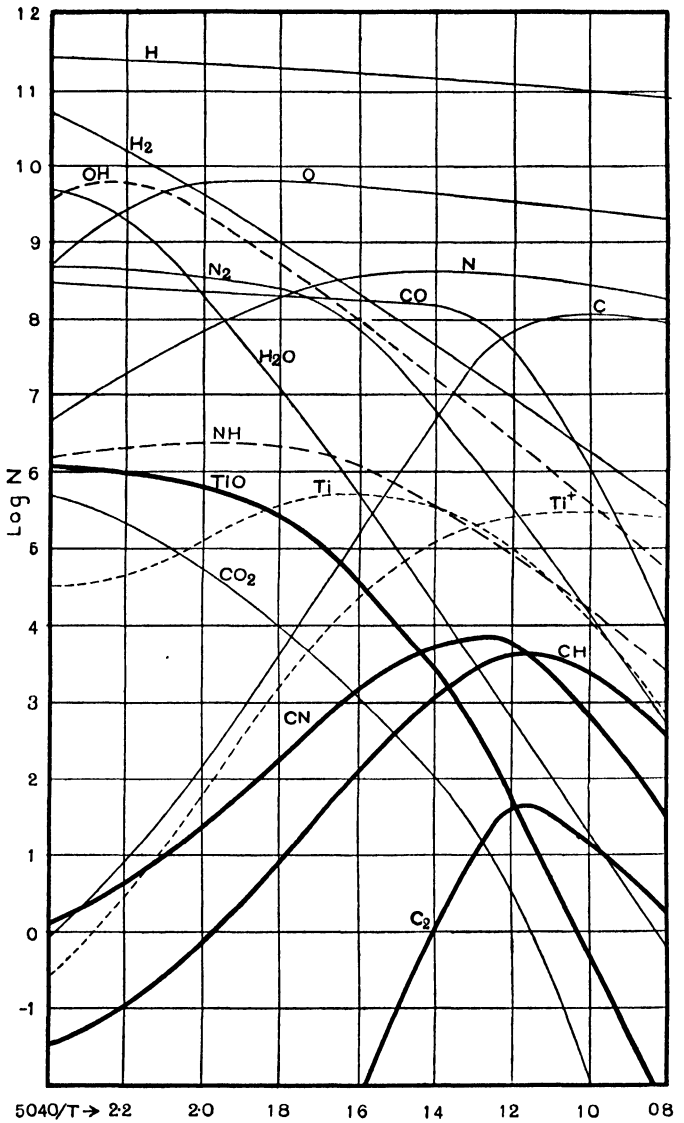


FIG. 27. Abundance of molecules and atoms in giant stars of the branch *K–M* (Russell, *Astrophys. J.* **79**). The heavy lines refer to molecular compounds actually observed. Thin lines give theoretical intensities of compounds which do not have observable bands in the accessible spectral region, and which hence cannot be observed directly. But their existence may be inferred from the presence of the related compounds.

the G band is likely to be higher than in the case of the CN bands, and that the result is affected by some uncertainty on this account.

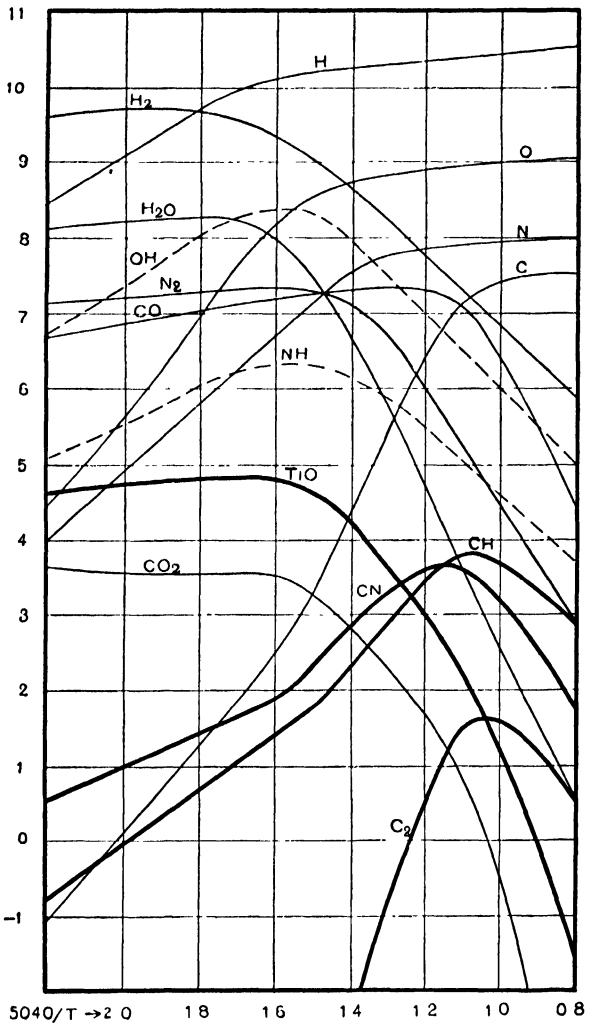


FIG. 28. Same as Fig. 27 except that the calculations refer to main sequence stars.

94. Carbon Compounds in Late-type Stars

Due to the large dissociation energy of CO, this compound dominates over all other carbon combinations, provided a sufficient supply of oxygen is available and the temperature is sufficiently low. This gives the clue to the understanding of the fact that C_2 and TiO

are mutually exclusive at low temperatures in the stars. It is of course possible that at temperatures of about 3,000°, where the TiO concentration begins to increase and the concentration of C_2 is

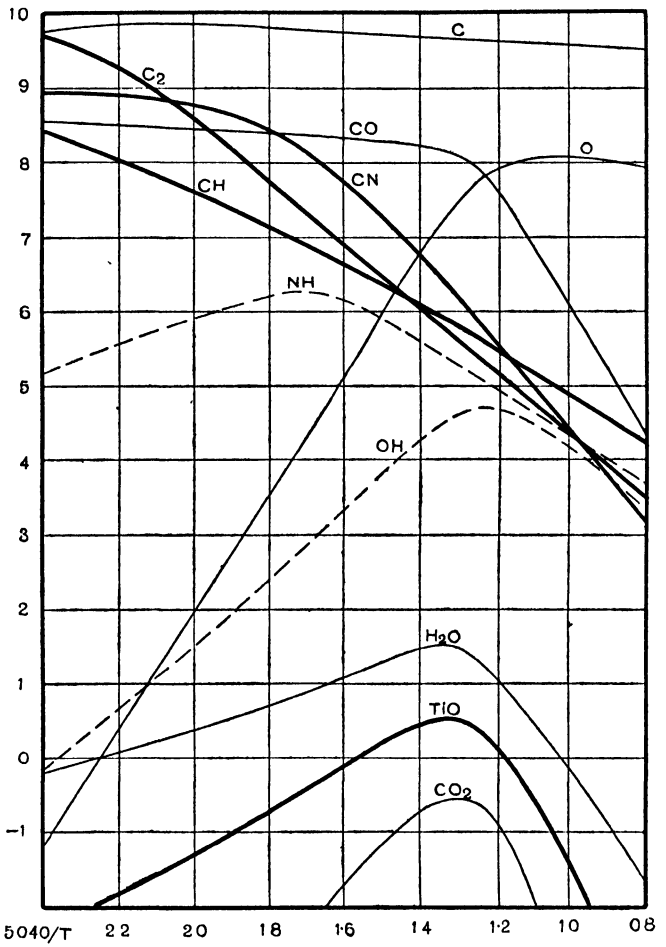


FIG. 29. Abundance of molecules in giant stars of the *R–N* branch (Russell, *Astrophys. J.* **79**).

falling sharply off, a star may show TiO as well as C_2. This seems actually to be the case of α Herculis.†

While thus CO must be present in large quantities in the atmospheres of some stars, it is a peculiar fact that in no case does it betray its presence in the spectra. In fact, by our formulae we should

† J. Stebbins, *Lick Obs. Bull.* **47** (1933).

expect CO to be about a million times more abundant than CN in low temperature dwarfs (cf. Fig. 28). The reason for this non-appearance of CO is the fact that the resonance bands all fall

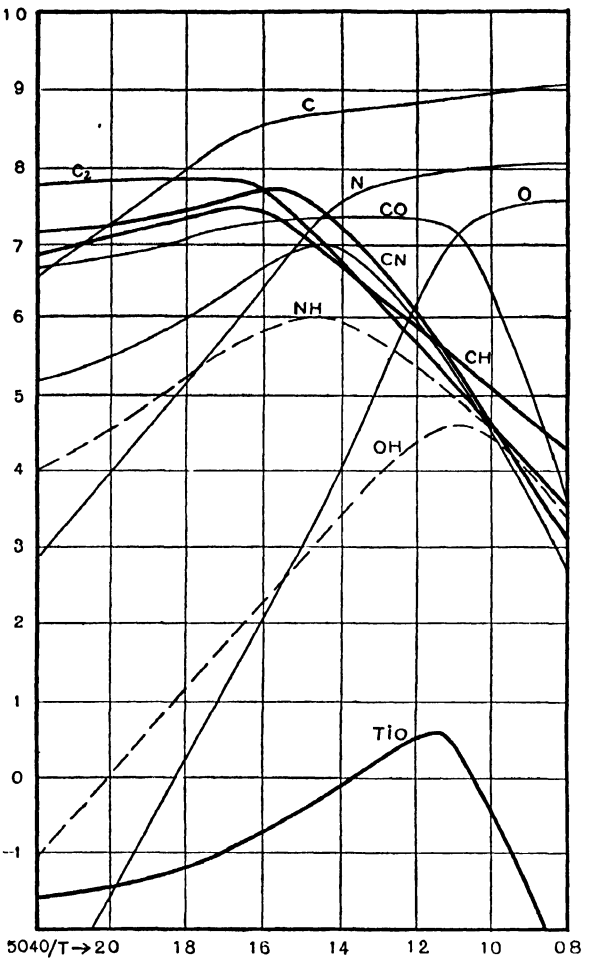


FIG. 30. Abundance of molecules in dwarf stars of the branch *R–N*. Cf. the preceding figure. (Russell, *Astrophys. J.* **79**.)

outside the visible region. The visible CO bands are primarily the Ångström bands, which have an excitation energy of about 8 volts. In the present temperature range this means, however, that only one molecule in every million will possess an excitation energy large enough for the absorption of these bands to be possible, so that they

should be several hundred times fainter than the CN bands, and this is sufficient to account for their absence.

95. Stellar Compounds in General

Russell† has carried out very extensive calculations on the state of dissociation of stellar compounds, based on his own determinations of the relative abundance of elements in stars. It would lead us too far here to enter into the detailed assumptions underlying his calculations, but it may be stated that in general they are similar to those used in this outline. Russell finds a fair agreement between theory and observation. In particular the excessive abundance of hydrogen is definitely confirmed.

96. Molecular Compounds in Planetary Atmospheres

1. The Earth. The terrestrial atmosphere is essentially made up of oxygen and nitrogen with some admixture of water vapour in the lower strata, and with traces of carbon dioxide, argon, and helium. Near the ground the constitution of the atmosphere may be found by direct chemical analysis, but for higher levels inferences must be drawn from spectroscopic observations. First of all there are the terrestrial lines and bands in the solar spectrum, which are recognized as such by increasing intensity with increasing optical path in our atmosphere, by the absence of a Doppler effect due to the sun's rotation, and, in the case of ozone, by the otherwise inexplicable petering-out of the solar spectrum near 2,900 A., and the varying intensity in this region with latitude and with meteorological conditions which point to a terrestrial origin. There is further the evidence from spectroscopic observation of the aurora borealis to be mentioned, and possibly also the observation of the zodiacal light.

The terrestrial bands of O_2 in the solar spectrum are represented by a strong band at 7,594, a weaker band near 6,867, and a still weaker band at 6,277 A. These bands are due to forbidden transitions, which explains why their intensities are so low in spite of the fact that the whole air mass of the atmosphere takes part in their formation. The principal bands of oxygen are situated in the ultra-violet beyond 1,800 A. and prevent any study of the solar or stellar spectra in this region. The intensity-distribution in the oxygen bands is such as to indicate a low average temperature of the

† *Astrophys. J.* 79 (1934), 317.

atmosphere, in accordance with the meteorological observations of the upper air.

The ozone absorption consists of some weak bands in the red and the orange, a series of stronger bands extending from 3,300 and towards shorter wave-lengths, the Huggins bands, and finally a strong absorption setting in at about 2,900 A. and screening off nearly everything beyond. This last absorption is connected with the breaking up of the ozone molecule into O_2+O, while the earlier isolated bands are of the ordinary rotational type.

This ozone is probably formed in the higher strata of the atmosphere, from 20 km. and up, by the ionizing influence of ultra-violet sunlight. The total amount of ozone corresponds to a layer of 0·2 to 0·4 cm. thickness at standard temperature and pressure. The detailed theory of this formation of ozone is as yet only in its infancy, and for this reason it will be passed over here.

Atomic oxygen does not exist in measurable quantities in the lower strata of the atmosphere. From 80 km. above the ground and up its existence is indicated by the green auroral lines 5,577 in the spectrum of the night sky and the polar aurorae. The identification of this line as belonging to atomic oxygen is due to McLennan.† It is due to a transition, which is forbidden by the ordinary selection rules. The closer details of its excitation in the upper atmosphere does not seem to be cleared up as yet. Recent work of Kaplan‡ on the production of this line in arbitrary sources shows that a trace of oxygen in an atmosphere of nitrogen, say 1 per cent. O_2 to 99 per cent. N_2, exposed to intermittent electrical discharges, will reproduce the general features of the auroral spectrum. This may be taken to indicate that nitrogen preponderates much more in the high strata of the atmosphere than near the ground where the $O_2:N_2$ ratio is nearly as 1:4.

Nitrogen does not reveal itself by any terrestrial lines and bands in the solar spectrum, but it shows up very strongly in the spectrum of the aurora borealis, the most conspicuous bands being those belonging to the so-called negative series which are due to N_2^+, while the positive bands, due to the neutral N_2 molecule, are generally weaker. The detailed condition of excitation of these bands in the auroral spectrum is still a matter of debate.

† *Proc. Roy. Soc.* A, **120** (1928), 327.
‡ J. Kaplan, *Nature*, **135** (1935), 229.

Water vapour absorption is confined to the long wave-length part of the spectrum, extending from about 7,200 A. to 18,000 A., causing very great trouble in the study of solar and planetary spectra in this region.

2. **Venus and Mars.** That these planets have atmospheres is evident from the presence of twilight on Venus, and the presence of clouds and the waxing and waning of the polar caps on Mars. The spectra of these planets are remarkably similar to the solar spectrum, and very great efforts have been made in order to discover individual features giving a clue to the chemical constitution of their atmospheres. This has thus far only succeeded for Venus, where Adams and Dunham† in 1932 discovered three bands due to carbon dioxide in the infra-red. These bands have their heads at about 7,820, 7,883, and 8,689 A. and had not up to that time been discovered in terrestrial sources. The identification was, therefore, first done from the identity of the moment of inertia given by the bands with that of the CO_2 molecule, and next Dunham succeeded in reproducing the bands by passing light through a pipe 40 m. long, containing CO_2 at a pressure of 10 atmospheres. Further experiments indicated that the layer of carbon dioxide above the visible surface of Venus is at least equivalent to a layer 3 km. thick at standard temperature and pressure, which is about twice as much as the amount of oxygen in the atmosphere of the earth. The weak intensity of these bands is due to the fact that they involve high harmonics of the fundamental vibration frequencies.

Very careful measurements by Adams and Dunham‡ in 1934 indicate that the atmosphere of Mars does not contain oxygen to a measurable extent. They estimate that an amount of oxygen equal to a thousandth part of that above an equal area on earth could have been detected. This applies also to Venus. The test was made by looking for the Doppler effect due to the relative motion of the planet and the earth.

3. **The Major Planets.** The atmospheres of the major planets form a problem of their own which gave no promise of a solution until a few years ago. Jupiter shows a conspicuous band in the orange, which was discovered visually by Huggins. There are fainter bands in the green region of the spectrum. These bands recur

† *Pub. Astron. Soc. Pacific*, **45** (1932), 243. ‡ *Astrophys. J.* **79** (1934), 308.

with increasing intensity in the spectra of Saturn, Uranus, and Neptune.

The identification of these bands remained obscure until 1932, when the puzzle was solved by Wildt,† who proved that they were due to ammonia and methane. Later work of Dunham‡ and Adel and Slipher‖ confirmed this result. Again, we find that the reason why the bands have not been identified before is due to the fact that they correspond to high harmonics of the fundamental vibration frequencies of the molecules, and as such show up with a very small intensity under ordinary conditions. By far the greater part of this absorption is due to methane, which Dunham estimated to be equivalent to a layer of 1·5 km. thickness for Jupiter, while the corresponding amount of ammonia does not amount to more than 10 m. (standard temperature and pressure).

† *Göttinger Veröff.* **22** (1932).
‡ *Pub. Astron. Soc. Pacific*, **45** (1933), 42; **46** (1934), 231.
‖ *Phys. Rev.* **46** (1934), 902; **47** (1935), 681.

XIX

THE SOLAR ENVELOPE

97. Extended Atmospheres

THUS far it has been assumed that the atmospheres are in equilibrium under the opposing influences of gas pressure and gravity. These assumptions seem to work well in the lower atmospheric strata, but in the highest atmospheric layers conditions seem to become rather different. For the sun this inference may be drawn from the complex phenomena exhibited by the chromosphere and the corona. For the stars, bright lines, which appear especially in early type spectra, indicate that the outer atmospheric envelope has assumed very much larger proportions than for the sun. Finally, for novae, planetary nebulae, and diffuse nebulae the nebulous envelope has outgrown the star or the stars involved in them completely, although there is still every reason to believe that the luminosity of the envelope is derived directly from the stars.

The study of this gradual growth of an envelope is of peculiar interest. From the point of view of physics it illustrates a transition from a state of local thermal equilibrium into a state in which the application of temperature considerations has become entirely ambiguous, and must be administered with the greatest care. At the same time the atoms forming the envelope get a better and better chance to display individual properties, which otherwise remain hidden behind the veil of thermodynamic equilibrium. From the point of view of the astronomer the transition from star to nebula is interesting because of evolutionary ideas, which always are present in his work, consciously or subconsciously. Earlier astronomers were apt to believe the stars born out of nebulae; modern astronomers are more inclined to believe the nebulae born out of the stars. But still much more work will be required before a final verdict on the matter can be pronounced.

98. Chromosphere and Corona

The study of the sun reveals to us the incipient stage in the formation of a nebulous envelope. This stage may, therefore, be studied in much greater detail than the subsequent development, leading over into the Be stage, the *P* Cygni stage, and the nova stage. It is also

possible, although not at all certain, that an intensive study of the solar envelope will reveal a sufficient indication of the hidden causes of this outflow of matter, which in some stars assumes such formidable proportions.

From observations during solar eclipses we know that the normal extent of the solar atmosphere is of the order of 200–400 km. This means that the majority of the monochromatic images of the solar limb photographed during eclipses indicate a height of this order.

Where this rule begins to fail, the outer envelope begins. The most notable exceptions to the rule are hydrogen, helium, calcium, and strontium, among which calcium attains, apparently, the greatest heights. On Mitchell's photographs of the flash spectrum, taken in Spain during the eclipse of 1905, the H and K lines are still noticeable at a height of 14,000 km. above the photosphere.†

There is now collected an extensive material bearing on the distribution of different elements with height in the chromosphere.‡ The general verdict of these investigations is that the density-distribution of an element is well represented by exponential laws. In some cases a single term applies. Thus Pannekoek and Minnaert give for the distribution of intensity in the hydrogen lines the formula

$$I = I_0 e^{-Kz/R},$$

where $K = 1{,}127$ and $I_0 = 62.10^4$ ergs/cm.^{2} sec. and R stands for the radius of the sun, while z is the height above the reversing layer. Mitchell states that in some cases a better result is obtained by using two exponential terms.

The observations give the projected intensity per cm.^{2} To pass on from that to the emission of light per unit volume in the atmosphere is easy if we are entitled to neglect reabsorption. Fortunately the optical thickness of the chromosphere does not seem to be large, so this assumption is permissible.

Let $E(r)$ denote the radiation emitted per unit time and volume in the required spectral line at a distance r from the centre of the sun, and denote by r_0 the perpendicular distance from the centre of

† Mitchell, *Astrophys. J.* **38** (1913), 407.

‡ Davidson and Stratton, *Mem. Brit. Astr. Ass.* **64** (1927), 105. Davidson, Minnaert, Ornstein, and Stratton, *Monthly Not.* **88** (1928), 536. Pannekoek and Minnaert, *Verh. Kon. Akad. Amsterdam*, **13** (1928), No. 5. Mitchell, *Astrophys. J.* **71** (1930), 1; **72** (1930), 146.

the sun to the line of sight. The projected intensity at r_0 is then

$$I(r_0) = 2\int_{r_0}^{\infty} \frac{E(r)r\,dr}{\sqrt{(r^2-r_0^2)}},$$

since an element ds of the line of sight is given by

$$ds = \frac{r\,dr}{\sqrt{(r^2-r_0^2)}},$$

on assuming spherical symmetry. This is an integral equation for $E(r)$, which was first solved by Abel.† The solution is found by writing

$$r^2 = u, \qquad r_0^2 = v,$$

multiplying the integral equation by $(v-w)^{-\frac{1}{2}}dv$, where w is a constant, and integrating from $v = w$ to $v = \infty$. Changing the order of integration, by writing the equation in the form

$$\int_w^{\infty} \frac{I(v)\,dv}{\sqrt{(v-w)}} = \int_w^{\infty} E(u)\,du \int_w^{u} \frac{dv}{\sqrt{\{(u-v)(v-w)\}}},$$

and noticing that

$$\int_w^{u} \frac{dv}{\sqrt{\{(u-v)(v-w)\}}} = \pi,$$

the solution follows by differentiation, in the form

$$E(w) = -\frac{1}{\pi}\frac{\partial}{\partial w}\int_w^{\infty} \frac{I(v)\,dv}{\sqrt{(v-w)}}.$$

This solution may be applied to the observed projected intensity in the chromosphere. Writing

$$v = w+x^2,$$

the following approximate expression of I will suffice

$$\log I = \log I_0 + K\left(1-\frac{\sqrt{w}}{R}-\frac{x^2}{2R\sqrt{w}}\right).$$

This gives to a first approximation

$$E(r) = \frac{I_0}{R}\sqrt{\left(\frac{K}{2\pi}\right)}e^{K\{1-(r/R)\}}.$$

This relation was applied by McCrea‡ to H_γ. The initial state of

† Cf., for instance, G. Kowalewski, *Integralgleichungen*, p. 9.

‡ *Monthly Not.* **89** (1929), 483.

H_γ is $n = 5$. Denoting the number per unit volume of hydrogen atoms in this state by x_5, the frequency of the H_γ line by ν, and the transition probability $5 \to 2$ by a_{52}, the emissivity is given by

$$4\pi E_\gamma = x_5 h\nu a_{52}.$$

The stimulated emission may be neglected in this spectral region, so that a_{52} stands for the probability of spontaneous transitions. Introducing numerical values for the constants McCrea found the value of x_5 at the bottom of the chromosphere to be 132. McCrea assumed a temperature in the chromosphere of 5,800°. Multiplying up by the right Boltzmann factors the numbers of atoms in states 2 and 1 were found to be 6,200 and $1{\cdot}03.10^{12}$ respectively.

The essential point in the above analysis is that the decrease of density of hydrogen with height is given by the same exponential factor as the projected intensity I, so we may write

$$\rho = \rho_0 e^{K\{1-(r/R)\}},$$

where ρ_0 is 16.10^{-13} gm./cm.^{3}, if we accept McCrea's analysis.

The same procedure cannot be applied unaltered to the corona, since the coronal spectrum does not show bright lines which may be used for this purpose. Close to the sun the coronal spectrum shows bright lines on a continuous background. Farther out it changes into a continuous spectrum with absorption lines, similar to the spectrum of the disk. The absorption lines are not much wider than in the solar spectrum,† which shows that the scattering of the light is produced by heavy particles and not by free electrons.

The intensity of the continuous radiation from the corona may be represented as a function of r, but there is considerable disagreement between different observers as regards the proper form of this function. Thus Turner‡ gave the formula

$$I = I_0\left(\frac{R}{r}\right)^6 \quad (I_0 = \text{const.}).$$

Becker,‖ on the other hand, preferred an expression of the form

$$I = I_0\left(\frac{R}{r-\alpha R}\right)^4 \quad (\alpha = \text{const.}),$$

† J. H. Moore, *Pub. Astron. Soc. Pacific*, **46** (1934), 298.
‡ *Pop. Astron.* **14** (1906), 548.
‖ *Mem. Roy. Astr. Soc.* **57** (1908), 51, appendix.

while Bergstrand† interpreted his observations on the basis of the formula

$$I = I_0\left(\frac{R}{r-R}\right)^2.$$

No exact agreement can be expected for observations made at different phases of the solar cycle, since the corona is known to vary greatly in form and extent during a cycle. Leaving the exact value of the exponent undecided, we may use the following expression for the projected intensity in the corona at a distance r from the solar centre

$$I = I_0(R/r)^m,$$

where I_0 and m are positive constants. Using this expression we find

$$E(r) = -\frac{R^m I_0}{\pi r}\frac{\partial}{\partial r}\int^{\infty}\frac{z\,dz}{\sqrt{(z^2-r^2)z^m}} = \frac{I_0}{\pi}K_m\frac{R^m}{r^{m+1}},$$

where K_m is a constant depending on m.

Let W denote the solid angle subtended by the sun at the point in the corona under consideration, so that

$$W = 2\pi\{1-\sqrt{(1-R^2/r^2)}\}.$$

Let further σ_ν denote the atomic coefficient of scattering in the corona, and I the intensity of photospheric radiation, both for frequency ν. Assuming the continuous emission from the corona to be due to pure scattering, the emissivity E must have the form

$$E(r) = \frac{\sigma_\nu I_\nu W n}{4\pi},$$

where n is the number of scattering atoms per unit volume. The density of scattering particles in the corona will consequently fall off with increasing distance according to the law

$$n = \frac{\text{const.}}{W r^{m+1}}.$$

At sufficient distances from the sun the solid angle W will be given approximately by $\pi R^2/r^2$ and the density will therefore ultimately fall off proportionately to $r^{-(m-1)}$.

99. Theories of the Chromosphere and the Corona

The exponential laws found to hold in the chromosphere might suggest that it is, for hydrogen, a simple question of hydrostatic

† *Études sur la distribution de la lumière dans la couronne solaire*, Stockholm, 1919; *Monthly Not.* **95** (1935), 436.

equilibrium with a nearly constant temperature. It is easy to show, however, that this cannot be correct. For in this case the constant K used in previous formulae must be given by

$$K = \frac{Rm_{\mathrm{H}}g}{kT},$$

where g is the acceleration of gravity, T the temperature, and m_{H} the mass of a hydrogen atom. Introducing numerical values

$$R = 695.10^8 \text{ cm.}; \qquad g = 27{,}360 \text{ cm./sec.}^2; \qquad T = 5{,}000^\circ \text{ K.},$$

which should not be far from the truth, we find

$$K = 4{,}600,$$

which is four times too large. If hydrogen were completely ionized, this figure might be reduced to half this value. But it is doubtful whether hydrogen is so strongly ionized that the molecular weight will be appreciably influenced. An application of the formulae of dissociative equilibrium indicates that the ionization of hydrogen at the base of the chromosphere cannot surpass 10 per cent., and is probably much smaller. At greater heights the ionization will be larger, but it is of no use following this line any farther since an unexplained factor 2 still remains. That hydrostatic conditions prevail, in the way it is understood ordinarily, is therefore out of the question.

As a first step towards a dynamic theory of the chromosphere McCrea† advanced the suggestion that the chromosphere is supported by turbulent motion. It would not be right to criticize this idea from the point of view of the narrow definition of turbulence in hydrodynamics. It was intended more as a first attempt at the introduction of dynamic concepts in this field, and to point out that the violent motions observed in the chromosphere, particularly in the form of prominences, cannot fail to have some connexion with the support of the chromosphere itself. And it raises the question whether the hydrogen chromosphere in reality is anything else than the net result of bubbles and outbursts from the solar surface, which in solitary instances assume the form of isolated prominences. Taken in this sense there is good reason for believing that McCrea was right in his suggestion, although it may be that the use of the term 'turbulence' in this connexion should have been avoided.

† *Monthly Not.* **89** (1929), 483, 718.

A closer study of the observations reveals, in fact, that the observations themselves cry out against the application of simple hydrostatic theories to the chromosphere and corona. Thus on all good photographs of the *corona* it is evident that it consists of a tangle of narrow streamers emanating from the solar surface, and stretching out into space without suffering any marked blurring in details. The projection of a three-dimensional structure of this kind on a plane tends greatly to efface the more intricate details. When the details are nevertheless so striking as is shown on most eclipse photographs, it seems that the corona is intrinsically more to be likened to a brush discharge than to a quiet fluid.

It is therefore evident that *all* matter composing the corona cannot be in equilibrium, but must be in one-sided motion, there being an outward current of so great rapidity that matter does not have the time to perform any appreciable side-motion. The presence of a magnetic field may also tend to prevent the transverse diffusion of matter. The fact that the largest cross-motion observed during an eclipse amounts at most to a few kilometres per second shows that the streamers may be nearly stationary in space. However, it tells us but little about the motion along the streamers, which must be investigated by different means. Moore† has found that the Fraunhofer lines in the coronal spectrum indicate an outward motion of the order 26 km./sec. The great changes in the corona observed from maximum to minimum solar activity, and the connexion found to exist between prominences and coronal streamers and arches, show that fairly rapid changes do take place also in the position of the coronal streamers.

It seems probable that a clue to the better understanding of the chromosphere is provided by the fact that it too shows a hairy structure, like a billowing meadow, thus disclaiming the idea that it is a kind of atmosphere approximately in hydrostatic equilibrium. The fibrous structure of the chromosphere may not be so generally recognized as that of the corona, but the evidence given by the most experienced eclipse observers seems to be convincing. Chromosphere, corona, and prominences would in that case form a complex of dynamic phenomena, the theory of which must be based on considerations of the expansive motion of matter moving away from the sun. This motion seems even to extend perceptibly into the main atmosphere.

† *Pub. Astron. Soc. Pacific*, **35** (1923), 335; **45** (1933), 147.

This is revealed by a progressive red-shift of all solar lines on passing from the centre to the limb of the sun (limb effect). The theory of this effect was given by McCrea and Mitra,† but the suggestion that a radial expansion is the cause of the limb effect is due to St. John.

The necessity of this conclusion is not so generally realized among astronomers as among the geophysicists working on the problem of aurorae and magnetic storms. That this is so is natural, since the geophysicists are directly concerned with the effect of particles expelled from the sun, while the astronomers deal with phenomena where the dynamic features are less obvious to the eye. Dynamic theories of the solar envelope have, therefore, also been first proposed and discussed by geophysicists. The corpuscular theory of aurora borealis was first proposed by Birkeland,‡ and was discussed mathematically in various ways by Störmer, who also outlined a dynamic theory of the corona|| based on the assumption that the sun is surrounded by a weak magnetic field, and that electrons, or charged particles in general, are continuously emitted from the solar surface.

The idea of 'streams of particles from the sun' as the direct cause of aurorae has been discussed by many geophysicists, among whom we may further mention Vegard and Chapman. An attempt to follow up the consequences of these views for the structure of the chromosphere and the corona has been made by the author.††

It would be premature to assert that a definite solution has been reached. What appears certain is that clouds of particles are emitted from the sun from time to time, and it cannot fail to be that this process must have some influence on the structure of the chromosphere, and most likely is responsible for the development of the corona. That magnetic forces come into play is proved by the form of the corona, but it would be going too far to assert that this field is a permanent attribute of the sun in the same way as this is the case for the terrestrial field. While this latter field has for the most part its origin in the interior of the earth, it may well be that the magnetic field in the corona is a sort of skin effect, or is generated in the corona itself.

It does not seem easy at the present moment to suggest a simple

† *The Observatory*, **57** (1934), 379.

‡ Cf. *The Norwegian Aurora Polaris Expedition, 1902–8*, by Kr. Birkeland, Oslo 1908, Chap. 6.

|| *Comptes rendus*, Feb. 20, March 6, 1911.

†† *Pub. Oslo Univ. Obs.* (1933), No. 5.

theory of the chromosphere and the corona which will cover the different facts in a satisfactory manner, and it is necessary to make an attempt at an isolation of particular facts which appear to be amenable to a quantitative treatment. A case of this sort may be offered by calcium in the chromosphere, which stands out in a special way among the rest of the chromospheric elements. The fact that the *H* and *K* lines are resonance lines strategically situated in an intense part of the solar spectrum led Milne† to propose the view that the Ca II atoms are supported by light pressure. Milne has worked out this idea at great length for the case of static equilibrium, and Chandrasekhar‡ has investigated how such a calcium chromosphere will work out if the intensity is not uniform over the solar disk, but is mottled in some way, which Chandrasekhar mimicked by assuming a periodic space-term in the intensity.

That light pressure should play a dominant part in the building-up of the chromosphere and the corona has been suggested from time to time; but it does not seem possible to construct a serious quantitative theory on such a basis, because hydrogen, the principal constituent of the chromosphere, has its resonance lines so far into the ultra-violet that it can only suffer a negligible light pressure in the solar atmosphere, *provided* the ultra-violet intensity in the solar spectrum follows a Planck law corresponding to a temperature of the order 6,000° K. The major cause of the development of the solar envelope is thus left unexplained. But although light pressure can at most give a clue to the understanding of the isolated facts of the calcium chromosphere in the sun, it may well happen that in early-type stars it will play a dominant part. And it has been suggested by Gerasimovic,‖ supported by quantitative arguments, that in Be stars light pressure must be taken into account in the theory of the extended hydrogen envelope. For planetary nebulae conditions are still more in favour of producing a strong light pressure on hydrogen, and may actually be responsible for the outward motion of the planetaries, which is indicated by the doubling of their emission lines. This idea has been advanced and advocated by Zanstra.†† It is therefore worth while to consider the theory of an atmosphere supported by light pressure.

† *Monthly Not.* **84** (1924), 354; **85** (1924), 111; **86** (1925), 8.
‡ Ibid. **94** (1933), 14; **94** (1934), 726.
‖ Ibid. **94** (1934), 737.
†† Ibid. **95** (1934), 84.

100. A Calcium Chromosphere Supported by Light Pressure

Let F_ν as usual denote the monochromatic flux of energy in the atmosphere and κ_ν the absorption coefficient. The force due to light pressure is then per unit volume

$$f = \frac{1}{c}\int_0^\infty \kappa_\nu F_\nu \, d\nu.$$

In Milne's theory the H and K lines of calcium are considered as a single line only, and other lines are neglected. It is then a question of pure scattering of light, and the flux F_ν is independent of optical height. The line profile is also simplified into a columnar form, so that

$$\int_0^\infty \kappa_\nu F_\nu \, d\nu = \Delta\nu \kappa_\nu F_\nu,$$

$\Delta\nu$ being the width of the line.

Let n_1 and n_2 denote the number of Ca II atoms in unit volume in the first and second quantum states respectively, ρ_ν the energy-density of radiation in the region of the line, and A_{21}, B_{21}, B_{12} the Einstein coefficients of the transitions $1 \rightleftarrows 2$, which then satisfy the relations

$$B_{12}/B_{21} = \sigma, \quad \text{and} \quad A_{21} = B_{21}\, 8\pi h\nu^3/c^3,$$

where σ is the ratio of the weights of the second and first states. We further write

$$n_1 B_{12}\rho_\nu = n_2(A_{21} + B_{21}\rho_\nu)$$

or

$$n_2/n_1 = \frac{\sigma\bar{\rho}_\nu}{1+\bar{\rho}_\nu}. \tag{123 a}$$

Since it is a question of pure scattering,

$$\bar{\rho}_\nu = \frac{1}{c}\bar{F}(2+3\tau), \tag{123 b}$$

where τ is the optical depth in the chromosphere, and a bar denotes multiplication by $c^3/8\pi h\nu^3$.

The momentum of the radiation emitted spontaneously vanishes in the mean, but the forced transitions contribute an amount

$$n_1 \frac{h\nu}{c^2} B_{12} F_\nu - n_2 B_{21} \frac{h\nu}{c^2} F_\nu = \frac{h\nu}{c^2} B_{12} F_\nu (n_1 - n_2/\sigma)$$

to the light pressure.

The force of gravity per unit volume is, on the other hand,

$M(n_1+n_2)g$, where M is the mass of a calcium atom, and the ratio of light pressure to gravity becomes

$$\eta = \frac{h\nu F_\nu B_{12}}{Mc^2 g}\,\frac{n_1-n_2/\sigma}{n_1+n_2}.$$

Consider next the equation of hydrostatic equilibrium

$$\frac{\partial p}{\partial z} = -M(n_1+n_2)g+\frac{h\nu}{c^2}F_\nu B_{12}(n_1-n_2/\sigma),$$

and let us expand the right-hand side for the case when n_2 is small in comparison with n_1, which is likely to be a tolerable approximation for the sun and similar stars of medium effective temperature. We then find

$$\frac{\partial p}{\partial z} = -\rho g\left\{1-\frac{h\nu F_\nu B_{12}}{c^2 Mg}\left(1-\frac{\sigma+1}{\sigma}\frac{n_2}{n_1}\right)+\ldots\right\},$$

where ρ is the density. Using (123 a) and (123 b) we find further, on neglecting the square and higher powers of $\bar{\rho}_\nu$,

$$\frac{n_2}{n_1} = \frac{\sigma\bar{\rho}_\nu}{1+\bar{\rho}_\nu} = \sigma\frac{\bar{F}_\nu}{c}(2+3\tau),$$

and finally

$$\frac{\partial p}{\partial z} = -\rho g\left(1-\frac{h\nu}{c^2}\frac{F_\nu B_{12}}{Mg}\left\{1-(1+\sigma)\frac{\bar{F}_\nu}{c}(2+3\tau)\right\}+\ldots\right). \qquad (124)$$

On the top of the chromosphere the pressure cannot increase with height, and can at most vanish; the chromosphere will then hold as much matter as it possibly can. The condition for the vanishing of the pressure gradient when τ vanishes is the following

$$\left\{1-(1+\sigma)\frac{\bar{F}_\nu}{c}\right\}\frac{h\nu}{c^2}\frac{F_\nu B_{12}}{Mg} = 1, \qquad (125)$$

which links up the net flux of radiation within the line with the absorption coefficient and the gravitational acceleration. In order to integrate the equation of hydrostatic equilibrium it is necessary to know the relation between z and τ. We therefore define an average absorption coefficient by the relation

$$\kappa = \rho\frac{h\nu}{c}\frac{B_{12}}{M\bar{\Delta}\nu},$$

on neglecting n_2/n_1 in comparison with unity. This means that

$$dz = -\frac{cM\bar{\Delta}\nu}{\rho h\nu B_{12}}d\tau, \qquad (126)$$

which, when introduced in (124), gives, on taking (125) into account,

$$\frac{1}{\tau}\frac{\partial p}{\partial \tau} = 3(1+\sigma)F_\nu \bar{F}_\nu \Delta\nu/c^2,$$

or integrated

$$p = \tfrac{3}{2}(1+\sigma)c^{-2}F_\nu \bar{F}_\nu \Delta\nu\tau^2. \tag{127}$$

Dividing this expression for the pressure by kT/M gives us the density, which introduced into (126) gives the relation between optical thickness and height. This relation can only be integrated when the dependence of temperature on height is known. For the present purpose it is sufficient to consider the case of constant temperature, which at once gives the integrated relation

$$\tau^{-1} = \frac{3}{2}\frac{(1+\sigma)F_\nu \bar{F}_\nu h\nu B_{12}}{c^3 kT}(z+z_0), \tag{128}$$

where z_0 is a constant of integration. A combination of (127) and (128) next gives the dependence of pressure on height:

$$p = \frac{16}{3}\frac{\pi c\nu\,\Delta\nu}{(1+\sigma)h}\left(\frac{kT}{\sqrt{F_\nu \bar{F}_\nu}\,B_{12}(z+z_0)}\right)^2.$$

The constant z_0 has the significance of a 'homogeneous height'. Milne finds for z_0 the value 1,800 km. for the solar chromosphere, on the assumption that $z = 0$ corresponds to the place where the specific intensity of the radiation in the combined H and K lines corresponds to the effective temperature of the sun according to the laws of thermal radiation.

Milne has refined the theory in various directions by taking account of the multiplicity of the H and K lines and the presence of the infra-red triplet $1D \rightarrow 1P$. He has also developed a theory of the velocity-distribution of the atoms which results from the absorption and emission processes in the H and K lines. According to his results the residual intensities in the H and K lines should stand in the same ratio as the corresponding weights.

We have seen earlier that the intensity-distribution within the Ca^+ lines is of the same type as for other lines in the solar spectrum. From the point of view of Milne's theory this must be considered a very remarkable thing, since the genesis of these lines differs radically from those of other lines. The residual intensity is to a first approximation given by

$$F_\nu = \frac{c^2 Mg}{h\nu B_{12}}, \tag{129}$$

and is thus solely determined by the Einstein B-coefficient and

gravity. According to calculations of Zwaan† it seems that the relation is not far from being true, although there are divergences in detail. It would be surprising, however, if these quantities could be adjusted in such a manner that the residual intensities of these lines should not differ essentially from those of other lines. This point may probably be cleared up without shifting the ground too much. For instance, we have assumed the calcium chromosphere to consist of Ca II atoms only. This cannot well correspond to facts, and the atomic mass M is therefore an average value depending upon the degree of ionization and the electronic mass. It may therefore be possible to interpret relation (129) by saying that the ratio B_{12}/M will adjust itself to correspond to the residual intensity which is normal for average absorption lines.

There is next the question of the stability of the chromosphere to be considered. The above calculations are based on a quite special state, in which the density-distribution is intimately dependent on the radiative flux. In fact, if the flux suddenly should fall essentially below the value demanded by (129), the chromosphere would suddenly contract and would merge with the ordinary reversing layer. Even when admitting the formal possibility of a state in which the equilibrium is maintained by radiation pressure, legitimate doubts may be entertained about its possible realization in nature.

These objections may be met, partially at least, by a consideration of the excitation processes which take place in the chromosphere. When supernumerary atoms in the chromosphere fall down towards the photosphere they will acquire a higher temperature, and hence become more transparent to the radiation in the region of the H and K lines, so that more radiation will be poured into the chromosphere at the base, and the upper part will experience a stronger light pressure. If the optical thickness of the layer is too small, the opposite process will take place, so that the limiting case considered may be a natural state of equilibrium. The ionization processes work in the same direction as the excitation processes.

The density found by Milne at the base of the chromosphere (10^6 Ca^+ ions per cm.^{3}) is so low that there must be present a much greater density of free electrons in order to prevent calcium from suffering a strong second-stage ionization. It was suggested by McCrea‡

† *Dissertation*, Utrecht, 1929. ‡ *Monthly Not.* **89** (1928), 483.

that these electrons were obtained by ionization of hydrogen, which is present in much greater quantities than calcium.

An alternate form of Milne's theory has been studied by J. Woltjer, Jr.,† who assumed the calcium chromosphere to have a slow outward drift, light pressure being assumed to be stronger than gravitation. That a certain fraction of the Ca^+ atoms are likely to suffer large outward accelerations was early realized by Milne.‡ If an atom, for example, acquires an outward velocity so large that it absorbs outside the absorption line proper, it will become exposed to the full photospheric radiation. According to Milne's calculations the limiting velocity which an atom may acquire in this way is of the order of 1,600 km. $sec.^{-1}$

Milne's theory of light pressure has been further developed by Chandrasekhar,‖ who considered the case when the photospheric flux varies periodically along the solar surface, and who also took account of the darkening towards the limb. In this way it was possible in a rough way to obtain some indication of the influence of a non-uniformly illuminated photosphere on a chromosphere supported by light pressure. As was to be expected, the result seems to be that a stable chromosphere is still a possibility, while at the same time the departures from a steady state encountered in the theory provide a formal possibility of bringing such sporadic phenomena as the generation of prominences within the realm of the theory.

101. A Chromosphere supported by the Emission of Corpuscular Rays

As emphasized before, it is difficult to believe that the solar chromosphere, apart from calcium, is supported directly by light pressure. Indirectly it may play some part, however, by producing swift particles, which in their outward course collide with gases not affected by light pressure. In this way the momentum imparted by light pressure may be transferred to atoms which are insensitive to light pressure themselves. If the light pressure on some selected gas is sufficiently strong, and the density in the envelope not too small, light pressure on calcium, for example, might entail a considerable distension also of inert gases like hydrogen. This is already sufficient reason for an investigation of the dragging effect of swift particles

† *Bull. Astron. Inst. Netherlands*, **167, 180, 182, 213,** and **262.** Also *Nature*, **129** (1932), 580.

‡ *Monthly Not.* **86** (1926), 459. ‖ Ibid. **94** (1933), 14; **94** (1934), 726.

emitted from the photosphere of a star exerted on inert gases in the atmosphere.

However, it is surely premature as yet to assert that we know enough about the stars to assert that light pressure is the only agency which can lead to the expulsion of particles. I think it is fair to say that the evidence of solar observations, taken as a whole, is against this idea, and to such a degree that it fully justifies an attempt at the development of a non-committal formal analysis of the solar phenomena from a wider point of view. For this reason we shall also consider briefly the main points of a theory in which an envelope is kept distended by the emission of corpuscular rays, which was first developed by the author.† The conditions in the envelope will turn out rather differently according as the emitted corpuscles are supposed to be electrically neutral or charged. If the escaping corpuscles are electrically neutral, conditions are simple enough, since the effect for strictly stationary conditions, spherical symmetry, etc., will mainly have the character of a diminution in gravity. It is therefore evidently possible to represent any envelope in this way.

Conditions are more involved when the escaping particles are negatively charged, and this case we shall treat at greater length. For this purpose we select as a typical model of such an envelope one consisting of two elements A and B say, of different atomic weight, say m and M, A being singly ionized throughout, while B is electrically neutral.

We denote the number of atoms A and B per unit volume by n_A and n_B respectively, and the corresponding number of free electrons by n_e. We consider throughout a stationary state of the envelope in which neutral atoms and free electrons remain in hydrostatic equilibrium, while the positive atoms perform a diffusive outward motion, so as to compensate the steady loss of negative particles expelled by non-thermal agencies. We might, of course, also have been considering the opposite case when the non-thermally expelled particles were positive. This would lead to an entirely different theory, in which ionized elements experienced no lifting force.

The number of positive particles generated (and expelled) in consequence of the loss of negative electricity may be written in the form

$$\operatorname{div}(n_A \mathbf{V}) = \epsilon \tag{130}$$

† *Publ. Oslo Univ. Obs.* **5** (1933). Cf. also J. Woltjer, Jr., *Nature*, **129** (1932), 580.

per unit volume, where ϵ is a positive quantity, which we need not specify any further for the moment, while $\mathbf{V}$ is the velocity of the positive ions.

Denoting the electric potential by ψ, the Poisson equation for this potential assumes the form

$$\nabla^2\psi = -4\pi e(n_A - n_e).$$

The outward drift of the ions will be resisted by collisions with free electrons and neutral matter. We represent this resistance by a force

$$-(n_B Q + n_e q)n_A \mathbf{V}$$

per unit volume, Q and q being coefficients proportional to the cross-section for collision of the neutral atoms and free electrons respectively, which may be derived by gas kinetic considerations.† In general q will be very small in comparison with Q.

The resistance experienced by the positive particles must result in an equally large lifting force on electrons and neutral atoms of magnitudes

$$qn_e n_A \mathbf{V} \quad \text{and} \quad Qn_B n_A \mathbf{V}$$

respectively. The lifting force of the primary electrons has been neglected, because very fast electrons are known to lose less energy per unit length of path than slow ones, about in the ratio of the inverse first power of the velocity.

We may now write down the hydrodynamic equations of the problem. The equation of motion of the positive particles will be

$$n_A \nabla(\tfrac{1}{2}mV^2) + \nabla p - n_A(m\mathbf{g} - e\nabla\psi) - (Qn_B + qn_e)n_A \mathbf{V} = 0;$$

for the neutral particles:

$$\nabla p_0 - n_B M\mathbf{g} - n_B Qn_A \mathbf{V} = 0; \tag{131}$$

and for the electrons:

$$\nabla p_e - n_e e\nabla\psi - n_e qn_A \mathbf{V} = 0.$$

Here p, p_0, and p_e denote the partial pressures of ions, neutral atoms, and free electrons respectively:

$$\frac{p}{n_A} = \frac{p_0}{n_B} = \frac{p_e}{n_e} = kT.$$

Differentiation with respect to the radius vector r is denoted by ∇. The mass concentrated in the envelope will be so small that the divergence of the acceleration of gravity may be neglected. Multiplying

† Cf. R. L. Rosenberg, *Zeits. f. Astrophys.* 8 (1934), 147.

equation (131) by kT/p_0 and taking the divergence, we then find, on using (130), the following equation for the determination of p_0:

$$\operatorname{div}\left(\frac{kT}{p_0}\nabla p_0\right) = Q\epsilon. \tag{132}$$

This equation brings out the lifting effect on the neutral particles in an interesting way. The right-hand term $Q\epsilon$ figures in this equation as the divergence of a repulsive field of force, acting on electrically neutral atoms, and which thus partly neutralizes gravity. Putting $Q = 0$ the equation reduces to the ordinary equation of hydrostatic equilibrium.

The above equation has a simple solution when the state is isothermal and $Q\epsilon$ is proportional to p_0, say

$$\frac{Q\epsilon}{p_0} = \frac{kT}{2}w,$$

where w is a new constant. The solution is then

$$p_0 = \frac{4\lambda^2 C^2 e^{\lambda z}}{(C^2 e^{\lambda z} - w)^2},$$

the layers being assumed to be plane parallel, z denoting the height, while λ and C are integration constants.

At the base of the envelope ($z = 0$) the pressure gradient must conform to the ordinary equation of hydrostatic equilibrium. This condition gives the equation

$$\frac{Mg}{kT} = \lambda\left(1 - \frac{2C^2}{C^2 - w}\right),$$

which gives one relation between the integration constants. As a second relation we may use a definite value of the pressure at the base, say P_0, which then gives

$$P_0 = \frac{4C^2\lambda^2}{(C^2 - w)^2}.$$

Solving for C and λ we now find

$$C = \alpha + \sqrt{(\alpha^2 - w)} \quad \left(\alpha = \frac{Mg}{kT\sqrt{P_0}}\right),$$

and

$$\lambda = \frac{Mg}{kT}\sqrt{\left(1 - \frac{wP_0 k^2 T^2}{M^2 g^2}\right)}.$$

These expressions show how the vertical structure of the atmosphere changes with increasing electrification. For $w = 0$ the atmosphere exhibits the usual exponential decrease of pressure. When w increases,

the pressure gradient decreases, and when w has reached the limiting value $M^2g^2/P_0k^2T^2$, the exponent vanishes, and the pressure becomes distributed according to a simple power law. Performing the passage to the limit in the expression of p_0 one finds

$$p_0 = \frac{4}{w}(z+z_0)^{-2} \quad (z_0 = 2kT/Mg).$$

The constant z_0 gives the homogeneous height of the envelope. Calculating this height when λ differs from zero we find

$$z_0 = \frac{1}{P_0}\int_0^\infty p_0\,dz = \frac{2}{C\sqrt{P_0}} = \frac{2kT}{Mg}\,\frac{1}{1+\sqrt{(1-w/\alpha^2)}}.$$

This equation shows how the homogeneous height increases from kT/Mg in case of no electrification ($w = 0$) to twice this value when $w = \alpha^2$.

From the above solution it is possible to build up a composite solution corresponding to a discrete set of values of w. Thus when the generation of positive ions is concentrated in a thin layer at the base of the atmosphere, the pressure p_0 is easily found to be given by

$$p_0 = Ae^{-\frac{M\phi+QKa/r}{kT}} \quad (A = \text{const.}).$$

Here $M\phi$ is the gravitational potential energy of an atom, r is the distance to the centre, and K a constant in the equation of continuity of positive ions

$$r^2n_A V = K.$$

For sufficiently small values of QK this gives an ordinary exponential decrease of pressure with height, in the limit when QK/r^2 is equal to the gravitational acceleration the pressure is uniform, and for still greater values of the constant the pressure will increase with height. This only means that for larger values of the constant the gas will no longer remain in hydrostatic equilibrium, but will acquire an outward velocity and become expelled from the star.

102. Influence of Magnetic Fields

The study of the corona during the last century has revealed that at times of minimum solar activity its long streamers show a remarkable flattening towards the solar equator. This feature can scarcely be interpreted otherwise than as a direct effect of magnetic fields.

From the work of Hale mentioned earlier (§ 75) it is known that

the sun is surrounded by a general magnetic field of axial symmetry, the axis being nearly parallel to the axis of rotation. Observations indicated, however, that the intensity of this field decreased rapidly outwards in the solar atmosphere, and it may be premature to assert that it is this field which extends into the outer corona, and influences the direction of the coronal streamers.

For magnetism to play any part in the motion of a gas it is first of all necessary that the gas is ionized and already in motion. The presence of a magnetic field may then alter the distribution of density considerably, as was pointed out by Cowling,† in connexion with solar theories. The case considered by Cowling was preferably that of a rotating star, but similar effects are also met with in the case of radial expansions.

Consider the simple example of a star with an axially symmetrical magnetic field, charged up to a certain electrical potential, and surrounded by an envelope of slowly varying density. Consider the motion of an electric particle in these crossed electric and magnetic fields, account being taken of the deceleration by the gases in the envelope.

Let $-\kappa\mathbf{V}$ be the resistance due to collisions of the particle, and assume the motion to be so nearly uniform that the acceleration term plays no part beside the forces. Let $\mathbf{E}$ and $\mathbf{H}$ denote the electric and magnetic field intensities, and e the charge of the particle. The equation of motion is then

$$\kappa\mathbf{V} = e\mathbf{E} + \frac{e}{c}\mathbf{V}\times\mathbf{H}. \tag{133}$$

The solution of this equation has the form

$$\mathbf{V} = \lambda\mathbf{E} + \mu\mathbf{E}\times\mathbf{H} + \nu\mathbf{H},$$

where λ, μ, and ν are constants, which may be determined by introducing this expression into the equation of motion. We then find first

$$\mathbf{V} = \frac{e}{\kappa}\mathbf{E} + \lambda\frac{e}{\kappa c}\mathbf{E}\times\mathbf{H} + \mu\frac{e}{\kappa c}(\mathbf{E}\times\mathbf{H})\times\mathbf{H}.$$

But by ordinary vector theory

$$(\mathbf{E}\times\mathbf{H})\times\mathbf{H} = (\mathbf{E}.\mathbf{H})\mathbf{H} - \mathbf{H}^2.\mathbf{E},$$

in consequence of which the above expression of $\mathbf{V}$ becomes

$$\mathbf{V} = \left(1 - \mu\frac{1}{c}\mathbf{H}^2\right)\frac{e}{\kappa}\mathbf{E} + \lambda\frac{e}{\kappa c}\mathbf{E}\times\mathbf{H} + \mu\frac{e}{\kappa c}(\mathbf{E}.\mathbf{H})\mathbf{H}.$$

† *Monthly Not.* **90** (1929), 140.

Comparing this expression of **V** with the one first given (133), we find at once

$$\lambda = \left(\frac{\kappa c}{e}\right)\mu = \left(\frac{\kappa c}{e}\right)^2 \frac{\nu}{\mathbf{EH}} = \frac{e/\kappa}{1+e^2H^2/\kappa^2c^2}.$$

This solution shows how the velocity of the electron is composed of three components, which we have denoted by $\lambda\mathbf{E}$, $\nu\mathbf{H}$, and $\mu\mathbf{E}\times\mathbf{H}$ respectively. The first component represents a radial motion, which we assume to be directed outwards. The second component, $\nu\mathbf{H}$, represents a tendency of the particle to move along the magnetic lines of force. Assuming **E** to be directed radially, and **H** to correspond to the field of an elementary magnet, the coefficient ν of this component will be positive in one hemisphere and negative in the other, and thus represents in both cases a tendency of the path to bend towards the magnetic equator. That the tendency is towards the equator and not away from it is ensured by the fact that (**EH**) is positive in the hemisphere where the magnetic field intensity is directed towards the equator. If $e\mathbf{E}$ were directed inwards, the paths would concentrate near the poles. This would correspond to an influx of particles from without.

The stream lines of these two components thus show the same general features as the long coronal streamers during sun-spot minimum.

The third component $\mu\mathbf{E}\times\mathbf{H}$ is directed perpendicularly to the meridian planes. As it preserves the same sign throughout it represents a tendency for the stream lines to spiral round the sun. Whether a tendency of this kind is present or not in the coronal streamers is not easy to say, because of the difficulties introduced by the projection of the three-dimensional curve system into a plane.

Consider next how the magnetic force may influence the quasi-static state of a chromosphere consisting of an ionized gas. The neutral component may be neglected in this connexion. We allow for a slight difference in velocity of rotation of ions and electrons, and write the equations of motion of ions and electrons separately, as

$$\frac{1}{n_A}\nabla p - \left\{m\mathbf{g} - e\nabla\psi + \frac{e}{c}\mathbf{V}\times\mathbf{H}\right\} + n_e q(\mathbf{V}-\mathbf{V}_e) = 0,$$

$$\frac{1}{n_e}\nabla p_e + \left\{-e\nabla\psi + \frac{e}{c}\mathbf{V}_e\times\mathbf{H}\right\} + n_A q(\mathbf{V}_e-\mathbf{V}) = 0.$$

Adding these equations we find

$$\frac{1}{n_A}\nabla p+\frac{1}{n_e}\nabla p_e+\frac{e}{c}(\mathbf{V}_e-\mathbf{V})\times\mathbf{H}+(n_A-n_e)q(\mathbf{V}_e-\mathbf{V})-m\mathbf{g}=0. \quad (134)$$

Since the gas must be electrically neutral, practically, we may write

$$n_A\sim n_e \quad \text{and} \quad \frac{1}{n_A}\nabla p+\frac{1}{n_e}\nabla p_e\sim\frac{2}{n_A}\nabla p,$$

and equation (134) becomes

$$\frac{1}{n_A}\nabla p=\frac{m}{2}\mathbf{g}-\frac{e}{2c}(\mathbf{V}_e-\mathbf{V})\times\mathbf{H}.$$

In the equatorial regions, where $\mathbf{H}$ is sensibly perpendicular to $\mathbf{V}$, it is thus possible by a suitable choice of relative velocity $\mathbf{V}_e-\mathbf{V}$ to make the pressure gradient vanish, or even reverse its sign. This means that the magnetic force may support an extensive equatorial ring of matter by postulating a suitable relative velocity of positive and negative particles. But this is only strictly true when the relative motion reduces to a simple rotation.

In the discussion of solar magnetic fields (§ 75) it was shown that the observations demand such a current system to be present in the lower atmospheric strata, and it is of interest to ask what mechanical effects this current may have.

We write then for the current density

$$J=n_A e(V-V_e)=-\frac{c}{4\pi}\frac{\partial H'}{\partial r},$$

to a sufficient approximation (cf. § 75), H' denoting the horizontal component of H, which is supposed situated in the meridian plane. Introducing this value of J in the equation of hydrostatic equilibrium it assumes the form, p denoting now the total pressure,

$$\frac{\partial p}{\partial r}=-\rho g-\frac{1}{8\pi}\frac{\partial}{\partial r}H'^2.$$

Integrating from a lower level 1 to a higher level 2 we find

$$p_1-p_2=M_{12}g-\frac{1}{8\pi}\{H_1'^2-H_2'^2\},$$

an equation first given by Ferraro.† Here M_{12} is the mass per cm.^{2} in a vertical column extending from level 1 to level 2.

As was noted by Ferraro, these equations can be used to draw

† *Monthly Not.* **95** (1935), 280.

interesting conclusions with regard to the minimum density and minimum mass per cm.^{2} above the photosphere. The minimum density consistent with a decrease of pressure with height is evidently

$$\rho = -\frac{1}{8\pi g}\frac{\partial H'^2}{\partial r},$$

and the minimum mass per cm.^{2} of the photosphere is

$$M = \frac{1}{8\pi g} H_p'^2,$$

H_p' being the value of H' in the photosphere. In the equatorial region H_p' must be of the order 150 gauss (§ 75). Using this value and assuming an average molecular weight of 1·5 we find the total number of atoms above the photosphere to be

$$N \geqslant 1{\cdot}4 . 10^{22}.$$

The minimum density in the reversing layer is more difficult to evaluate. But assuming an exponential drop of H' it appears that the minimum density at the bottom of the reversing layer is of the order 10^{-8} gm./cm.^{3} Ferraro obtained a value 9 times smaller, owing to his using a value $H_p' = 50$ gauss. It appears, therefore, that magnetic fields should probably be taken into account also when considering the vertical distribution of matter in the solar atmosphere.

103. Kiepenheuer's Theory of the Corona

Quite recently Kiepenheuer† has suggested a new theory of the corona, which is intrinsically different from all those mentioned previously, although it bears resemblance to some of them. Although it appears a little premature as yet to take a definite stand as regards its intrinsic value, it contains so many interesting possibilities that it should be mentioned here.

Kiepenheuer does not postulate a permanent magnetic field in the corona, in which the electrical particles emitted from the sun are moving. On the contrary, he assumes the permanent solar magnetic field to be very nearly perfectly screened off quite close to the solar surface. However, when an ionized cloud of gas passes rapidly through such a strongly inhomogeneous field, the cloud must become electrically polarized. But outside the solar atmosphere there is no longer any agency tending to maintain this polarization against the mutual electric attraction of the free particles. The return of the cloud to

† *Zeits. f. Astrophys.* **10** (1935), 260.

an electrically neutral state is, however, necessarily accompanied by the generation of a magnetic field, and this field will last as long as there exists a residual polarization of the cloud.

The point made by Kiepenheuer is that if the dimensions of the cloud are sufficiently large, and of the order to be postulated for solar clouds, several hours will be required for a cloud, once polarized, to return to its electrically normal state. This means that it will be magnetic during the whole time of its transit through the corona. Since all clouds must be polarized in the same sense, they will exert a magnetic attractive force on one another, and it is to this magnetic interaction Kiepenheuer appeals in order to explain the equatorial preference of the long streamers during minimum solar activity.

It is a strong point of this theory that it gives a natural explanation of the difference in form between the maximum and minimum types of the corona. It is, namely, natural to suppose that the velocity of expulsion of the clouds during the time of low activity will be smaller than in times of maximum activity. The magnetic forces might therefore have more chance to act on the cloud in the former than in the latter case. The trajectories calculated by Kiepenheuer from his theory for different values of the velocity of expulsion reproduce very suggestively the characteristic maximum and minimum forms of the corona.

XX

STARS WITH EXTENSIVE ENVELOPES

104. Envelopes of Giant Stars

THERE is reason to believe that the development of envelopes analogous to the solar chromosphere and corona is more pronounced in giant stars than in dwarfs. This is revealed in various ways. Since calcium has been recognized as a typically chromospheric element, it is suggestive to find that the calcium lines show a characteristic strengthening with increasing luminosity. This increased absorption is to be expected on any chromospheric theory, and is, in particular, accounted for by assuming the calcium chromosphere to be supported by light pressure. This follows from the relation found for the average flux of radiation in the lines, which was found to be proportional to the acceleration of gravity.

There are also other elements whose absorption lines behave in such a way as to suggest that the absorption takes place in a tenuous envelope rather than in the atmosphere proper, where the majority of the lines originate. In some stars it is found that although the majority of the lines are widened in such a way as to suggest a rapid rotation of the star, other lines are quite narrow. These latter may then originate at high levels in a quiescent and tenuous envelope. Stars like o Aqr, ϵ Cap, and ϕ Per are of this type. It is also fair to mention 17 Leporis in this connexion. The spectrum of this star has been interpreted by Struve† by assuming the star to be surrounded by an expanding shell of gas, which is responsible for narrow absorption lines of Fe II, Ti II, Fe I, Sc II, etc., in the spectrum.

An obvious way of testing out this hypothesis would be to study the spectrum of eclipsing binaries carefully during the beginning or the ending of the eclipse. If the eclipsed star is of early type, and the eclipsing star a diffuse giant of later type, the chromosphere of the latter might produce narrow absorption lines in the spectrum of the former.

A spectacular instance of this sort was shown by the eclipse of ζ Aur in the fall of 1934. In this case the primary component is a $K5$ supergiant and the secondary a $B2$ star. The elements of the

† *Astrophys. J.* **77** (1932), 85.

system are, according to Hopmann,

$R_B = 2{\cdot}4\,\odot$; $a_B = 7.10^9$ km.; $M_B = 12\,\odot$; $d_B = 0{\cdot}9$ gm./cm.$^{-3}$;
$R_K = 355\,\odot$; $a_K = 3.10^9$ km.; $M_K = 28\,\odot$; $d_K = 0{\cdot}8.10^{-6}$ gm./cm.$^{-3}$.

The period is 973^d. The $B2$ star went into eclipse on August 24, and came out again about October 2. While the time from outer to inner contact of the stars is only about $0{\cdot}86^d$, it was found on spectrograms

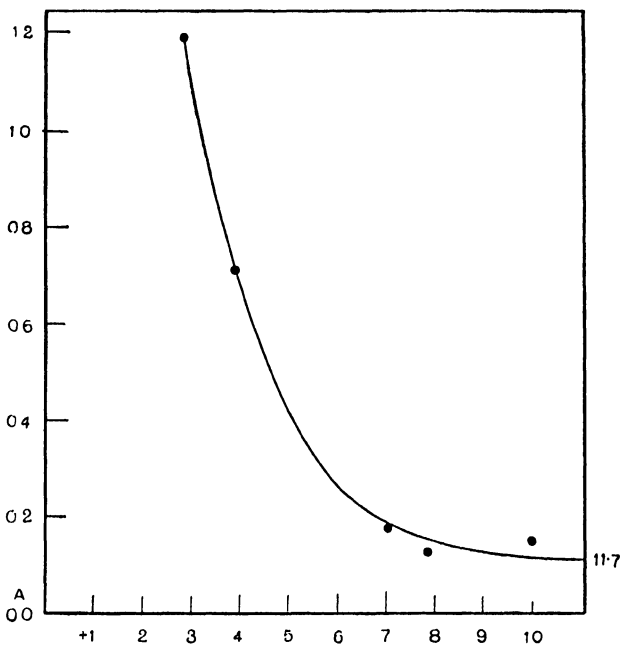


FIG. 31. The intensity of the Ca^+ K line as a function of phase after emergence from total eclipse in ζ Aur (Beer, *Monthly Not.* **95**). Abscissa is the time after ending of the eclipse in days. Ordinates give the intensity in arbitrary units.

obtained at the Cambridge Solar Physics Observatory that sharp H and K lines persisted in the spectrum of the B star about a whole week after the end of the main eclipse. These lines are obviously due to absorption in a very extended calcium envelope surrounding the supergiant $K5$ star.†

The behaviour of the hydrogen lines in stellar spectra suggests similar conclusions. In fact, the hydrogen lines are also strengthened in absorption with increasing luminosity in late-type stars. It is, on

† Cf. A. Beer, *Monthly Not.* **95** (1934), 24, and also William H. Christie and O. C. Wilson, *Astrophys. J.* **81** (1935), 213.

the whole, very remarkable that these lines show up at all in stars of so low effective temperature as, say, Betelgeuse or Antares; but there is scarcely a single late-type giant known which does not betray the presence of hydrogen lines. In late-type dwarfs the lines are either very weak or wholly absent.

It would probably be too much to say that hydrogen is a chromospheric element in the sun as far as the absorption lines are concerned. When we nevertheless feel inclined to attribute the strengthening of the lines with increasing luminosity to the development of an envelope, it is because it is not possible to explain the great strength of these lines in late-type spectra on the basis of the thermodynamic theory of the atmospheres, which ordinarily works well for most elements, as was shown earlier in this book.

105. Bright-line Stars

This impression is further strengthened when passing on to the behaviour of the hydrogen lines in early-type stars. In many cases the hydrogen lines are here found to be reversed into emission lines. This is an indication that the envelope has outgrown the star to such an extent that the selective emission from the envelope has become comparable in strength with the emission from the photosphere of the star in the corresponding spectral regions.

The following table, given by R. H. Curtiss,† shows the frequency of bright-line stars among early types. The bright lines belong for

TABLE 9

Class	*No. of H.D. Catalogue stars*	*Emission-line stars*	*Ratio*
O 5–*O* 9	55	7	1 : 8
B 0–*B* 5	1,996	136	1 : 15
B 8	1,604	13	1 : 123
B 9	2,752	3	1 : 917
A 0	6,320	1	1 : 6,320

the most part to hydrogen. But other elements also occasionally appear in emission. This is the case of He I and He II, Fe I and Fe II, and some other metals. The bright hydrogen lines in *B* spectra are usually superposed on wide and shallow absorption lines, indicating a rapid rotation of the stars. They are of three types. Either the

† *J. Roy. Astron. Soc. Canada*, **20** (1926), 23.

emission is in a single, fairly narrow line at the centre of the underlying absorption line. Or the emission line is double, symmetrically centred with respect to the underlying absorption. These are the typical forms of hydrogen lines in *B* spectra. The third type is exemplified by the star *P* Cyg, where the absorption is found on the violet side of the emission lines, indicating a relation to novae.

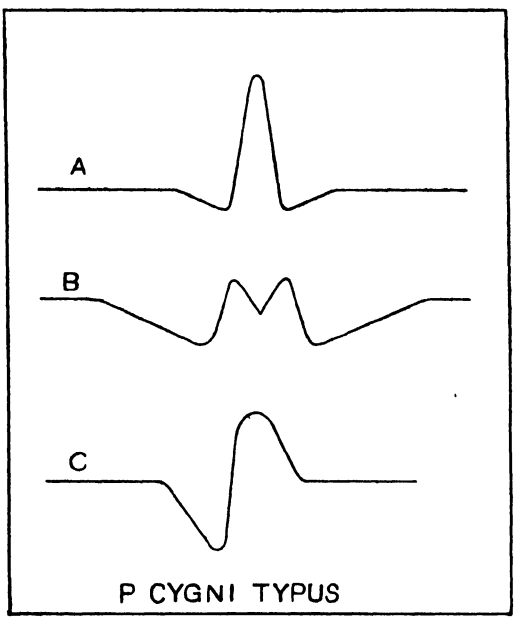


FIG. 32. Typical profiles of bright lines in *B* spectra (Merrill, Humason, and Burwell, *Astrophys. J.* **61**). Types *A* and *B* are typical for bright-line *B* stars, the bright line being superposed on a very broad and shallow absorption line. Type *C*, on the other hand, is typical for stars of the *P* Cyg type, which are closely related to novae.

The widths of the emission lines increase with increasing wavelength. R. H. Curtiss† gave the following formula for this relation:

$$\Delta\lambda = 0{\cdot}000628(\lambda - 3{,}270)(H_\beta - 2{\cdot}61) + 2{\cdot}61.$$

Units are angstroms, and H_β stands for the width of H_β, defined as the distance between the steepest outer gradients.

The intensity of the central emission lines decreases rapidly with increasing order of the line. It is by far the strongest for H_α. The higher members of the Balmer series appear usually like normal absorption lines, occasionally with bright borders.

† *Pub. Univ. Obs. Michigan*, **3** (1923), 1.

There is a general tendency for the lines to be variable in some way or other. In binaries like β Lyr or o Per the variations of the bright lines have the same period as the revolution of the stars. An interesting case of periodicity was discovered by Merrill,† where the violet component of the H_β emission varied periodically, while the red component remained unaltered. (Other lines were not investigated.)

The variations are mostly aperiodic in character. The bright lines in Pleione disappeared some years ago, and similarly in μ Cen and f^1 Cyg. In the latter star the bright lines were present in 1895, they were definitely absent in 1912 and 1918, just barely visible in 1920, and definitely present in 1927.‡

In some stars the emission lines exhibit a very complicated fine structure. In the stars γ Cas, β Lyr, and v Sgr the wings of the broad absorption lines are split up into components, symmetrically spaced about a central frequency. The scale of the pattern is the same for all hydrogen lines when measured in terms of a hypothetical Doppler velocity.‖

Passing on to stars of earlier classes, the emission lines change considerably in character. The duplicity, so characteristic of the lines in *B* spectra, disappears. In *O* stars the lines are without particular peculiarities. Besides hydrogen here also He I and He II lines appear in emission. In fact the lines of ionized helium were found for the first time in the *O* star ζ Puppis by E. Pickering (1896).

In *WR* stars (Wolf-Rayet stars) the emission lines widen out into broad bands, many angstroms in width, which only recently were recognized as being due to true atomic lines. Besides hydrogen a number of other elements are represented,†† first of all He, C, N, O, Al, Si, S, and Si III, all in high stages of excitation. The lines show a peculiar structure. Some of them show flat tops with steep slopes, but profiles of that form are not the rule. Usually the profiles are rounded, unsymmetrical, and occasionally with a central reversal.

The next link in the chain is provided by the novae, where the lines of hydrogen and also other elements are broadened to widths which measured in Doppler velocity may amount to more than three thousand kilometres per second.

† *Phys. Rev.* **25** (1925), 717.

‡ R. H. Curtiss, *Monthly Not.* **88** (1928), 205.

‖ Cf. J. S. Plaskett, ibid. **87** (1926), 31; *Pub. Dom. Astrophys. Obs. Victoria*, **4** (1927), No. 4; and also R. H. Curtiss, *Monthly Not.* **88** (1928), 203.

†† Cf. C. H. Payne, *Zeits. f. Astrophys.* **7** (1933), 1.

That in all these cases it is a question of different stages in the formation of tenuous envelopes round the stars is clearly shown by the association of *B*, *O*, and *WR* stars with planetary and diffuse nebulae, and also the fact that novae in the latest stages are surrounded by diffuse nebulosities.

106. Rotation Hypothesis for *Be* Stars

It was first pointed out by O. Struve† that the absorption lines in *Be* stars are of the type which indicates rapid rotation of the star. When the emission lines are double and the components widely spaced, the absorption lines are especially broad and shallow. When the emission is in a single line, the absorption is more of the usual type. Struve suggested therefore that the formation of the envelope round the *Be* stars, in which the bright lines originate, is brought about or very much facilitated, by a rapid rotation of the star.

The duplicity of the bright emission components suggests that the lines originate in a rotating gaseous shell or ring. The violet component is then supposed to be produced in the part of the shell advancing towards the earth, while the red component originates in the receding part of the shell. This idea of a rotating gaseous ring is not *per se* so improbable as it may look at first sight. We have seen in a preceding section (§ 101) that magnetic forces may have a strong tendency to support ionized gases in the magnetic equatorial plane, but much less in the vicinity of the poles. And Jeans‡ has shown that rotating rings of gas under the influence of light pressure and centrifugal force are in a certain sense possible configurations. This side of the question may therefore be left aside provisionally, and the observations analysed in a formal way, in order to see if a rotating ring or shell may be made to fit the facts.

In the section on rotating stars it was shown that a closed shell of uniform emissivity will give flat, dish-shaped emission lines without central depressions. However, a detached shell would rather appear brightened towards the limb. An illustration of such an intensity distribution is obtained by using the ordinary formulae for the intensity distribution over a stellar disk, but with a negative coefficient of darkening towards the limb, $-v$ say. The simplified solution for rotationally broadened profiles given on p. 205 then

† *Astrophys. J.* **73** (1931), 94. Cf. also Swings and Struve, ibid. **75** (1932), 161.
‡ *Monthly Not.* **83** (1923), 481.

assumes the following form

$$O(v) = \text{const.}\left(1+v-\frac{\pi}{4}\frac{v\eta}{a}\right)\eta \qquad \left[\eta = a\sqrt{\left\{1-\frac{c^2}{a^2W^2}\left(\frac{\nu-\nu_0}{\nu_0}\right)^2\right\}}\right],$$

where a is now the radius of the envelope and W its angular velocity. For $v \gg 2/(\pi-2)$ this profile will show two maxima near $\eta = 2a/\pi$. The solution for a narrow equatorial ring, on the other hand, will be nearly linear in η, and will therefore have the intensity maxima at the edges of the profile ($\eta = 0$),

A central depression of the profile will also arise from the fact that the star screens off some light from the back side of the luminous shell or ring. This effect will obviously be most important for ring-formed envelopes.

For the 'occultation effect' to be fully developed the inclination of the plane of the ring to the tangent plane of the sky must be near to 90°. For decreasing inclinations, the star will eclipse less and less of the ring, which means that the central depression must tend to fill out. This is fairly well in accordance with observation, and indicates that this effect may be of importance.

But there are other points where the simple rotation hypothesis seems to fail. Thus Gerasimovic† has pointed out that the observed increase in the intensity on passing from narrow to broad emission lines‡ goes against the idea of a rotating ring. In fact the width of the line should be a simple function of the inclination of the ring, and there is no loop-hole for a systematic dependence of intensity on inclination, apart from what is due to the screening of the star.

The perfect symmetry of the emission profile demanded by the rotation hypothesis is another weak point, which at once leaves the *P* Cygni stars out of the problem.

The rotation hypothesis in the rather restricted form formulated by Struve may perhaps need reconsideration also in other respects, as he considered the rotating ring to be the product of a rotational breaking up of the star in the sense of Jeans's theory. But even if the special model of a rotating ring should have to be dropped, the idea may still prove correct that the formation of bright lines in *B* stars is influenced, and facilitated, by a rapid rotation of the stars.

As an alternative Gerasimovic suggested that Struve's rotation

† *Monthly Not.* **94** (1934), 737.
‡ Merrill, *Pub. Astron. Soc. Pacific*, **45** (1933), 49.

hypothesis should be restricted to *Be* stars with nebulous lines, and he suggested that in *Be* stars with sharp emission lines the most important feature is an expansion of the envelope, the gases being repelled from the star by radiation pressure. Gerasimovic could show, by a revision of earlier calculations by Johnson,† that in the temperature range of the *Be* stars radiation pressure on hydrogen might well counterbalance gravity.

Although the calculations of Gerasimovic on the expansion of a *Be* envelope as a consequence of light pressure did not lead to results which could be immediately interpreted in terms of observations, it seems to the present author that this mode of approach is likely to become more fruitful in the future. The further step of considering the *Be* emission as a magnified form of the emission from the solar chromosphere and the corona seems natural. In both cases there is ground for believing that rotation (through the generation of a magnetic field), radiation pressure, and other repulsive forces may come into play, and an attack on the problem from the point of view of the identity in nature of the two phenomena would seem to be a fruitful field of work.

107. Expanding Envelopes. *WR* Stars and Novae

Passing on to *WR* stars and novae conditions become clearer, as there is scarcely any doubt that the profiles of the emission lines in this case must be interpreted in terms of a preponderant outward motion of an extensive gaseous cloud enveloping the star. It may still be true that a rotation of the star plays a part in bringing about and maintaining the outflow. But there is little chance of proving such a proposition from the spectrum, where the widths of the emission lines correspond to Doppler velocities from 500 km./sec. for *WR* stars and up to 3,000 km./sec. for novae. That rotational velocities of this order should be found in stars is wellnigh impossible.

The first author who seriously suggested that the width of emission lines in the spectra of *WR* stars and novae reveals the existence of envelopes expanding with a prodigious velocity was Beals,‡ who thought that the envelope was expanding under the influence of radiation pressure. However, a simple uniform expansion cannot

† *Monthly Not.* **85** (1925), 813, and **86** (1926), 300.
‡ *Pop. Astron.* **37** (1929), 577; *Monthly Not.* **90** (1930), 202.

explain the many irregular profiles observed in bands of novae, and Genard† suggested that the observations rather indicate a series of violent eruptions from the surface. These two points of view are not mutually exclusive, the latter appearing as a kind of refined edition of the former, as far as kinematics is concerned.

In developing a formal theory of expanding envelopes it is natural to begin with the point of view of Beals, and consider a spherically symmetrical envelope, in which the velocity is directed radially and is a function of the radius vector r to the centre only. Preliminary work in this field has been done by Gerasimovic‡ and Chandrasekhar.|| The first attempt at such a theory was, however, given much earlier by Carroll.††

Consider then a star surrounded by a spherically symmetrical expanding envelope. The state is taken to be stationary, which answers well to conditions in *WR* stars, but is a poorer approximation in the case of novae.

Consider a volume element dv in the form of a ring round the line joining the observer and the centre of the star, the direction to the centre forming an angle θ with the line of sight. This volume element is then

$$dv = 2\pi r^2 dr \sin\theta \, d\theta.$$

The radiation of rest-frequency ν_0 emitted from this element in the direction of the observer will have the apparent frequency

$$\nu = \nu_0(1+u/c),$$

where u is the radial velocity in the line of sight, positive in the direction towards the observer, so that

$$u = V\cos\theta.$$

Let us, in the expression of the volume element, perform a transformation of coordinates from r, θ to r, ν. The Jacobian of this transformation is

$$\frac{\partial\theta}{\partial\nu} = -\frac{c}{\nu_0 V \sin\theta},$$

and the expression of the volume element becomes

$$dv = \frac{2\pi c}{\nu_0}\frac{r^2 dr}{V} d\nu.$$

Let ρ' denote the emissivity of matter for radiation of natural frequency ν_0. The emission from the given volume element is then

† *Monthly Not.* **92** (1932), 396.
|| Ibid. **94** (1934), 522.
‡ Ibid. **94** (1934), 737.
†† Ibid. **88** (1928), 548.

$\rho' dv$ and the total emission in the frequency interval dv is obtained by integrating this expression over those parts of the envelope which contribute to this emission in the direction of the observer. The question of the limits of integration requires some care. We therefore leave these limits arbitrary for the moment. Neglecting reabsorption we may then write for the observed intensity per unit frequency range:

$$I_\nu = \frac{2\pi c}{\nu_0} \int \rho' \frac{r^2\, dr}{V}.$$

The problem in hand is thus reduced to the evaluation of this integral for different assumptions with regard to V and ρ'.

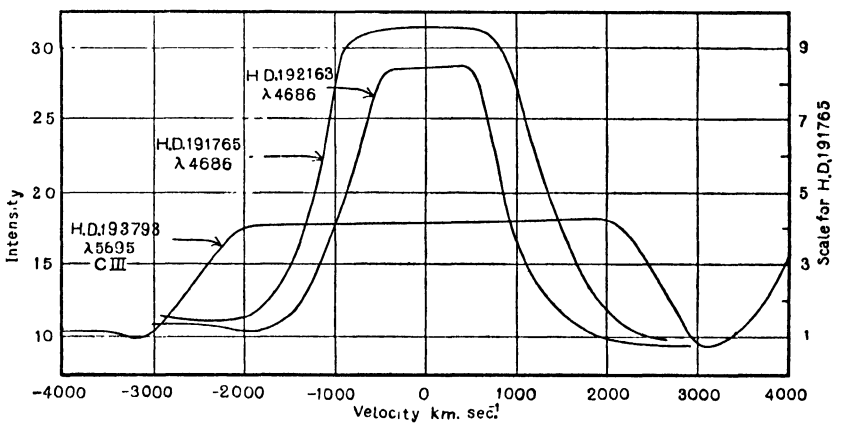


FIG. 33. Typical flat-topped profiles of emission lines in *WR* stars (Beals, *Monthly Not.* **91**).

108. Envelope expanding with Constant V.

In so far as the dimensions of the star are negligible in comparison with the dimensions of the luminous envelope, the integral is just a constant, and the intensity per unit frequency range is also a constant within the limits

$$\nu = \nu_0(1 \pm V/c).$$

This holds independently of the functional form of ρ'. The emission line is then drawn out into a band of width $\Delta\nu = 2\nu_0 V/c$, and with uniform intensity inside these limits. The profile is of the so-called flat-topped form. A few pure cases of this sort are known, but flat-topped profiles are exceptions and not the rule.

When the star is not negligible in dimensions compared with the envelope, the red wing of the line is rounded off by the so-called

'occultation effect', which simply means that the star screens off the light emitted by the part of the envelope which is hidden behind the star. Since this part of the envelope is moving away from the observer, the occultation effect will only affect the red wing of the line. Instead of the stellar radius a, the lower limit of integration for a direction θ is $a/\sin\theta$. Let R denote the outer radius of the shell, and assume the emissivity to be a constant. Then

$$I_\nu = \text{const.}(R^3 - a^3/\sin^3\theta),$$

$\sin\theta$ being a function of ν by Doppler's principle.

A better approximation may be had by assuming the emissivity proportional to mass density, instead of assuming it constant. Denote mass density by ρ as usual. The equation of continuity for radial motion is

$$\frac{\partial\rho}{\partial t} + \frac{1}{r^2}\frac{\partial}{\partial r}(\rho r^2 V) = 0,$$

which for a stationary state may be written

$$\rho r^2 V = K = \text{const.}$$

Taking $\rho' = \rho$ leads to the intensity formula for the red wing

$$I_\nu = \text{const.}(R - a/\sin\theta),$$

corresponding to a much more gradual rise in intensity than in the former case. This is easily understood, since the emission is now more concentrated towards the surface of the star. There is also a displacement of the centre of gravity of the line towards the violet, since the intensity drops to zero at an angle θ_0 given by $\sin\theta_0 = a/R$.

If the luminous part of the envelope forms a detached shell of inner radius R_1, the intensity remains constant also in the red wing of the line until an angle θ_1 is reached, for which

$$\sin\theta_1 = a/R_1.$$

From then on the intensity drops off as before. If the shell is sufficiently thick optically, there may occur reabsorption effects. As will be shown later (§ 123), these must preferentially be sought near the centre of the line, where the line of sight passes the luminous shells tangentially, and the optical paths in the shell are a maximum. If the *Be* envelopes are also assumed to belong to the expanding type, the central depression between the two emission components of double emission lines in *Be* spectra must also be interpreted as a reabsorption effect.

It is also worth noticing that the shell itself will produce absorption

lines in the stellar spectrum. The sharp absorption lines on the violet edges of emission bands in novae and *P* Cygni stars are typical instances. More doubtful, but still suggestive in this connexion, is the occurrence of very narrow absorption lines in *Bne* spectra.†

This rough model of an expanding atmosphere only gives a first approximation to the truth, as flat-topped profiles are scarce.

This, of course, was to be expected. For the moving gases are subjected to the influence of the gravitation of the star, which will retard their motion, and make the velocity smallest at the greatest heights. Or, conversely, it may be that the gases are accelerated away from the star, by radiation pressure, for example, or electrical forces, in which case the velocity will be the larger, the greater the height. As a suitable illustration of these two opposite cases we shall consider the theory of the line profile when the atmosphere is accelerated outward by an inverse square force, or retarded in its outward motion by such a force.

109. Accelerated Expansion

Assume the matter to leave the stellar surface with zero velocity, and let V_0 denote the limiting velocity at infinity. The equation of conservation of energy then gives the following expression for the velocity:

$$V = V_0\sqrt{(1-a/r)}.$$

For the sake of convenience we may assume the emissivity ρ' to be proportional to $V^n r^{-4}$,

$$\rho' = AV^n r^{-4},$$

where A and n are constants. This particular choice has the advantage of analytic simplicity, as it permits the integration to be carried out.

From the given expression of V we find, in fact,

$$\frac{r^2\,dr}{V} = r^4\frac{2\,dV}{aV_0^2}.$$

Changing the variable from r to V, $\rho'\,dv$ becomes

$$\rho'\,dv = \frac{4\pi cA}{\nu_0 aV_0^2}V^n\,dV,$$

from which the integral follows directly.

The above choice of the emissivity function is also sufficiently flexible to represent a detached shell of luminosity. By differentiation

† Cf. Struve and Swings, *Astrophys. J.* **75** (1932), 176 et seq.

it follows that the emissivity has a maximum at

$$r = a\left(1+\frac{n}{8}\right).$$

This maximum will thus lie outside the star, and have a real importance, provided n is positive. The larger n is, the more pronounced is the concentration of the luminosity into a shell. The given expression is thus also able to give a first approximation to an expanding planetary ring nebula. Corresponding results might, of course, also be obtained by different analytical expressions of ρ'. It is only

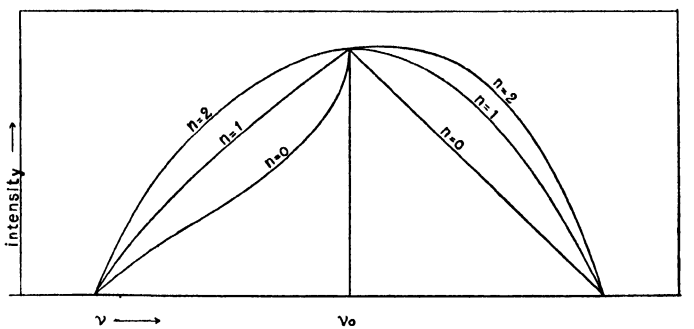


FIG. 34. Theoretical emission-line profiles of an expanding atmosphere. The motion is supposed accelerated outward according to the law discussed in the text, and the emissivity is assumed to be proportional to $V^n r^{-4}$ (Chandrasekhar, *Monthly Not.* 94). Note the rounded form of the profiles in contrast to the flat-topped profiles, which correspond to an expansion with constant V.

necessary that $\rho' r^4$ is given as an integrable function of V. If it be desired, for instance, to study a shell with a more sharply defined outer boundary, it is possible to assume the emissivity to have the form

$$\rho' = \text{const.}\, r^{-4} V^n e^{-kV},$$

where k is a new constant. In the following, however, we shall be concerned with the simple power law only, which was first studied by Chandrasekhar, although in an analytically different form. The assumed dependence of ρ' on the velocity need not give rise to any objections. In fact in these calculations V stands only for a certain function of r, and the separating out in the expression for ρ' of terms in V and terms in r is made only from the point of view of convenience. We assume n to be larger than -1, so as to make the integral converge at the stellar surface. The intensity is then

$$I_\nu = \frac{4\pi c A V^{n+1}}{\nu_0 a V_0^2 (n+1)} + \text{const.}$$

The limits of integration will turn out differently in the two halves of the line, owing to the occultation effect. In the violet wing there is no question of a screening by the star, and the upper limit of V is the limiting velocity V_0, and the lower limit is the radial velocity u, which corresponds to the given frequency interval

$$u = c\frac{\nu-\nu_0}{\nu_0}.$$

Hence

$$I_\nu = \frac{4\pi cA}{\nu_0 a V_0^2(n+1)}\{V_0^{n+1} - u^{n+1}\}.$$

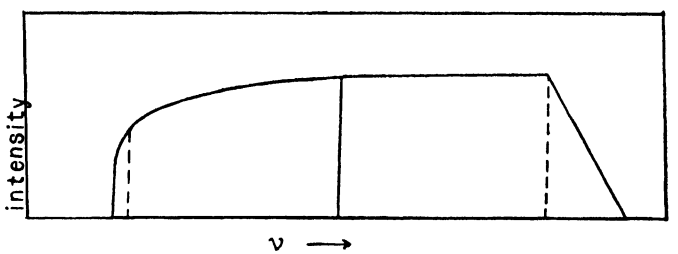


FIG. 35. Profile of an emission line originating in an expanding atmosphere with a decelerated motion. The matter is supposed to be ejected with a constant velocity, and the motion to be slowed down by gravity. The emissivity has the same functional form as in Fig. 34. The profile is now seen to approach much more to the ideal flat-topped form. This is due to the assumed functional form of the emissivity, essentially, which confines the velocity within much more narrow limits, provided the initial velocity of ejection is sufficiently large, as is assumed in the computed example (Chandrasekhar, *Monthly Not.* **94**, 522).

In the red wing of the line the upper limit is still V_0, but the lower limit is higher than u, because $\sin\theta$ has the lower limit a/r, corresponding to the boundary of the 'shadow cylinder' of the star. Denoting this r by r_s we have

$$\cos\theta = \sqrt{(1-a^2/r_s^2)} = \frac{u}{V_0\sqrt{(1-a/r_s)}},$$

or

$$(1-a/r_s)\sqrt{(1+a/r_s)} = u/V_0.$$

This equation must be solved for a/r_s and the proper root introduced into the expression of V, which then gives the lower limit of the integral.

The conditions are illustrated in Fig. 34. The curves illustrate how the profile becomes more and more rounded with diminishing effective thickness of the shell, i.e. for increasing values of n.

110. Decelerated Expansion

The case of the ejection of particles from the stellar surface with an initial velocity V_i, say, the motion being subsequently subjected to

a retardation according to the inverse square law, may be treated in nearly the same way, by a little rearrangement of the analytical machinery. The energy equation becomes in this case

$$V^2 = V_i^2\frac{a}{r} + V_0^2\left(1-\frac{a}{r}\right),$$

V_0 being as before the velocity at infinity. The transformation from r to V gives in this case

$$\frac{r^2\,dr}{V} = \frac{2r^4\,dV}{a(V_0^2-V_i^2)},$$

and the intensity becomes

$$I_\nu = \frac{4\pi c A V^{n+1}}{\nu_0 a(V_0^2-V_i^2)(n+1)} + \text{const.}$$

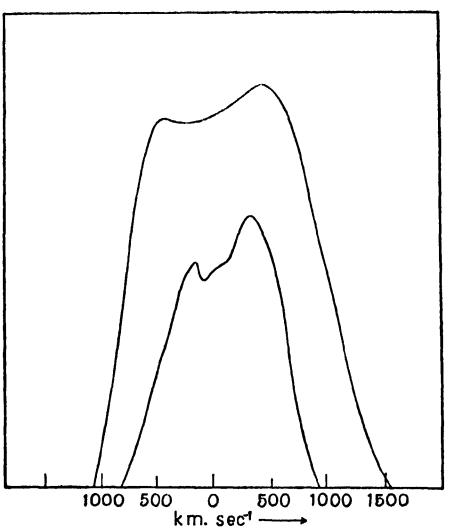


FIG. 36. Profiles of H_β in Nova Cyg 1920, August 27 (upper curve), and in Nova Lac 1911, January 29 (Genard).

The question of the integration limits is this time a little more complicated, as it is necessary to distinguish between the cases $u < V_0$, and $V_0 < u < V_i$. In the violet wing the upper limit is V_i, as it is always possible to find a $\cos\theta$ small enough for $\cos\theta = u/V_i$ to hold. If $u < V_0$, the lower limit is V_0 itself, corresponding to a limiting $\cos\theta = u/V_0$. The intensity is therefore constant in this range.

In the range $V_0 < u < V_i$

the lower limit is $V = u$, giving

$$I_\nu = \text{const.}(V_i^{n+1} - u^{n+1}).$$

In the red wing a similar distinction between $u \gtrless V_0$ must be made. When

$$u < V_0$$

the lower limit of integration is V_0. The upper limit is determined in the same way as for accelerated motion, by first solving the equation

$$\cos\theta = \sqrt{(1-a^2/r_s^2)} = \frac{u}{V} = \frac{u}{\sqrt{\left\{V_i^2\frac{a}{r_s} + V_0^2\left(1-\frac{a}{r_s}\right)\right\}}}$$

for a/r_s, and introducing this root into the expression for V, which then gives the proper lower limit.

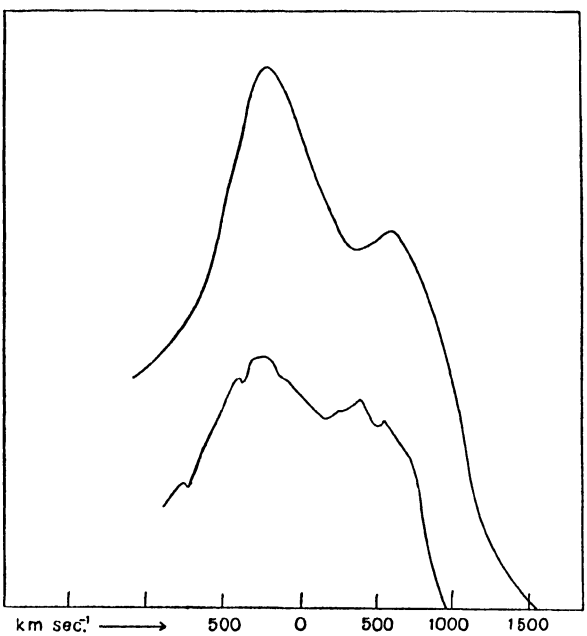


FIG. 37. Profiles of H_γ in Nova Cyg 1920, August 27 (upper curve), and in Nova Lac 1911, January 29 (Genard, *Monthly Not.* **92**).

When $u < V_0$ this equation has but one positive real root. When $V_0 < u < V_i$ the equation has two positive real roots, which give both integration limits. For increasing values of u these roots will come closer and closer together till the intensity finally vanishes when the roots coincide. This double root is therefore obtained by calculating the maximum value of u. By a simple calculation one finds the critical value of r to be given by

$$r = a\{\alpha+\sqrt{(\alpha^2+3)}\},$$

where $$\alpha = \frac{V_0^2}{V_i^2 - V_0^2}.$$

A typical profile of the sort discussed above is shown in Fig. 35.

The above discussion of possible profiles of emission lines from expanding envelopes brings out the fact that there is a large variety of possibilities to be considered. In particular, it appears premature

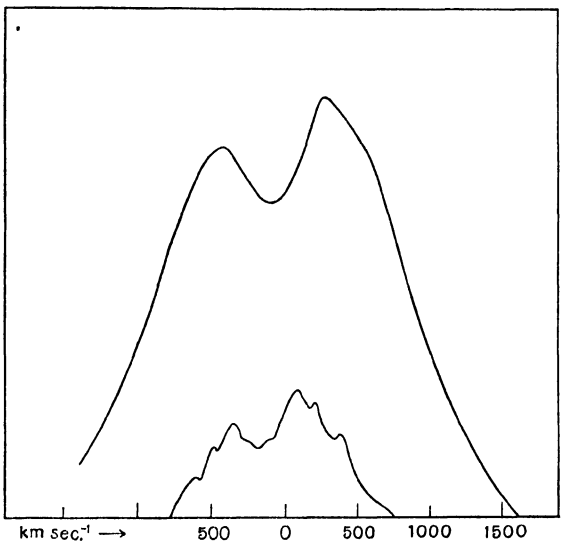


Fig. 38. Profiles of the nebular line 4,658 A. in Nova Cyg 1920, August 27 (upper curve), and in Nova Lac 1911, January 29 (Genard, *Monthly Not.* **92**).

for Genard† to consider it necessary to assume series of irregular eruptions in order to account for the profiles of emission bands in novae, as the general features of his measured profiles fall well in line with what is to be expected in spherically symmetrical envelopes, although modifications may be required in detail.

† *Monthly Not.* **92** (1932), 396.

XXI

FORBIDDEN TRANSITIONS

111. Transitions from Metastable States

ALTHOUGH there may be divergence of opinion about details, it must be admitted that the study of the emission bands in the spectra of *WR* stars and novae has shown conclusively that the outer atmospheres of these stars are expanding freely into interstellar space. The velocities of outflow are so large that it is clear that the ejected gases will not return to the star again, but form a nebula of indefinite extent around the star, of the same apparent constitution as the gaseous nebulae long observed in the sky.

There is reason to believe that these nebulae are generated by the activity of *O* stars, *WR* stars, and novae. The evidence for this is given by the fact that the stars embedded in diffuse nebulae are all of the *O* type, or occasionally the *B* type, while the nuclear stars of planetary nebulae are mostly *WR* stars. All planetaries are perhaps post-novae. Thus the Crab nebula in Taurus is suspected of being identical with a nova recorded by Chinese astronomers 900 years ago. The organic connexion between stars and nebulae appears on the whole to be a settled fact.

The transition of matter from atmospheric density at the surface of a star to the ultra-vacuum of interstellar space, where the illuminating star is situated at a great distance, entails considerable changes in the physical properties of the gas. The gradual development of these changes may be followed step by step for the novae, and the final state is permanently recorded in the spectra of planetary and diffuse nebulae.

The most significant feature about this transformation is the appearance of lines corresponding to 'forbidden transitions' in the spectra of novae and nebulae, which, during the last ten years has made the study of transitions from metastable states an important part of astrophysics.

The theory of transition probabilities is usually based on the calculations of harmonic amplitudes, in the way shown earlier in this book. It is by this procedure that the ordinary selection rules of spectroscopy have been verified theoretically. On passing to higher approximations, however, it is usually found that transitions forbidden by

the selection rules will take place with a finite probability, although at a rate of the order a million times less than for ordinary transitions.

In most cases forbidden transitions are of no importance, but a certain class of such transitions, those taking place from a metastable state, may under certain conditions play a considerable role. By a metastable state we understand a quantum state from which all spontaneous transitions to lower energy states are forbidden by the ordinary selection rules. A state of this sort is the state $2s$ in hydrogen. This is a level from which all spontaneous transitions lead to the normal state. But since this latter is a $1s$ state, transitions from $2s$ are forbidden by the selection principle for the azimuthal quantum number.

The reason why transitions from metastable states may acquire a special importance for nebulae lies close at hand. Considering the Orion nebula as a convenient example, it is clear that its emission of radiation is stimulated by the high-temperature O stars which are embedded in the nebula. The radiation from the embedded stars will be absorbed by the nebular material, and mostly go to ionize the atoms, the nebular radiation being emitted for the most part during the recombination process. In the course of this process a number of atoms will be caught in metastable states. And they will remain there until something happens to them.

Under ordinary experimental conditions the first thing happening to such an atom would be a collision, either with another gas atom or with the surface of adjoining solid bodies. In the Orion nebula, however, there are no solid bodies to collide with, and the density of the gas is so small that serious collisions will only take place a few times per year.

One may imagine, however, that the atom will be removed from the metastable state by absorption of energy, in consequence of which it will be lifted to a higher state, from which it may suffer transitions to the normal state. But the radiation density in the nebula is so small that the probability of such absorption acts is unimportant in comparison with the residual probability of spontaneous transitions in violation of the selection principle. Conversely, when anywhere in the stellar or nebular spectra forbidden emission lines from metastable states are discovered, we know that they originate in tenuous matter situated at some distance from the exciting light source.

When the transition probabilities have been calculated, or found by experiments of some sort, it is possible to say something more definite of the physical conditions underlying the emission process.

112. Metastable States of Astrophysical Significance

It is not possible here to enter into great detail on the theory of forbidden transitions, because it is very complicated. We must therefore be content with indicating the principal points and a discussion of the results.

The forbidden transitions of astrophysical interest belong for the most part to three types, which are illustrated by the diagram given in Fig. 39.

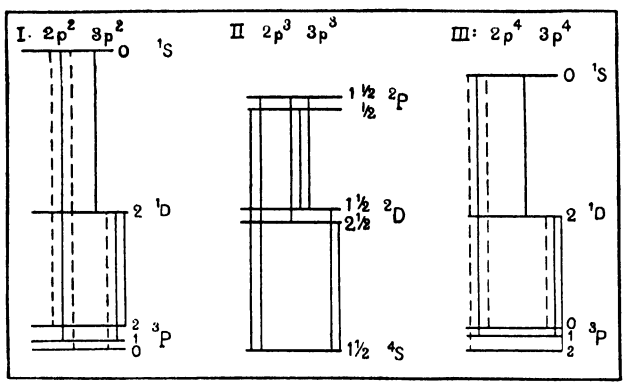


FIG. 39. Level diagram of low metastable states. Boyce, Menzel, and Payne (*Proc. Nat. Acad. Washington*, **19**, 581) suggested a special terminology for transitions between such states. Transitions from middle to low were called 'nebular' transitions, from high to middle 'auroral' transitions, and from high to low 'trans-auroral' transitions. For the atomic configurations which correspond to the above level schemes compare Table 10.

It will be noticed that transitions to the lowest level give rise to inter-system combination lines, while transitions to the middle level are forbidden by the selection rule for the azimuthal quantum number. This difference means that the theory of these transitions will also be different.

The list on p. 304 includes those atomic configurations which have the above level systems as the lowest levels, and which prove to have astrophysical importance.

The quantum theory of forbidden transitions has been considered by a series of investigators, but still the theory appears far from complete. As mentioned above there are several ways in which a

TABLE 10

I: $2p^2$	$3p^2$	II: $2p^3$	$3p^3$	III: $2p^4$	$3p^4$
C I	Si I	N I	P I	O I	S I
N II	P II	O II	S II		
O III	S III	F III	Cl III	F II	Cl II
F IV	C IV	Ne IV	A IV		

transition may be forbidden in the first approximation and will reappear in the next. The occurrence of inter-system combination lines means that there exists a departure from the ideal coupling system assumed in the classification of the levels. We have here assumed the multiplets to be normal in the sense of the Russell-Saunders coupling system. It may be recalled that this system rests upon the assumption that the interaction between orbital magnetic moments and the electronic spins separately is sufficiently strong to make them combine into quantized resultants L and S respectively. These two vectors enter into interaction, and combine into a resultant J, but this latter interaction is supposed small in comparison with the one producing the quantized vectors L and S.

The presence of inter-system combination lines means that there is a departure from this coupling scheme, a lessening of some sort of the bonds between the individual spins or orbital momenta, and a corresponding strengthening of the bond between L and S. As was pointed out by Condon† this may give rise to transitions enforced by a spontaneous radiation which may be said to be due to the oscillation of a magnetic dipole in the atom. The magnetic moment of the atom is proportional to $L+2S$, and if there is a sufficient lessening of the bonds, this moment will no longer be entirely quantized, but, for a combination of two states of different multiplicity, correspond to an oscillation, and a consequent transition.

Although we are not going to develop the details of the theory of inter-system combinations, it may be of interest to recall some properties of magnetic dipole radiation. The classical theory of such a dipole runs exactly parallel to the theory of an electric dipole. Denote the moment of the former by M and of the latter by P, and further the electric and magnetic field intensities by E and H. To pass from an electric to a magnetic dipole means to replace P, E, and H by M, H, and $-E$ in all formulae. Carrying this formalism directly over

† *Astrophys. J.* **79** (1934), 217.

into the quantum theory, the Einstein probability of spontaneous transitions from a state a to a state b is given by the usual formulae ((25) and (53)) provided the matrix element of the electric moment is replaced by the corresponding matrix element of the magnetic moment:

$$A_{ab} = \frac{64\pi^4\nu^3}{3hc^3}(a|M|b)^2,$$

where $(a|M|b)$ is the matrix element of M for the two states a and b. The analogy between electric and magnetic dipole is so close that much of the ordinary radiation theory is applicable also to the latter. Thus the selection rule $\Delta J = 0 \pm 1$ (except $J = 0 \to J = 0$) still holds. The magnetic moment is further given by

$$M = \frac{eh}{8\pi\mu c}(L+2S).$$

The problem is thus reduced to the calculation of the matrix components of $L+2S$, which, however, is no really simple matter. It was carried out by Condon in the paper referred to above for some interesting cases, which will be considered further below.

The transitions $^1S \to {}^1D$ are of a different type. The terms have the same multiplicity, so there is no question of inter-system combinations. In return there is a violation of the selection rule for the azimuthal quantum number, as the jump in L is by two units. This time the magnetic dipole radiation is of no particular consequence, and the transition will be enforced by what is called the electric quadripole radiation of the atom.

Consider the radiation from an electric system according to the classical theory and find the expression for the potentials at a great distance from the atom. Expanding the potential according to powers of the ratio of the atomic dimensions relative to the distance r from the atomic centre, the first term in the expansion will be given by the net charge divided by r. The second term will contain the dipole moment P multiplied by a function of the angle θ between P and r, divided by r^2. The third term will by definition give the quadripole moment of the electric system, again multiplied by a function of θ, and divided by r^3. If the first two terms are zero, or do not vary with the time, the quadripole term will come into play and determine the radiation from the system. Our usual selection rules are determined by the vanishing of the electric dipole moment, and may prove invalid when the quadripole term is also taken into consideration.

But it is usually found that the probabilities of transitions corresponding to quadripole radiation are millions of times smaller than for dipole radiations.

For transitions $^1S \rightarrow {}^1D$ it is the quadripole radiation which produces the main effect. Quadripole radiation will also contribute to the intensity of inter-system combination lines, but is small in comparison with the effect of the magnetic dipole radiation, and may be neglected in the first round. Stevenson† developed a theory of inter-system combinations in which both effects were treated indiscriminately, and it was Condon who first pointed out their relative importance. Otherwise his paper gave the same results as Stevenson's.

The quantum theory of quadripole radiation was first developed by Rubinowicz‡ and his collaborators. In particular he has shown that in quadripole radiation the selection rule for the azimuthal quantum number may be violated, and Bartlett|| has shown that the same is the case for the inner quantum number J, except for the transition $J = 0 \rightarrow J = 0$. Bartlett also made extensive calculations on the intensity of inter-system lines, although his work was not so complete as that of Stevenson and Condon. The selection rules for J which are expected to hold for transitions in J are, according to Rubinowicz,††

$$\Delta J = 0, \pm 1, \pm 2; \qquad \text{except } 0 \rightarrow 0, \tfrac{1}{2} \rightarrow \tfrac{1}{2}, \text{ and } 1 \not\rightleftarrows 0.$$

We next give the numerical data of interest to astrophysics as given in Condon's paper.

TABLE 11

Transition Probabilities from Metastable States (in sec.$^{-1}$)

Case I, the $2p^2$ configuration. d = magnetic dipole lines, q = quadripole lines

Transitions	C I	N II	O III
$^1D_2 \rightarrow {}^3P_2(d)$	$1{\cdot}5.10^{-4}$	24.10^{-4}	18.10^{-3}
$^1D_2 \rightarrow {}^3P_1(d)$	$0{\cdot}5.10^{-4}$	$8{\cdot}1.10^{-4}$	6.10^{-3}
$^1S_0 \rightarrow {}^3P_1(d)$	41.10^{-4}	66.10^{-3}	$10{\cdot}2.10^{-2}$
$^1S_0 \rightarrow {}^3P_2(q)$	$0{\cdot}6.10^{-5}$	$0{\cdot}5.10^{-4}$	$1{\cdot}5.10^{-4}$
$^1S_0 \rightarrow {}^1D_2(q)$	1·0	1·7	1·8

It will be noticed that in the transition $^1S_0 \rightarrow {}^3P_2$ the quadripole radiation is the most important, although it is really an inter-system line.

† *Proc. Roy. Soc.* A, **137** (1932), 298.

‡ See Rubinowicz and Blaton, *Ergeb. d. exacten Naturwiss.* **11** (1932), 170, for a full review of his theory.

|| *Phys. Rev.* **34** (1929), 1247.

†† *Zeits. f. Physik*, **61** (1930), 338.

In the next table we give the data for O II and O I, corresponding to the configurations II and III of Fig. 39, since only in these cases are there sufficiently accurate data at hand.

TABLE 12

Transition Probabilities from Metastable States in sec.$^{-1}$

Case II: $2p^3$	O II	*Case III*: $2p^4$	O I
$^2D_{5/2} \rightarrow {}^4S_{3/2}$	$3{\cdot}7.10^{-5}$	$^1D_2 \rightarrow {}^3P_2$	$7{\cdot}5.10^{-3}$
$^2D_{3/2} \rightarrow {}^4S_{3/2}$	$2{\cdot}42.10^{-5}$	$^1D_2 \rightarrow {}^3P_1$	$2{\cdot}5.10^{-3}$
$^2P_{3/2} \rightarrow {}^4S_{3/2}$	$4{\cdot}0.10^{-2}$	$^1S_0 \rightarrow {}^3P_2$	$1{\cdot}6.10^{-4}$
$^2P_{1/2} \rightarrow {}^4S_{3/2}$	$4{\cdot}9.10^{-2}$	$^1S_0 \rightarrow {}^3P_1$	$1{\cdot}8.10^{-5}$
$^2P_{3/2} \rightarrow {}^2D_{5/2}$	2·3	$^1S_0 \rightarrow {}^1D_2$	$2{\cdot}0.10^{-4}$
$^2P_{3/2} \rightarrow {}^2D_{3/2}$	1·7		
$^2P_{1/2} \rightarrow {}^2D_{5/2}$	1·0		
$^2P_{1/2} \rightarrow {}^2D_{3/2}$	3·3		

113. Origin of the Luminosity in Gaseous Nebulae

The theory of forbidden transitions first became of importance in the interpretation of spectra of gaseous nebulae, where the appearance of several lines of unknown origin was long considered an outstanding problem. The correct interpretation of these lines in terms of forbidden transitions from metastable states, notably in ionized oxygen and nitrogen atoms gave the key to the understanding of the physical constitution of the nebulae, and made it possible to make the first steps towards a quantitative theory of such nebulae.

It is here only a question of nebulae belonging to our galactic system, and among these only such as show an emission-line spectrum. The very extended and irregular nebulae, of the type of the Orion nebula or the Trifid nebula, form one extreme of the emission nebulae, and the small, very regular, round or annular planetary nebulae form the other extreme. All known emission nebulae are relatively near. The Orion nebula has a distance of the order 500 light years. The Trifid nebula in Sagittarius may have a distance of 3,000 or 4,000 light years, and most known irregular emission nebulae will be found between these extremes. As far as the parallaxes of planetary nebulae are known, it appears that the nearest are a few hundred light years off, while the farthest we know have distances of some thousand light years.

The emission of light from the nebulae is undoubtedly in all cases due to embedded stars or stars standing so close that their radiation is sufficient to stimulate the nebular emission. In the case of planetaries

the situation is fairly clear, since these nebulae have a central star, which obviously is the cause of the emission. As regards the irregular nebulae the situation is sometimes a little doubtful, but in most cases the active star is easily identified.

Hubble, whose work has contributed very much to the proper understanding of nebular radiation, has found that there is a definite relation between the spectrum of the active star and the character of the nebular radiation, which runs as follows:

Planetary nebulae are excited by stars of earlier class than *Oe* 5.

Emission nebulae are excited by stars of class *Oe* 5 to *B* 0.

If nebular stuff is illuminated by stars later than *B* 1 the nebulae will show up simply in reflected light—the spectrum of the nebula coinciding with the spectrum of the star. An example is furnished by the Pleiades (spectral class *B* 4–*K* 0). We call such nebulae reflection nebulae.

It appears that in all cases when an investigation may be profitably undertaken the nebular light intensity falls off with the inverse square of the distance from the exciting star, as would be expected when it is a question of a simple reflection of the light. This holds for reflection nebulae as well as for ordinary emission nebulae. The relation was first tested by Herzsprung† for the Pleiades. Later it was taken up on a broad basis by Hubble,‡ who extended the investigation to both reflection and emission nebulae.

When the stars are responsible for the nebular radiation it is to be expected that intrinsically very luminous nebulae must be associated with stars of high absolute luminosity. As shown by Hubble, this relation leads to the consequence that the apparent magnitude of the illuminating star must stand in a definite relation to the surface brightness of the surrounding nebula. Hubble gave this relation in the form that the apparent diameter A of the nebula (on the photographic plate) must be proportional to the apparent luminosity of the exciting star, or expressed logarithmically:‖

$$m + 5 \log A = 11{\cdot}09,$$

where m is the apparent magnitude of the star and α the apparent diameter in seconds of arc. The diagram in Fig. 40 shows the agreement between theory and observation for irregular emission nebulae.

† *Astron. Nach.* **195** (1913), 449. ‡ *Astrophys. J.* **56** (1922), 400.

‖ The numerical constant 11·09 was first given correctly by Zanstra, ibid. **68** (1927), 50.

It appears that the linear relation between m and $5 \log A$ corresponds to actual facts, and that the numerical values of the constants are also reproduced. That the slope of the line comes out right proves the correctness of the inverse square law in this case. That the height of the line also is correct shows that there is no light lost in the nebula. This would be easily intelligible for a pure reflection nebula, but is not so easily understood for emission nebulae, where the radiation

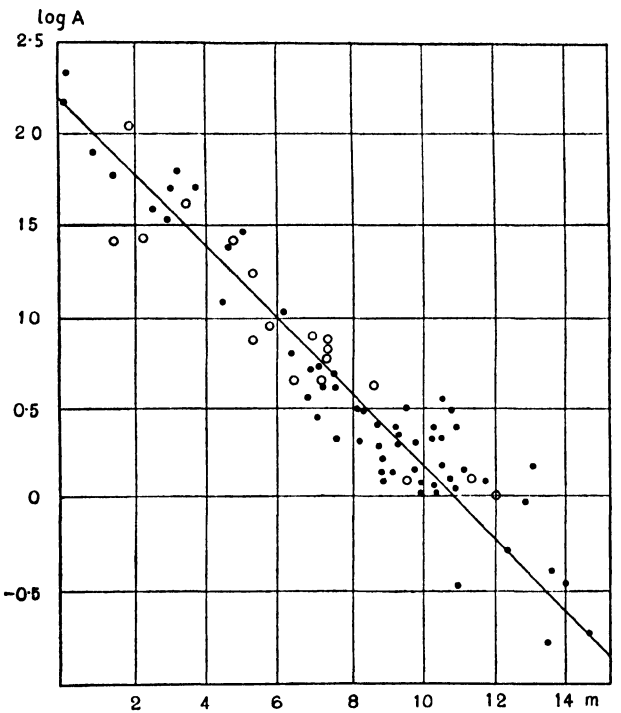


FIG. 40. Correlation between apparent magnitude of the embedded star and the extension of the surrounding nebula. The straight line gives the theoretical correlation.

received from the star is obviously transformed in many ways before it is re-emitted. The agreement is too good, and calls for a special explanation. Anyhow, the agreement in question is sufficient to show beyond doubt that nebulae have no intrinsic energy sources and are merely to be considered as transformers of stellar energy, in the same way as the earth, the planets, and the moons in our own solar system.

That part at least of this latter agreement is accidental is indicated, however, by the fact that the relation is very far from being true for

planetary nebulae, where the nebula is 7–8 magnitudes brighter than the exciting central star.

The true explanation of the discrepancy between the magnitudes of the central star and the surrounding nebula for planetaries should probably be looked for in the very high temperature of the central stars. It is difficult to obtain a very direct determination of these temperatures, but Wright has shown that the maximum intensity in the radiation for these stars is to be found far beyond the limit of atmospheric absorption.

The spectral type of these stars is pronouncedly of the early *O* or the Wolf-Rayet type, which further indicates very high temperature. It is a reasonable procedure to admit that the light radiation from the central star is the sole cause of nebular light, and to use the discrepancy in luminosity between star and nebula for a determination of the stellar temperature, a procedure first suggested by Zanstra.†

114. Spectra of Gaseous Nebulae

The spectra of the emission nebulae consist of sharp emission lines, possibly on a very faint continuous background. The Balmer lines of hydrogen are prominent in all nebular spectra, the first fourteen lines from H_α to H_ξ being easily distinguishable. Beyond the series limit 3,647·0 A. there is a strong continuous emission, especially in the spectra of many planetaries, which is clearly due to the recombination of free nuclei and free electrons to form neutral hydrogen atoms in the second quantum state. The hydrogen lines are sharp, and show no unusual peculiarities.

Helium is represented by both arc lines and spark lines. The relative intensity of these lines varies much from one nebula to another, as a consequence of different conditions of excitation. It was thus discovered by Max Wolf‡ that in slitless spectra of planetary nebulae the images were considerably larger in light of He I lines than in He II lines. This interesting feature provides a further proof that the nebular radiation is stimulated from the central star.

The lines of hydrogen and helium appearing in nebulae are quite ordinary lines, which by themselves would not attract much attention, although the study of these lines already gives many indications about the internal constitution of the nebulae. The rest of the

† *Astrophys. J.* **65** (1927), 50; *Publ. Dom. Astr. Obs. Victoria*, **4** (1931), 209.
‡ *Sitz. d. Akad. d. Wiss. Heidelberg*, **27** (1911).

nebular lines, however, are sensational, as they are a mixture of ordinary and forbidden lines, belonging for the most part to singly and doubly ionized oxygen and nitrogen, and to a minor degree to other atoms. This discovery was made by Bowen† in 1927 on the basis of experiments made by Bowen and Millikan, A. Fowler, Paschen, and other experimenters working with the spectra of strongly ionized atoms.

As most of the interest here is concentrated on the forbidden lines, we give a tabular survey of all lines of this character observed in nebulae, along with their series relationships and chemical origin.

TABLE 13

Forbidden Lines in Nebular Spectra

Element	*Multiplet*	*Wave-length*	*Intensity* N. G. C. 7,027	*Intensity* N. G. C. 6,572
N II	$D-S$	5754·8	4	2
	$P-S$	6348·1	10	10
	$P-D$	6583·6	30	30
O I	$D-S$	6302	3	5
	$P-D$	6364	1	2
O II	$S-D$	3726·16	15	8
	$S-D$	3728·91	7	4
	$D-P$	7319·9	3	3
	$D-P$	7330·4	2	2
O III	$D-S$	4363·21	20	80
	$P-D$	4958·91	80	200
	$P-D$	5006·84	250	800
Ne III	$D-S$	3342	..	..
	$P-D$	3868·74	50	70
	$P-D$	3967·51	50	70
Ne IV	$D-P$	4725·5	..	4
Ne V	$P-D$	3345·8	..	8
	$P-D$	3425·8	..	20
S II	$S-P$	4068·62	..	6
	$S-P$	4076·22	..	3
	$S-D$	6716	..	..
	$S-D$	6730	..	1
A III	$P-D$	7135·6	5	1
A IV	$S-D$	4711·4	..	..
	$S-D$	4740·2	20	5
A V	$P-D$	6313	2	..

The most conspicuous lines belong to oxygen and nitrogen, and are just of the types discussed earlier, arising by transitions from meta-

† *Pub. Astron. Soc. Pacific*, **39** (1927), 295; *Astrophys. J.* **67** (1928), 1.

stable states in the close vicinity of the ground state. In order to make the matter still clearer we reproduce diagrams of the energy levels in O III, O II, and N II which show the coordination of the more conspicuous forbidden lines, and some others which are permitted.

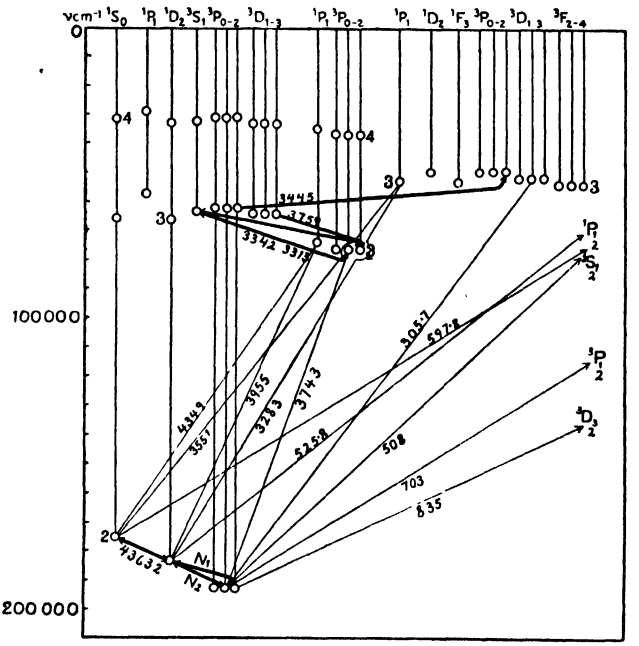


FIG. 41. Term diagram of low energy states in O III. This diagram shows the origin of the chief nebular lines N_1 and N_2, and also a number of permitted lines of much higher excitation energy which appear in nebular spectra. These latter lines are probably excited by the absorption of the resonance line 304 of He II. Cf. Fig. 46.

It appears that the most conspicuous lines originate in oxygen. The most intense pair, the doublet 5,007, 4,559, is emitted by O III, while the ultra-violet doublet 3,726, 3,729, which is also very intense, is emitted by O II. The first doublet lines are called the chief nebular lines, and are denoted by N_1 and N_2 respectively. The relative intensity of the lines in each doublet remains the same in all nebulae, as they should do. It is further clear that the chief nebular lines require a greater stimulus in order to appear than does the ultra-violet doublet. For this reason we should expect that in slitless spectra the nebulae should show smaller disks in the light of the former doublet than in the case of the latter, which they do. Also the

relative intensity of the doublets changes progressively in the same sense as the relative intensity of ionized and neutral helium lines, which is also to be expected.

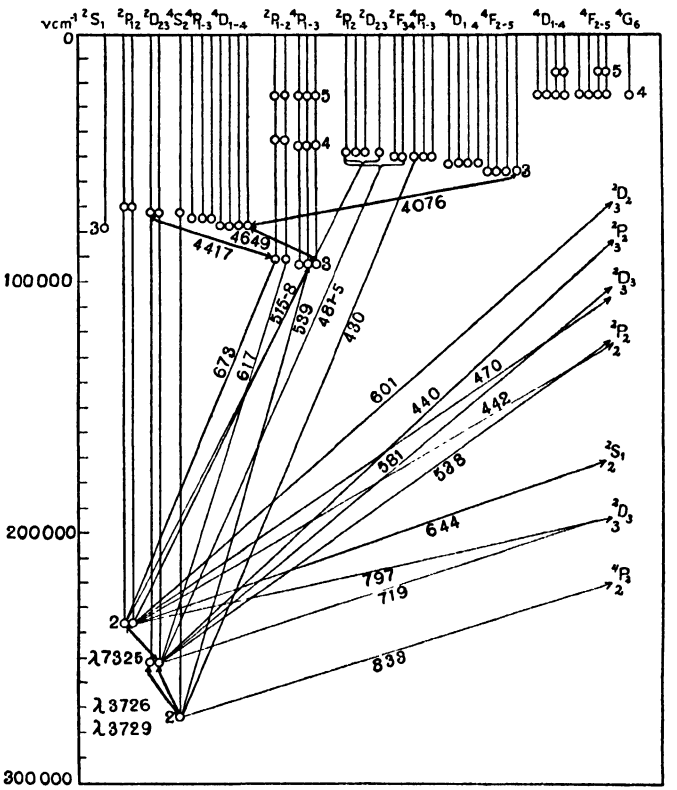


FIG. 42. Level diagram of low energy states in O II showing the origin of some forbidden nebular lines, particularly the ultra-violet doublet 3,726–9.

115. Forbidden Lines in the Spectra of Novae and other Sources

Well past maximum brightness the novae usually begin to show the chief nebular lines, and their spectra begin to resemble that of a planetary nebula in most respects, apart from a great lack of sharpness of the lines. The novae differ from the nebulae, however, in that they also show the green auroral line 5,577 A., which is due to the transition $^1D_2 \rightarrow {}^1S_0$ in O I. The recent nova in Hercules also showed the forbidden O I lines 6,300·23 and 6,363·88.† Also some

† F. J. M. Stratton and A. Beer, *Monthly Not.* **95** (1935), 432.

forbidden lines of CI, NI, FII, FIII, SiI, and SI appear in the spectra of novae, but not in ordinary nebulae. For a full discussion of these lines see the paper by Boyce, Menzel, and Payne.†

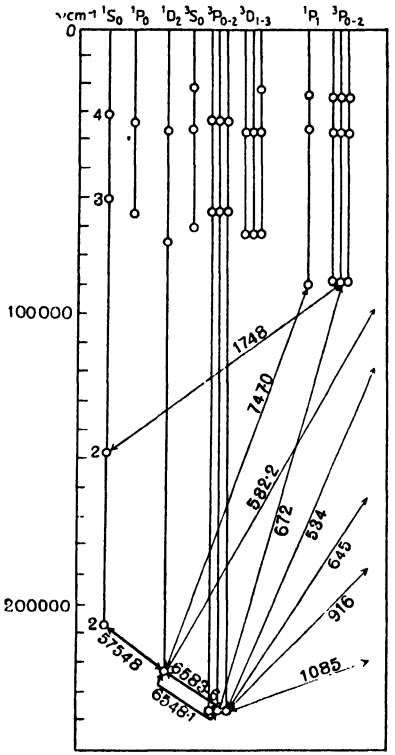


FIG. 43. Level diagram of low energy states in N II.

In some cases forbidden lines have been found in spectra of stars. It is then always a question of peculiar stars which are obviously connected with nebulae of some sort. Merrill‡ discovered forbidden lines in the spectrum of the long-period variable star *R* Aquarii. There is no doubt about the fact that the star is a genuine long-period variable, as it shows the characteristic titanium oxide bands, and bright hydrogen lines. Its nebular lines are the OIII doublet N_1 and N_2, the HeI line 4,471·5, the NeIII lines 3,869, 3,967, the SiII line 4,068, and the unidentified line 4,658·3.

The star is embedded in a luminous nebula, and it is almost the only known case in which such a nebula is physically connected with a late-type star. However, in the years 1921–7 the star has developed a spectrum of the *P* Cygni type, and it is clear that the star has an early-type companion which acts as a nucleus for a planetary nebula.‖ The system bears some resemblance to Hubble's variable nebula N.G.C. 2,261 which is physically connected with *R* Monocerotis.††

The star η Carinae is situated in a region where nebulae abound, and this star is also probably embedded in nebulous stuff. Its magnitude has been subjected to great variations. About 1843 it was one of the brightest stars in the southern hemisphere,‡‡ but

† *Proc. Nat. Acad. Sc.* **19** (1933), 581.
‡ *Astrophys. J.* **53** (1921), 375.
‖ Merrill, *Pub. Astron. Soc. Pacific*, **39** (1927), 48; *Astrophys. J.* **81** (1935), 312.
†† Cf. Slipher, *Lick Obs. Bull.* **3** (1918), 63.
‡‡ R. T. A. Innes, *Ann. Cape Obs.* **9** (1903), 78.

then it began to decline; in 1856 its brightness dropped suddenly, and in the course of some years it became invisible to the naked eye. It suffered another sharp decline in 1894. The spectrum of this star is variable and peculiar. According to Bok† its spectrum in 1892–3 was *cF* 5 with very bright hydrogen lines and faint iron spark lines in emission. In 1895 the continuous background was very faint, and the intensity of the bright lines had increased. Merrill‡ has found that about 20 of the iron lines appearing in the later cycle of the star are due to forbidden transitions. According to Merrill most of these lines appear also in the spectrum of the *Be* star H.D. 45,677.

These stars all seem to represent some sort of intermediate stages between gaseous nebulae and stars. The novae are particularly interesting in this respect, as the close study of the development of the nova spectrum reveals how the transition from a stellar to a nebular state takes place. The appearance of a definite forbidden line in a nova is clearly an indication that the density of the expanding shell of gas surrounding a nova has reached a definite limit, depending on the transition probability and excitation energy of the forbidden line in question. The detailed discussion of this question seems to promise considerable information about the constitution of novae.

† *Pop. Astron.* **38** (1930), 399. Cf. also Bok's dissertation, Groningen, 1932.
‡ *Astrophys. J.* **67** (1928), 391.

XXII

THEORY OF RADIATIVE TRANSFORMATIONS IN NEBULAE

116. Ionization Theory of Nebulae

THE gaseous nebulae bring out in a more pertinent form than the stars in general the problem of the radiative equilibrium of a tenuous gas, since we are here obviously dealing with very radical departures from a state of thermal equilibrium. It is only in this way that it is possible to stimulate forbidden transitions to such an extraordinary degree as is actually observed.

This problem has several aspects. First of all it is already surprising to find atoms of such a high degree of ionization as O^{++} and N^{++} playing an important part in nebular radiation. The solution of this problem is obviously to be sought, on the one hand, in the excessively low density which slows down recombination processes very much. On the other hand it is connected with the very high temperature of the embedded stars, which provides radiation of very great ionizing power.

We shall now show how these two causes in combination will suffice to explain the observed state of ionization.† We consider a gaseous mass at great distances from a hot star. Let n_0 and n_1 denote the number of atoms in the normal state and the ionized state per unit volume and neglect all intermediate states. In our earlier notation the state is then fixed by the single equation

$$n_0 a_{01} = n_1 a_{10}. \tag{135}$$

In order to calculate the coefficients a_{01} and a_{10} we must know the absorption coefficient of the atom in the continuous region beyond the series limit of the principal series, and also the corresponding emission coefficient for the converse process. The former we denote by σ_ν and the latter by $\beta_\nu\left(1+\frac{c^3}{8\pi h\nu^3}\rho_\nu\right)$, where ρ_ν is the energy density of the radiation field. The stimulating effect of the radiation field is as usual given by the factor $\left(1+\frac{c^3}{8\pi h\nu^3}\rho_\nu\right)$.

The quantity β_ν is proportional to σ_ν by the Einstein relations con-

† A. S. Eddington, *Proc. Roy. Soc.* A, **111** (1926), 424.

necting emission and absorption coefficients. It is not easy to see directly how these relations are to be modified in the continuous region, and it is therefore better to derive them *de novo*.

For this purpose we consider a state of thermal equilibrium, so that the ratio n_0/n_1 is known. The above equation then gives immediately the ratio of a_{10}/a_{01}, which is all we want. The number of absorbed quanta per unit time and volume is

$$n_0 \sigma_\nu c \rho_\nu \, dv/h\nu.$$

The corresponding number of recombination processes is, on the other hand,

$$n_1 \beta_\nu \left(1 + \frac{c^3}{8\pi h\nu^3} \rho_\nu\right) v \, dn_e,$$

where dn_e is the number of free electrons in unit volume, the velocity v of which is confined to the interval v to $v+dv$. The ions are assumed to be so heavy that their velocity may be neglected in the description of the recombination process. The velocity v is connected with the frequency ν of the absorbed radiation by the relation

$$h\nu = \tfrac{1}{2}\mu v^2 + \chi,$$

where χ is the ionization energy of the atom in the given state. Hence

$$dv = \frac{\mu v}{h} dv.$$

We need also the usual statistical formulae, which need no further explanation here (see p. 172),

$$dn_e = n_e \frac{4\pi\mu^3}{f_e h^3} e^{-\frac{1}{2}\mu v^2/kT} v^2 \, dv; \quad \frac{n_1 n_e}{n_0} = \omega_0 f_e e^{-\chi/kT}; \quad \rho_\nu = \frac{8\pi h\nu^3/c^3}{e^{h\nu/kT} - 1}.$$

The condition of equilibrium

$$n_0 \sigma_\nu c \rho_\nu \frac{dv}{h\nu} = n_1 \beta_\nu \left(1 + \frac{c^3 \rho_\nu}{8\pi h\nu^3}\right) v \, dn_e \tag{136}$$

then assumes the form

$$\beta_\nu = \frac{\sigma_\nu 2h^2\nu^2}{\mu^2 c^2 v^2 \omega_0}, \tag{137}$$

which is the desired relation.

When conditions do not correspond to thermal equilibrium, equation (136) must be multiplied by $h\nu$ and integrated over all frequencies on the left and all electron velocities on the right:

$$n_0 c \int_0^\infty \sigma_\nu \rho_\nu \, d\nu = n_1 \int_{v=0}^\infty \beta_\nu \left(1 + \frac{c^3 \rho_\nu}{8\pi h\nu^3}\right) h\nu v \, dn_e.$$

This equation expresses the conservation of energy for all transitions out of and into the given state. The transition coefficients a_{01} and a_{10} are therefore given by the expressions

$$a_{01} = c\int_0^\infty \sigma_\nu \rho_\nu \, d\nu \quad \text{and} \quad a_{10} = \frac{2h^3}{\mu^2 c^2 \omega_0}\int_0^\infty \sigma_\nu \nu^3 \left(1+\frac{c^3\rho_\nu}{8\pi h\nu^3}\right)v^{-1}\frac{dn_e}{dv}\,dv. \quad (138)$$

Let us apply these formulae to the following simple case. The radiation field is assumed to consist of dilute thermal radiation, so that

$$\rho_\nu = W\frac{8\pi h\nu^3}{c^3}(e^{h\nu/kT}-1)^{-1},$$

in the usual notation. The velocity-distribution of the electrons we assume to be given by Maxwell's distribution law, corresponding to a temperature T_0, which is different from the temperature T of the star. The absorption coefficient σ_ν is taken to be zero for all frequencies smaller than ν_m. For larger frequencies it is supposed to fall off proportionally to ν^{-3}, which is known to give good approximations for continuous X-ray absorption. On these assumptions we find for a_{01}:

$$a_{01} = W\frac{8\pi\sigma_0}{c^2}kT\log(1-e^{-h\nu_m/kT})^{-1}; \quad \sigma_\nu = \sigma_0\nu^{-3}; \quad \nu > \nu_m = \frac{\chi}{h}.$$

In the calculation of a_{10} it is obviously permissible to neglect the term in ρ_ν on account of the smallness of the dilution factor W. The integration then gives

$$a_{10} = \frac{8\pi\sigma_0 n_e kT_0}{\omega_0 c^2 f_e}.$$

Remembering that $f_e = (2\pi\mu kT_0)^{\frac{1}{2}}h^{-3}$ we find as the new equation of dissociation, on introducing our expressions for a_{01} and a_{10} into (135),

$$\frac{n_1 n_e}{n_0} = \frac{W\omega_0(2\pi\mu kT)^{\frac{1}{2}}}{h^3}\sqrt{\left(\frac{T_0}{T}\right)}\log(1-e^{-h\nu_m/kT})^{-1}. \quad (139)$$

A very much more elaborate theory of ionization in nebulae by Ambarzumian† leads to a nearly equivalent formula. The difference consists in replacing $\log(1-e^{-h\nu_m/kT})^{-1}$ by $e^{-h\nu_m/kT}$, and $\sqrt{(T_0/T)}$ by $2\tau_0+1$, τ_0 denoting the optical thickness of the nebula below the outer boundary in the continuous region.

It appears from this equation that the temperature of the gas does

† *Monthly Not.* 95 (1935), 469.

not influence the equilibrium very much—the dominating temperature is that of the exciting star. Since

$$\log(1-e^{-h\nu_m/kT})^{-1} \sim e^{-h\nu_m/kT},$$

in practical cases, one really passes from the state of ionization in thermal equilibrium at a temperature T to that of a nebula at temperature T_0 by reducing all particle densities in the ratio $W\sqrt{(T_0/T)}$. We shall soon indicate arguments which show that the T_0 to be introduced into this formula is not much different from T, say $\sqrt{(T/T_0)} = 2$. The main interest therefore centres on the dilution factor W.

117. The Trifid Nebula

A typical instance is afforded by the Trifid nebula in Sagittarius. This is a roundish object, which is probably illuminated by an embedded O star of the seventh magnitude. Since the absolute magnitude of O stars is about $-3{\cdot}5$, it follows from the usual relation between parallax and absolute and apparent magnitude that the distance of the nebula is of the order 3,300 light years. A moderate estimate of 30,000° K. for the temperature of the star then leads to a radius of 7·5 solar radii. The apparent diameter of the nebula is 10′, which with the given distance corresponds to a linear diameter of about ten light years. Let R denote the radius of the star and r the distance from the star to the point at which the ionization is calculated. The dilution factor W is then given by

$$W = R^2/4r^2,$$

and the value of this factor at a distance $\frac{1}{2}$ parsec from the star is found to be $2{\cdot}7.10^{-14}$.

The spectrum of this nebula shows, according to Hubble, the chief nebular lines of O III with about the same intensity as the ultraviolet doublet 3,726–9 of O II. As a rough guess we may then assume that the O^{++} atoms have the same abundance as the O^{+} atoms. The frequency limit ν_m follows in the usual way from the ionization potential of O^{+}, which is 34·94 volts. With these data we find from (139) the following value of the density of free electrons:

$$n_e = 235 \text{ cm.}^{-3}$$

No stress can be placed on this figure, except so far as it indicates the order of magnitude. But remembering that the corresponding figure for the number of air molecules in one cm.^{3} is 270.10^{17}, one

realizes that no tinkering with the data can explain away the excessively low density of the gaseous nebulae.

118. The Temperature in Nebulae

In the above calculation it was assumed that the temperature in the nebulae was roughly one-fourth of the effective temperature of the illuminating star. This somewhat startling assumption needs, of course, further qualification. To give a quantitative theory proves very difficult, however, and we shall be satisfied with the following rather qualitative considerations.

Let us first consider the mechanism by which the free electrons in the nebula acquire their velocities, and how they lose them. The conditions are in this respect rather different from the processes of exchange of velocities in a stationary gas in thermal equilibrium. In the gas both gains and losses of velocity are due to collisions between atoms in the gas. In the nebulae the free electrons will mostly acquire their initial velocity in ionization processes, taking place by absorption of radiation, and they will gradually lose their velocity by collisions with atoms and ions, and at an accelerated rate, leading to capture of the electron by an ion.

The average velocity of an electron on a 'free path' from its release in the absorption process to its capture by an ion will clearly depend very much on the probability of capture as a function of the relative velocity. The expression which we have derived for the coefficient of capture, β_ν, equation (137), shows that the probability increases proportionately to the inverse first power of the velocity—since β_ν varies as v^{-2}, and the number of collisions per unit time is proportional to v. Hence the probability increases very rapidly with decreasing v, and it is evident that during most of the time the electron is free its velocity will be high.

We may estimate this average velocity in the following rough way. According to the theory of Bohr† of the passage of electrified particles through matter, the loss of energy dE of a particle on passing a distance dx is given by an expression of the form

$$dE = -Kv^{-1}\,dx,$$

where v is the velocity of the particle and K a slowly varying function of v, which we here consider as a constant. Putting $E = \frac{1}{2}\mu v^2$ and

† *Phil. Mag.* **30** (1915), 581.

integrating we find for the total range of the particle, that is, the distance travelled from the moment it is started with an initial velocity v_0 until the velocity is reduced to zero:

$$x = \frac{\mu}{3K} v_0^3.$$

For the time spent on this free path we find similarly

$$t = \int \frac{dx}{v} = \frac{\mu}{2K} v_0^2,$$

and the average square of the velocity is

$$V^2 = \frac{1}{t} \int_0^t v^2 \frac{dx}{v} = \tfrac{1}{2} v_0^2.$$

It is easily found that the average energy of the electrons expelled by absorption of dilute ultra-violet radiation of temperature T is kT, so that

$$\tfrac{1}{2}\mu v_0^2 = kT.$$

Defining the temperature of the nebula by the equation

$$\tfrac{1}{2}\mu V^2 = \tfrac{3}{2} k T_0$$

it follows that

$$T_0 = \tfrac{1}{3} T.$$

This, of course, only indicates some kind of order of magnitude. Thus the assumed length of the free path in its dependence on the initial velocity is largely hypothetical. But altering the power of v in this expression within reasonable limits will not change the result much. The essential point is the basic assumption that at the end of the free path the electron is captured by an ion. When this is dropped conditions will change fundamentally, and T_0 will depend essentially on the coefficient of dilution.

We may get an indication of conditions in this next step by assuming the expelled electrons to remain free so long that most of them will have acquired a Maxwellian velocity-distribution as a consequence of collisions, and that the diffuse radiation from the gas corresponds to the temperature T_0 according to the laws of Kirchhoff and Planck. This is the case first considered by Fabry,† which for the first time brought out the fact that a cosmic nebula may have a high temperature, provided only that the temperature of the illuminating star is sufficiently high. As a first approximation we then

† *J. de Physique*, **6** (1916), 207.

assume the energy absorbed per second by a slab of gas of unit volume to be

$$\Omega \int_{\nu_m}^{\infty} \sigma_\nu c\rho_\nu(T)\, d\nu,$$

and the energy emitted by the gas in the same time to be

$$4\pi \int_{\nu_m}^{\infty} \sigma_\nu c\rho_\nu(T_0)\, d\nu.$$

Here Ω is the solid angle subtended by the star at the point in question. In the emission this is replaced by 4π since the emission is uniform in all directions. Under stationary conditions these two quantities must be equal, or

$$W \int_{\nu_m}^{\infty} \sigma_\nu \rho_\nu(T)\, d\nu = \int_{\nu_m}^{\infty} \sigma_\nu \rho_\nu(T_0)\, d\nu,$$

since $\Omega/4\pi$ is just the coefficient of dilution W. Assuming as usual σ_ν proportional to ν^{-3} on the short wave-length side of $\nu = \nu_m$, and zero on the long wave-length side, we find

$$T_0 = WT \frac{\log(1-e^{-h\nu_m/kT})}{\log(1-e^{-h\nu_m/kT_0})}.$$

For temperatures not higher than 40,000° K., and absorption limits at least as far into the ultra-violet as for the Lyman series, it is permissible to expand the logarithms and to retain the first term of the series only. Hence the above equation may be brought to the form

$$\frac{T-T_0}{T_0} = \frac{kT}{h\nu_m} \log(T_0/WT).$$

For the Trifid nebula considered above, assuming the typical element of the nebula to be hydrogen—which is true, the relevant data will be

$$T = 30{,}000° \text{ K.}; \quad W = 2{\cdot}7.10^{-14}; \quad h\nu_m = 2{\cdot}16.10^{-11} \text{ ergs.}$$

Solving the above equation for T_0 in successive approximations, one finds the corresponding value of T_0 to be about 4,500° K.

119. Transformation of Radiation in a Nebula

As a next step in the analysis of nebulae we may try to follow up the cycle of transformations suffered by the incident star light in the nebula. This latter we may picture as a vast expanse of hydrogen

gas for the most part, some helium, and a small amount of air and other gases. The gases may be assumed to be arranged in plane parallel strata, and the illuminating star to be well outside the gas. This simplified picture will do both for diffuse and planetary nebulae, as it is in most cases possible to approximate to the spherical shells of the latter by planes, as far as the transmission of radiation is concerned.

On the side of the nebula nearest to the star the incident radiation will at once begin to become depleted by absorption in the Lyman series, and in the continuum extending beyond the head of this series, which we shall refer to as L_c, meaning the Lyman continuum. As a consequence of this absorption there will be created, firstly, a certain number of atoms in excited states, and next a certain number of free electrons. The excited atoms will very rapidly regain the normal state, after having suffered a series of cyclical changes, in which the various series of the hydrogen spectrum are emitted. Considering the very low intensity of the incident star light, and the known short life-periods of the hydrogen states, it is at once clear that the number of hydrogen atoms in excited states in a nebula is excessively small, and would, when interpreted thermodynamically correspond to a very low temperature.

It is different with the absorption in the Lyman continuum by which free electrons are created. Here the mean free life-period of the ionized state will depend on the gas-density. If this is sufficiently small, the life-period, or time spent on a free path, may become very long, and the gas may simulate a high temperature, as was set out more elaborately above.

Here it interests us more to know what happens to the energy of the expelled electrons. As long as the electron is free, it spends its energy in collisions, where it is partly emitted as low-frequency continuous radiation, but mostly imparted to other particles, and to some degree is used for exciting atoms to higher states. This may be particularly the case with atoms of the type O^+, O^{++}, etc., in their normal state, where the excitation energy of the next state is very low—much lower than the lowest excitation energy of hydrogen. Finally, however, the electron gets caught by an ion, mostly a free hydrogen nucleus. It may choose among all the available states of the atom, but once caught it will very soon by some route arrive back in the normal state.

By such circular processes the incident dilute radiation of the star is converted into radiation of longer wave-lengths. It is in this way that the hydrogen emission in the nebulae, the Balmer series, and the continuous radiation beyond the head of this series is generated. In so far as hydrogen is concerned, the nebula is nearly perfectly transparent to light emitted in the Balmer series and higher series, since the number of atoms in quantum states above the normal state is vanishingly small. This radiation is therefore lost to the further balance of radiant energy in the nebula.

The radiation which passes into Lyman lines will suffer further transformations by reabsorption. A Lyman-β, for instance, will soon be reabsorbed, whence it may either return as Lyman-β or it may return as an $H\alpha$ and an $L\alpha$. The $H\alpha$ generated in this way is lost to the nebula. The $L\alpha$, on the other hand, will go on being reabsorbed and re-emitted all the way through the nebula, diffusing slowly outwards, and finally escaping when the surface has been reached.

We are thus led to consider the simplified problem of a hydrogen nebula, when all transitions are reduced to a three-state cycle, in which only the first, the second, and the Lyman continuum participate, the continuum playing the same part as an ordinary quantum state.

The theory of such cycles was considered earlier in connexion with the problem of the hydrogen lines in the sun. In that case the incident radiation is so little diluted that before the lines have become deepened somewhat the probability of a closed set of transitions does not differ much for two opposite cycles. For a nebula, however, the probability is overwhelmingly in favour of a unidirectional cycle already from the start. In order to realize this we may, in the theory of a three-state cycle given earlier (cf. § 48), take states 1 and 2 to represent the first and second quantum states of the hydrogen atom, and state 3 to represent the Lyman continuum. It was then shown that the ratio of the probability of a successive climbing up towards the ionized state and with a single return jump, $a_{12}a_{23}a_{31}$, is smaller than a direct jump to the continuum with a step-wise return, $a_{13}a_{32}a_{21}$, by a factor

$$y = \frac{WF_{31}}{F_{32}F_{21}}.$$

Here W is the dilution coefficient and F is given by

$$F_{ik} = 1-(1-W)e^{-h\nu_{ik}/kT},$$

while ν_{ik} denotes the radiation frequency corresponding to a transition $i \to k$. Provided each $h\nu_{ik}$ is in no case vanishingly small in comparison with kT, it follows that y is, for the case of a nebula, a negligible quantity. We have in fact seen earlier that the order of magnitude of the dilution coefficient W in a nebula is of the order of 10^{-13}.

It is interesting, however, that the further course of events in a nebula is the opposite of the course in a stellar atmosphere. In the sun the gradual working out of the spectral lines serves to intensify the difference in probability of opposite cycles. In the nebulae, on the other hand, the transformation of radiation will serve to decrease the probability a_{13} by depleting the continuous flux in the Lyman continuum, and to increase a_{23}, by throwing new radiation into the Lyman-α line. By this process the ratio of the probabilities may be reduced by as much as a factor 10^5, but still this does not alter the main course of events.

120. Theory of a Three-State Cycle

It is worth while to go into some detail with regard to the transformation in a hydrogen nebula in order to substantiate the above conclusions quantitatively. To fix the ideas we consider a star enveloped by a concentric, thin, spherical shell of gas, representing the nebula. Assuming the shell to be sufficiently thin in comparison with its radius, we may use the equation of radiative transfer for flat layers

$$\mu \frac{dI}{dz} = -\kappa I + E,$$

where the symbols have their usual significance.

Since we have seen that transitions to the continuum from state 2 are of negligible importance, we are left with resonance transitions in Lyman-α, and transitions to the Lyman continuum and back. In this latter case a continuous range of frequencies come into play, which means a very great complication to the theory. In a first survey of this theory the practice has therefore been introduced of working with an 'average' frequency ν for the continuum, and throwing all the difficulties of the problem into the calculation of this average value. That such a procedure only means a rather rough approximation need scarcely be emphasized. But the state of the theory in general is such that it justifies the use of simplifications of this kind.

Consider first the transmission of radiation in the continuum. The emission function E has then the form

$$4\pi E = n_3 a_{31} h\nu.$$

The equation of equilibrium for the transitions to and from state 3 is

$$n_3(a_{31}+a_{32}) = n_1 a_{13}.$$

Transitions from state 2 to the continuum have been neglected, because we have seen that a cycle in this direction has a negligible probability. Hence it follows that

$$n_3 a_{31} = \frac{a_{31}}{a_{31}+a_{32}} n_1 a_{13}.$$

Writing

$$\frac{a_{31}}{a_{31}+a_{32}} = p,$$

the emission function becomes

$$4\pi E = p h\nu n_1 a_{13}.$$

But $h\nu n_1 a_{13}$ is just the energy absorbed per second per unit volume from the radiation field by atoms in the lower state. It is necessary to distinguish between the diffuse radiation in the nebula and the flux of radiation from the central star. Denoting this latter flux by S on the inside boundary of the nebula, the flux at optical height τ above the inner boundary is clearly

$$Se^{-\tau}.$$

Since the radiation from the star is practically parallel at the distance of the nebula, there is no need to distinguish between flux and total intensity.

In case of the diffuse flux, on the other hand, it is necessary to make such a distinction, and we write as usual F for the flux and J for the total intensity. Having already denoted the absorption coefficient by κ, it follows that the emission function E has the form

$$4\pi E = p\kappa(J+Se^{-\tau}).$$

The equation of radiative transfer has consequently the form

$$\mu \frac{dI}{d\tau} = -I + \frac{p}{4\pi}(J+Se^{-\tau}),$$

where

$$J = 2\pi \int_{-1}^{+1} I \, d\mu,$$

I being the intensity of diffuse radiation.

In order to solve this equation we proceed in the usual way. Multiplication by $2\pi\, d\mu$ and integration with respect to μ gives

$$\frac{dF}{d\tau} = -J + p(J + Se^{-\tau}),$$

while multiplication by $2\pi\mu\, d\mu$ and integration gives

$$F = -2\pi \frac{\partial}{\partial \tau} \int_{-1}^{1} I\mu^2\, d\mu = -\frac{1}{3}\frac{\partial J}{\partial \tau},$$

since the diffuse radiation is distributed over various directions to such a degree that the usual approximation

$$2\pi \int_{-1}^{1} I\mu^2\, d\mu = \tfrac{1}{3} J$$

may be adopted. We are then left with the equation

$$\frac{d^2J}{d\tau^2} = 3(1-p)J - 3pSe^{-\tau} = \lambda^2 J - 3pSe^{-\tau}, \tag{140}$$

the solution of which is

$$J = Ae^{\lambda\tau} + Be^{-\lambda\tau} + \frac{3pSe^{-\tau}}{\lambda^2 - 1},$$

where A and B are integration constants, and λ is given by

$$\lambda^2 = 3(1-p).$$

Since the corresponding flux is $-\frac{1}{3}\frac{\partial J}{\partial \tau}$, we find

$$F = -\tfrac{1}{3}\lambda(Ae^{\lambda\tau} - Be^{-\lambda\tau}) + \frac{pS}{\lambda^2 - 1} e^{-\tau}.$$

The case $\lambda^2 = 1$, i.e. $p = \frac{2}{3}$, demands separate treatment. Putting $p = \frac{2}{3}$ in the original equation one finds

$$\frac{d^2J}{d\tau^2} = J - 2Se^{-\tau},$$

the solution of which is

$$J = Ae^{\tau} + Be^{-\tau} + S\tau e^{-\tau}.$$

Further the corresponding flux is

$$F = -\tfrac{1}{3}(Ae^{\tau} - Be^{-\tau}) - \tfrac{1}{3}S(1-\tau)e^{-\tau}.$$

121. The Boundary Conditions

In order to determine the integration constants A and B it is necessary to specify the conditions to be satisfied at the outer and inner boundary of the shell.

On the outer boundary the conditions are clearly the same as in the case of a stellar atmosphere. Denoting the optical thickness of the shell by τ_0 we must therefore have

$$J(\tau_0) = 2F(\tau_0),$$

which ensures that there will be (practically) no back radiation at the boundary.

The condition to be satisfied on the inner boundary is the vanishing of the diffuse flux. This is a consequence of the conservation of energy, no diffuse radiation being generated inside the empty shell. This means that the diffuse intensity J will be constant inside the shell.

The condition at the inner boundary gives the equation

$$A = B + \frac{3pS}{\lambda(\lambda^2-1)}.$$

On the outer boundary the condition is

$$A(1+\tfrac{2}{3}\lambda)e^{\lambda\tau_0} + B(1-\tfrac{2}{3}\lambda)e^{-\lambda\tau_0} + \frac{pSe^{-\tau_0}}{\lambda^2-1} = 0,$$

which gives the explicit expression of B:

$$B = -\frac{3pS}{\lambda(\lambda^2-1)} \frac{(1+\tfrac{2}{3}\lambda)e^{\lambda\tau_0} + \tfrac{1}{3}\lambda e^{-\tau_0}}{(1+\tfrac{2}{3}\lambda)e^{\lambda\tau_0} + (1-\tfrac{2}{3}\lambda)e^{-\lambda\tau_0}}.$$

The corresponding equations for the critical value $p = \frac{2}{3}$ are

$$B = A + S$$

and

$$A = -\frac{S(\tau_0+3)e^{-\tau_0}}{e^{-\tau_0}+5e^{\tau_0}}.$$

Chandrasekhar† has made extensive calculations of the emergent flux under different conditions. In Figs. 44 and 45 we reproduce some of his diagrams.

122. The Resonance Line

We imagine the profile of the line to have a flat-topped columnar form, so the absorption coefficient in the line is a constant, which we denote by $\omega\kappa$. The corresponding optical height is $\omega d\tau$, and the equation for the total intensity in the line becomes, by the usual machinery of approximations,

$$\frac{d^2J'}{d\tau^2} = 3\omega^2(J' - 4\pi E/\omega\kappa),$$

† *Zeits. f. Astrophys.* 9 (1935), 266.

where a dash indicates that the quantities belong to the radiation in the resonance line.

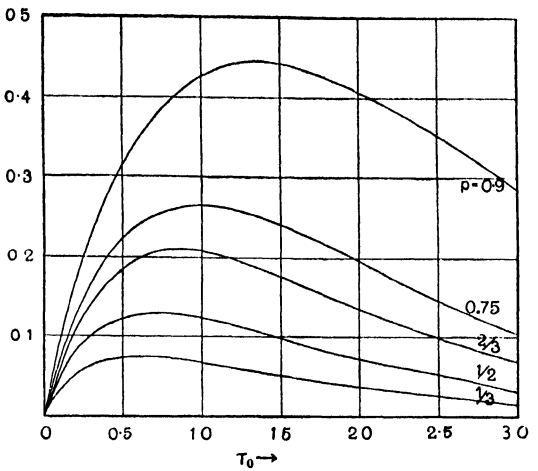


FIG. 44. These curves give the emergent diffuse ultra-violet flux for different values of p.

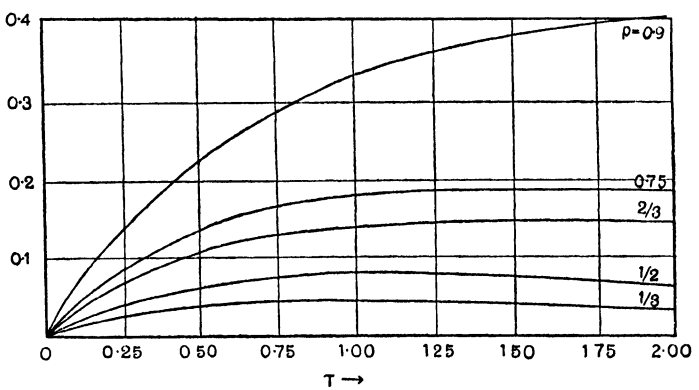


FIG. 45. Diffuse flux in units of S for different values of p. The optical thickness is assumed to be $\tau_0 = 2$.

The emission function E is by definition given by

$$4\pi E = h\nu n_2 a_{21}.$$

The equation of equilibrium of the normal state may be written

$$n_2 a_{21} = n_1 a_{12} + n_1 a_{13} - n_3 a_{31} = n_1 a_{12} + (1-p) n_1 a_{13}.$$

Multiplying by $h\nu'$ it is seen that the first term on the right is just the total absorption in Lyman-α, and may therefore be written

$\omega\kappa J'$. Similarly, the second term on the right becomes $(1-p)\nu'/\nu$ times the total absorption in the continuum, which was written as $\kappa(J+Se^{-\tau})$. Hence, finally,

$$4\pi E = \omega\kappa J' + (1-p)\frac{\nu'}{\nu}\kappa(J+Se^{-\tau}),$$

and the equation determining J' becomes

$$\frac{d^2J'}{d\tau^2} = -\omega\frac{\nu'}{\nu}(\lambda^2 J + \lambda^2 Se^{-\tau}),$$

on remembering that $\lambda^2 = 3(1-p)$. Expressing $\lambda^2 J$ by $d^2J/d\tau^2$, on using (140) we find

$$\frac{d^2J'}{d\tau^2} = -\omega\frac{\nu'}{\nu}\left(\frac{d^2J}{d\tau^2} + 3Se^{-\tau}\right).$$

Integrating twice we find

$$J' = a + b\tau - \omega\frac{\nu'}{\nu}(J + 3Se^{-\tau}),$$

where a and b are integration constants. The corresponding expression for the flux is

$$F' = -\frac{1}{3\omega}\frac{\partial J'}{\partial\tau} = -\frac{b}{3\omega} - \frac{\nu'}{\nu}(F + Se^{-\tau}).$$

Since F and F' both vanish at the inner boundary,

$$b = -3\omega\frac{\nu'}{\nu}S,$$

and the expression for the flux becomes

$$F' = \frac{\nu'}{\nu}\{S(1-e^{-\tau}) - F\}.$$

At the outer boundary both $J = 2F$ and $J' = 2F'$, which demands that

$$a = \frac{\nu'}{\nu}(\omega-1)J(\tau_0) + \frac{\nu'}{\nu}S\{3\omega\tau_0 + 2 + (3\omega-2)e^{-\tau_0}\}.$$

The problem is thus formally solved.

The above expression for the flux in Lyman-α might have been written down directly. But we have found it better to let it appear in its proper place. Apart from close to the inner boundary, F' is clearly of the order of magnitude of $S\nu'/\nu$, which shows how the radiation absorbed in hydrogen degenerates into Lyman-α radiation. The above expression for F' may also be written in the form

$$S = Se^{-\tau} + F + F' + \left(\frac{\nu}{\nu'} - 1\right)F'.$$

This equation expresses the incident flux as the sum of the flux of direct star light, $Se^{-\tau}$, the continuous flux F, the Lyman-α flux F', and an additional flux

$$F'' = \left(\frac{\nu}{\nu'} - 1\right) F' = \left(1 - \frac{\nu'}{\nu}\right)\{S(1-e^{-\tau}) - F\},$$

which clearly gives the flux of all other radiation emitted in the recombination processes, which do not go directly to the ground state.

This theory of a three-state cycle cannot be used for a calculation of the intensities of the Balmer lines, which are important data in the theory of nebulae. An approximate theory of the relative intensities of these lines was developed by Cillié,† by considering a seven-state cycle of transitions. But in that case it did not prove possible to cast the theory into such a clear form as is the case with the theory of a three-state cycle.

123. Radiation Field in an Expanding Nebula

The above results are only applicable to a nebula at rest, and will suffer serious modifications for an expanding nebula. Unfortunately, the theory is much less developed in this case, which is regrettable, because several interesting problems awaiting a solution are intimately connected with the internal motion in the nebula.

The effect of motion will be less important for the spectral region beyond the absorption limits than for the field of the spectral lines. In the former case a small shift of the absorption limit by Doppler effect does not make much difference, but in the latter the shift may take an atom out of the region in which the emergent flux is appreciable, so that it is unable to absorb energy from the flux.

At present only first attempts at a theory of reabsorption in expanding envelopes are available,‡ and it is therefore better here only to consider the qualitative aspects of the problem.

The amount of reabsorption in an expanding nebula will depend intimately on the distribution of macroscopic velocities on the one hand, and the intrinsic width of the spectral line under consideration, as it is determined by damping or thermal Doppler effect, on the other. The effect will also be very different according as the nebula is 'optically thin', as was assumed in the theory of the contours of

† *Monthly Not.* **92** (1932), 820.

‡ H. Zanstra, ibid. **95** (1934), 84; J. Woltjer, Jr., *Bull. Astron. Inst. Netherlands*, **7** (1934), 217.

emission lines in novae and *WR* stars, or 'optically thick', as was assumed in the preceding theory of radiative equilibrium of a nebula at rest.

For the problem of the reabsorption effect in observed emission lines of novae and *WR* stars the interesting case is clearly that of a thin nebula. The reduction of the spherical problem to a plane one has then no longer any meaning, since the mean free path of a quantum is of the order of the cross-section of the nebula itself.

Consider the two cases discussed earlier of a uniformly accelerated nebula and a uniformly decelerated nebula, as these two opposite cases will bring out the main point. For the problem of reabsorption of a given line it is clearly essential that the intensity in the diffuse radiation shall fall within the region of the line, as seen from a coordinate system at rest relative to the atom. In an optically thin envelope this requires that the radial velocity along a given line of sight shall assume the same value at several different points.

This is not the case when the envelope is uniformly accelerated outward, but it may take place when the outward motion of the atmosphere is uniformly decelerated. In order to see this we write the radial velocity, in units of the velocity at infinity, in the form

$$u = \cos\theta\sqrt{(1+A/r)},$$

where A is positive for a decelerated and negative for an accelerated envelope. Along a given line of sight we have

$$\sin\theta = r_0/r,$$

where r_0 is the perpendicular distance from the stellar centre to the line of sight. Hence

$$u = \cos\theta\sqrt{\left(1+\frac{A}{r_0}\sin\theta\right)}.$$

When A is negative, u increases uniformly when θ varies from π to 0, and no real maximum of u will be found. When A is positive, u has two maxima, symmetrically situated about the zero position, and determined by the equation

$$\sin\theta = \frac{r_0}{3A}\{\sqrt{(1+3A^2/r_0^2)}-1\}.$$

On both sides of a maximum there will exist corresponding regions with the same value of the radial velocity, and the radiation emitted from the farther half along the direction of the line of sight may be absorbed in the nearer region.

No reabsorption can occur for $|u|$ less than the limiting velocity at infinity. Hence an inner core of the line, the width of which measured in velocity units will be twice the velocity at infinity, will remain unaffected. Outside this core reabsorption will take place right out to the border of the line. The reabsorption effect considered here will therefore preferentially be effective in the line wings. The resulting line profile will, moreover, be influenced by the occultation effect, especially in the red wing.

All preceding considerations involve the assumption that the nebula is optically thin in the frequency range under consideration. If the nebula is optically thick, the reabsorption depending on the distribution of macroscopic velocities will become negligible, and the influence of the intrinsic width of the line will become of the first importance. This case, which is of interest in connexion with the radiation pressure due to resonance absorption, was discussed by Zanstra,† who showed that the reduction in radiation pressure due to the motion under certain circumstances is in the ratio of the square of V/v, V being the macroscopic velocity and v the mean thermal velocity in the gas.

124. Excitation of Permitted Lines of O III and N III

The admixture of small amounts of foreign gases to the hydrogen content does not change the essential features of the problem. The effect will in all cases be that radiation is transformed into the lowest resonance line of the element, which is kept, so to say, imprisoned in the nebula and only leaks slowly away, while the rest of the transformed radiation is lost at once.

There is one interesting exception to this state of affairs, which was recently pointed out by Bowen,‡ and which serves to emphasize the correctness of the analysis. This is the case in which there is an accidental coincidence of some lines emitted by one element with lines absorbed by another. This is the case with a mixture of He^+ with ionized air. The lowest resonance line of He^+ has the wave-length 303·780 A. and this wave-length lies in the close vicinity of the O III line 303·799 A. There is no necessity for an exact coincidence for the effect to become important, since a moderate relative motion of the different parts of the nebula will suffice to bring the frequencies of some O III atoms into the critical frequency interval.

† *Monthly Not.* **95** (1934), 84. ‡ *Astrophys. J.* **81** (1935), 1.

This coincidence is of particular interest, as it appears to give the clue to the appearance of some permitted lines of O III, which otherwise would be very difficult to explain. These lines are indicated in the diagram, Fig. 46. The process of excitation of these lines is then first the generation of a strong emission in the helium resonance line 303·780. The next step is the absorption of this line by O III atoms in the state $2p\,^3P$, which brings the atoms to the state $3d\,^3P$.

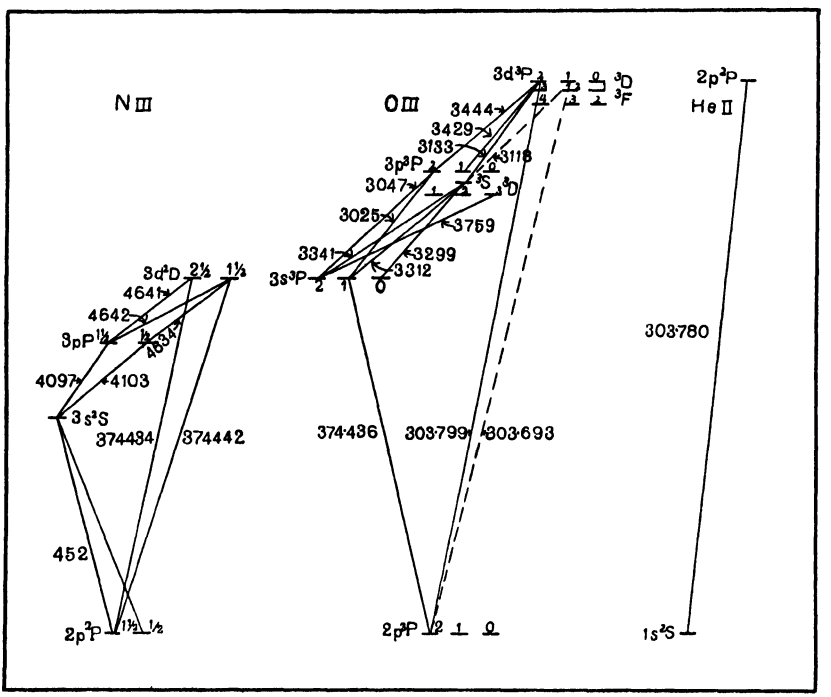


FIG. 46. Diagram illustrating the method of excitation of permitted lines of N III and O III in nebular spectra. The He II resonance line 303·780 is represented on the right in the figure. Owing to its resonance character an abnormal amount of radiation will be thrown into this line. By a chance coincidence this line is absorbed by the $3d\,^3P$–$2p\,^3P$ line in O III, the difference in wave-length being small enough to be adjustable by Doppler effect. This absorption starts the emission of an ultra-violet line 374·436, which again sets a similar cycle going in N III, which is shown on the left (Bowen, *Astrophys. J.* **81**).

From this state the atoms then suffer permitted spontaneous transitions to $3p\,^3P$ and farther on to $3s\,^3P$. From this state the only permitted transition leads back to the $2p\,^3P$ state, which makes the cycle a closed one.

It is now a curious fact that the ultra-violet line 374·436 A. emitted

during this last transition lies in the close vicinity of the resonance line 374·422 A. of N III. Here the game repeats itself. The ground state of this latter line is $2p\,^2P$, and the absorption of the line brings the N III atom into a $3d\,^2D$ state, from which it may suffer spontaneous transitions first to $3p\,^2P$ or $3s\,^2S$, and thence finally back to $2p\,^2P$, from which the cycle started. This cycle also gives rise to a number of lines observed in the spectra of nebulae.

125. Excitation of Forbidden Nebular Lines

The next question is to find out how the forbidden nebular lines are excited. At the first moment it might be thought that these lines are stimulated by a similar cycle, the atoms being lifted to a higher state by absorption of radiation, from which they suffer transitions to the metastable upper states of the forbidden lines from which the forbidden transitions take place.

A cycle of this type cannot take place within the multitude of discrete quantum states because it would necessarily involve forbidden transitions before the initial state of the lines is reached. It is, in fact, necessary for the atoms to include the ionized state in the cycle.

This is important, and for a double reason. First of all we should expect that if the forbidden lines are generated in a cycle including the ionized state some permitted lines might also be included in the cycle. But this does not appear to be true. Apparently all permitted lines are generated by separate cycles which are set going by the absorption of some resonance lines. Secondly, it must be realized that the ionization of atoms like O III requires quite a high energy. The ionization potential of O III is, in fact, 54·8 volts, about four times larger than that of hydrogen, and the ionizing radiation is confined to wave-lengths shorter than 250 A. In order to set the cycle going fast enough for the production of forbidden lines at the observed rate, it is then clear off-hand that quite high values of the stellar temperature are required to get the radiation forced over into sufficiently short wave-lengths.

In this way, however, one gets into difficulties with various attempts at a determination of the temperature of the nuclear stars. It appears quite certain that most stars associated with diffuse emission nebulae have an effective temperature of the order 30,000–40,000° K. But Bowen estimated that in order to stimulate the

forbidden lines by a purely radiative cycle the temperature of the exciting star would have to be of the order 100,000° K.

It seems therefore probable that the purely radiative cycle can be ruled out as a cause of excitation of forbidden lines in nebulae. As a second alternative there is only left the possibility of excitation by electronic impact, which was first suggested by Bowen.† It is then most natural to think of the electrons set free by continuous absorption in hydrogen, which process was discussed earlier. Since the excitation potentials of the forbidden lines are so small, most free electrons in the nebulae will have energy enough to excite an atom to a metastable state.

However, no really quantitative work has thus far been done on this problem, and a closer attention to its details is very much to be desired.

126. Chemical Composition of Nebulae

After this survey of the excitation processes it is now possible to form an opinion on the relative abundance of the elements in the nebulae. It is clear, namely, that no reliable conclusion can be drawn from a comparison of the intensity of a forbidden line and a permitted line when the difference in the excitation process is not taken into account. And a comparison of the Balmer lines of hydrogen with the permitted O III lines would also lead to a misleading view of the relative importance of the two elements. The lines compared must belong to the same class as regards the excitation process before one may hope that the relative intensity will give an indication of the true relative abundance, a point of view which was first emphasized by Bowen.‡ Leaving out forbidden lines, and lines excited by absorption of resonance lines, one is effectively left with lines of hydrogen and helium, two very faint carbon lines, and one oxygen line. This shows clearly the predominance of hydrogen and helium in the nebulae. It is not possible to give definite figures for the abundance of other elements, but it is evident that they are present only in small quantities.

† *Astrophys. J.* **67** (1928), 1. ‡ Ibid. **81** (1935), 1.

XXIII

ABSORPTION BY INTERSTELLAR GASES

127. Detached Lines

THUS far the constitution of nebulae has been studied from the properties of their emission lines. It would be of great importance to be able to study the properties of interstellar gases also by using transmitted light. But this is only possible in some quite exceptional cases.

Two conditions must be satisfied before it is possible to detect interstellar absorption lines. First of all, the element must have absorption lines in the accessible part of the stellar spectrum. Next, stars must be available which have a spectrum free of absorption lines in the given region, of so high a luminosity that their spectra can be photographed even though they are very far off from the earth. The most propitious stars are therefore *B* and *O* stars, *WR* stars, and novae. Which elements to look for it is more difficult to say beforehand, but observations have decided in favour of calcium and sodium.

This happened as early as 1904 when Hartmann† discovered that the *H* and *K* lines in the spectrum of the spectroscopic binary δ Orionis were double. One pair of components he found to oscillate back and forth as they should do in virtue of the orbital motion of the stars. The second pair of lines, however, remained quiet, or very nearly so.

Lines of this sort became known as 'stationary calcium lines', an appellation which some years ago was changed into 'detached lines'. By and by detached lines were discovered in quite a number of early-type binaries. In most cases only the calcium lines *H* and *K* are stationary, but Miss Heger‡ at the Lick Observatory discovered a few cases of detached sodium lines. Recently Merrill‖ has announced the discovery of four new interstellar lines in the red and yellow region of the spectrum. The origin of these lines is thus far unknown. However, the fussy character of the lines makes it probable that they are really narrow molecular bands.††

† *Astrophys. J.* **19** (1904), 268.
‡ *Lick Obs. Bull.* **10** (1919), 59; **11** (1924), 141.
‖ *Pub. Astron. Soc. Pacific*, **46** (1934), 206.
†† H. N. Russell, *Monthly Not.* **95** (1935), 635.

It is not necessary for a star to be a binary in order to reveal stationary lines. It is sufficient that the star has no prominent calcium lines of its own, and has an appreciable radial velocity—appreciable in comparison with the radial velocity of the interstellar gases responsible for the absorption.

128. The Galactic Rotation and the Intensity of Interstellar Lines

The study of interstellar lines is of interest in two respects. First the displacement of the lines from their zero velocity position gives an average measure of the velocity of the interstellar gases, which has served well in the study of the galactic rotation. Next it is possible to obtain some information about the density and temperature of the interstellar gases from the line intensities.

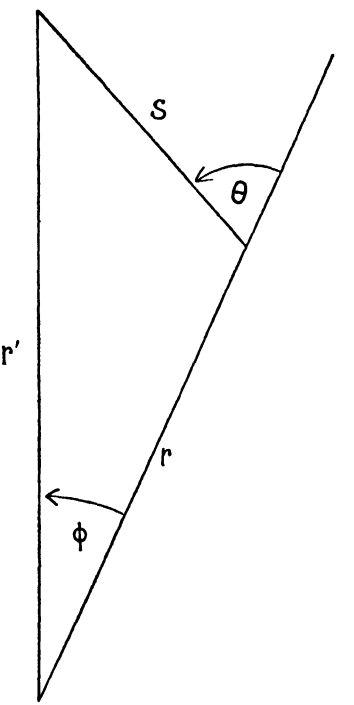


FIG. 44. Orientation of coordinates used in the theory of galactic rotation.

In Oort's theory† of the differential galactic rotation the radial velocity along a given line of sight should increase or decrease proportionately to distance.

It is easy to see why this must be so. Let r and r' denote the distances from the galactic centre to the sun and to a star in the galactic plane respectively, ϕ the angle between r and r', s the distance from the sun to the star, and θ the angle between r and s. Let finally W denote the angular velocity of rotation, which is assumed to be a function of the distance to the galactic centre only.

From the equation

$$s^2 = r^2 + r'^2 - 2rr'\cos\phi$$

we find the radial velocity u of the star to be given by

$$u = \frac{ds}{dt} = s^{-1}rr'\sin\phi\frac{d\phi}{dt},$$

† *Bull. Astron. Inst. Netherlands*, **3** (1927), 275; **4** (1927), 79.

since r and r' remain constant during a circular revolution. But to a first approximation

$$\frac{d\phi}{dt} = W(r') - W(r) = \frac{\partial W}{\partial r} s\cos\theta + \ldots, \qquad r'\sin\phi = s\sin\theta.$$

Hence, finally,

$$u = s\frac{r}{2}\frac{\partial W}{\partial r}\sin 2\theta + \ldots.$$

The observed value of $\frac{1}{2}r\partial W/\partial r$ is about -17 km./sec. per kiloparsec, corresponding to a rotation period of the galaxy of about 250 million years.

It is clear from the symmetry of the problem that, assuming the interstellar gases to participate in the differential galactic rotation, the displacement of a line from its zero position expressed in km./sec. is just half the radial velocity due to differential rotation at the position of the observed star. It is important that this velocity is measurable without measuring the width of the line. This latter is, in fact, also affected by the density and temperature of the absorbing gases, which fact may perhaps lead to a determination of these quantities.

But this second problem is difficult to solve, because the line intensities are very insensitive to changes in density and temperature in the region where the lines are accessible to observations. This point was brought out by Eddington.†

In order to see this clearly we may neglect the Doppler width of the line due to thermal motion, and assume its intrinsic width to be due to radiation damping alone. Write the absorption coefficient in the usual form

$$\kappa_\nu = \frac{K\delta}{\delta^2 + (\nu - \nu_0')^2},$$

ν_0' being the proper frequency of the line relative to the observer. If ν_0 is this frequency relative to an atom of radial velocity u, we have

$$\nu_0' = \nu_0(1 - u/c).$$

We hence find for the optical length of the line of sight

$$\tau_\nu = nK\int_0^s \frac{\delta\, ds}{\delta^2 + (\nu - \nu_0')^2},$$

where n is the number of absorbing atoms in unit volume. Since s is

† *Monthly Not.* **95** (1934), 2.

proportional to u, and hence also to $\nu_0'-\nu_0$, this integration may be carried out, giving

$$\tau_\nu = nK\frac{ds}{d\nu_0'}\left\{\tan^{-1}\frac{\nu-\nu_0'}{\delta}-\tan^{-1}\frac{\nu-\nu_0}{\delta}\right\},$$

where ν_0' corresponds to the distance s of the observed star. This expression may be further simplified by introducing the radial velocity U due to galactic rotation at the distance of the star, and by using the addition theorem of tan functions. Writing

$$\delta\bar{\nu} = \nu-\nu_0, \qquad \delta\bar{u} = \nu_0 U/c,$$

we find easily

$$\tau_\nu = nK\frac{ds}{d\nu_0'}\tan^{-1}\frac{\bar{u}}{1+\bar{\nu}\bar{u}+\bar{\nu}^2}. \tag{141}$$

The profile of the line is now given by the formula for the transmitted intensity:

$$I = I_0 e^{-\tau_\nu}.$$

In order to fix ideas we may assume the radial velocity U to be positive, so that the absorption line is displaced towards the red. Since ν_0/δ is of the order 10^7, $\bar{u}$ will be a large quantity, provided U exceeds one km./sec., which we assume to be the case. The variations of τ_ν with ν is very slow in the interval $\bar{\nu}=0$ to $\bar{\nu}=-\bar{u}$, which corresponds to the Doppler width of the line. In this region the $\tan^{-1}$ expression only varies from nearly $\frac{1}{2}\pi$ at the ends of the interval to nearly π in the middle. Outside this region τ_ν will fall off ultimately as $(\bar{\nu})^{-2}$.

Whether the line width will grow appreciably broader than the Doppler width or not depends on the value of

$$A = nK\frac{ds}{d\nu_0'}.$$

If A is small in comparison with unity, it is possible to write $1-\tau_\nu$ instead of $e^{-\tau_\nu}$, and the intensity of the line,

$$R = I_0\int_0^\infty (1-e^{-\tau\nu})\,d\nu,$$

becomes approximately $\quad R = \frac{1}{2}\pi I_0 A\delta\bar{u},$

since $\delta\bar{u}$ is the Doppler width. In this case the intensity is proportional to the density of absorbing atoms, and this latter quantity is then easy to determine. As a matter of fact this is the only case in which this density can be determined in a trustworthy manner. But

unfortunately it has proved very difficult to find stars where the line intensities fall in this range.

When A is large in comparison with unity, the transmitted intensity will be very nearly zero over the interval in which τ_ν is large. Somewhere in the wing of the line, however, τ_ν has decreased to the order unity, and from then on it drops so fast that we may take the intensity of the line to be given by

$$R = I_0(\nu_2 - \nu_1),$$

where ν_2 and ν_1 are the two values of ν which are roots of the equation

$$\tau_\nu = 1.$$

Since A has been assumed to be large, tan and $\tan^{-1}$ need not be distinguished, this equation assumes the form

$$\bar{\nu}^2 + \bar{\nu}\bar{u} + 1 = A\bar{u},$$

the solution of which is

$$\bar{\nu} = -\tfrac{1}{2}\bar{u} \pm \sqrt{(\tfrac{1}{4}\bar{u}^2 + A\bar{u} - 1)},$$

so that

$$\bar{\nu}_2 - \bar{\nu}_1 = \sqrt{(\bar{u}^2 + 4A\bar{u} - 4)}.$$

The expression of the line intensity becomes

$$R = I_0 \delta \sqrt{(\bar{u}^2 + 4A\bar{u} - 4)}.$$

It should be noted that the term $A\bar{u}$ is independent of galactic longitude, and only dependent on distance. The term $\bar{u}^2$ depends on longitude, however, and as long as $4A$ is small in comparison with $\bar{u}$ we should therefore expect a pronounced longitude effect in the intensity. But this question has not been investigated very closely from the observational side as yet, and the existence of such an effect has thus far not been established.

129. The Density of Interstellar Calcium and Sodium

The determination of density from the line intensities is in this case an elusive task, because the intensity is so closely linked up with the radial velocity. At present the information revealed about the intensity in this way is only sufficient to indicate a lower limit to the density.

In order to see this it is convenient to distinguish between three different ranges of density, as was done by Eddington in his discussion of this question. The first range (range A) extends from zero up to a density corresponding to

$$\tau_\nu = 2,$$

say, where τ_ν corresponds to the theoretical central intensity of the line. The next range of the density (B) extends from this value and to an upper limit where $4A$ becomes comparable with $\bar{u}$. The third range (C) extends from this latter value to infinity.

In range A the intensity is proportional to density, and it would be an easy matter to determine density for various constituents if it were confined to this range. In range B the intensity depends very little on density, as it only enters as a small correcting term in the greater part of the range. In C, on the other hand, the intensity is proportional to the square root of the density, and the mutual relationship is again of a type allowing the density to be fairly well determined.

The first problem to solve is to find out in which range the densities of calcium and sodium are to be placed. This is easy as range B is distinguished from A and C by the fact that all individual peculiarities of the lines are blotted out. There can, for instance, no longer be any difference in intensity between the H and the K lines, and if both calcium and sodium are in this range their lines should be equally intense.

The obvious way of applying this test is to consider the relative intensity of the H and K lines in the same spectrum. In range A the relative intensity should be the ratio of the respective transition probabilities, and in range C the relative intensity should be the square root of this ratio. As far as observations go, however, there is no difference to be found between these lines, which at once indicates that the density of Ca^+ atoms is in range B.

Consider next the limits of range B. These are determined on one hand by $\tau_\nu = 2$ at the centre of the line, i.e. by (141)

$$n_A = \frac{2}{\pi} \frac{dv_0'}{ds} \frac{1}{K}$$

very nearly. The value of ds/dv_0', corresponding to $\sin 2\theta = 1$ and a radial velocity of 17 km. per kiloparsec, is

$$\frac{ds}{dv_0'} = 7{\cdot}10 . 10^{10} \text{ cm. sec.}$$

The value of the constant K is for the K line

$$K = \frac{e^2}{\mu c} f = 5{\cdot}54 . 10^{-3} \text{ cm.}^2 \text{ sec.}^{-1}$$

Introducing these values in the expression of n_A we find

$$n_A = 1{\cdot}6 . 10^{-9} \text{ cm.}^{-3}$$

Multiplying this number by the mass of a Ca atom, $6{\cdot}4 . 10^{-23}$ gm., the mass-density is found to be about 10^{-31} gm. cm.$^{-3}$

The upper limit to the density in range B, as we have defined it, is

$$n_B = n_A \tfrac{1}{8}\pi\bar{u}.$$

Expressing u in km./sec., we find, by using the value

$$\delta = 1{\cdot}15 . 10^{8} \text{ sec.}^{-1}$$

for the damping constant,

$$n_B = n_A 87u.$$

For a distance of one kiloparsec ($u = 17$ km./sec.), the upper limit to the density is thus about 1,500 times larger than the lower limit:

$$n_B = 2{\cdot}4 . 10^{-6} \text{ cm.}^{-3},$$

or 2·4 atoms per cubic metre.

Somewhere in this range the actual density of Ca^+ is to be expected.

For sodium the corresponding range in density is lower by a factor about 0·6. The sodium lines are about as intense as the calcium lines,† which indicates that sodium is also in range B.

Using the equations of dissociative equilibrium, modified to suit conditions in interstellar space, it is now possible to derive the corresponding ranges for the atoms in other stages of ionization as well as the ranges of total abundance.

It is then found that, due to the rather wide limits of range B, the data are consistent with the view that the relative abundance of Ca and Na is the same in interstellar space as in the stellar atmospheres. This result was obtained by Eddington.‡

Further information about the interstellar gas may be obtained from a detailed study of the form of the stationary lines in relation to galactic longitude and latitude.

130. Cosmic Dust and the Reddening Effect

While the atoms of calcium and sodium producing the stationary lines are undoubtedly in a gaseous state, there is definite evidence that a considerable amount of matter in interstellar space occurs in the form of fine dust. This is shown, first of all, by the presence of dark nebulae, or such nebulae as reflect the incident stellar light

† C. S. Beals, *Monthly Not.* **93** (1933), 585.
‡ Ibid. **95** (1934), 2.

without change of quality. But it is also shown by a reddening of the light coming from distant objects in or close to the Milky Way.

The information to be obtained about the state of this matter from observations of dark nebulae or reflection nebulae is rather vague. We can do little more than infer that the dust particles must have diameters comparable with the wave-length of visible light. The absence of any marked reddening of the light transmitted through dark nebulae indicates that the diameters of the particles are not much smaller than the wave-length of light, and considerations based on the estimated mass of the material indicates that they cannot be much larger either. Schalén† has tried to substantiate these conclusions more definitely by making extensive calculations based on Mie's‡ theory of absorption and scattering of light by small metallic droplets.

The reddening effect was first definitely discovered by Trumpler§ in a study of open galactic clusters. From a statistical study of the angular diameters of such clusters in relation to the apparent magnitudes and types of their brightest stars Trumpler concluded that a space absorption corresponding to 0·67 m. per kiloparsec in the photographic region and 0·32 m. in the visual region would be consistent with the observed facts.

Later work has served to confirm Trumpler's conclusions. The most direct confirmation came from a study of the colour indices of globular clusters in low galactic latitudes which was made by Stebbins,‖ and by Emma T. R. Williams and A. N. Vyssotsky.†† In general, the closer the clusters are to the galactic plane, the redder they are. This indicates definitely that the light suffers a marked selective absorption in the galaxy, and that this absorption is produced by particles of smaller size than those collected in dark clouds.

Although this discovery is of rather recent origin, there is already collected a large amount of material bearing upon the problem. The general trend of the material is to confirm the existence of the obscuring matter, and to show that it is very irregularly distributed in the galactic plane. In some places, as for instance in the direction of the Scutum star cloud, regions free from obscuring matter appear to exist, extending right into the centre of the galaxy itself. But apart from such isolated regions some absorption and reddening of light

† *Med. Upsala Obs.* No. 58 (1934) and No. 64 (1936).
‡ *Ann. d. Phys.* **25** (1908), 377.
§ *Lick Obs. Bull.* **14** (1930), 154.
‖ *Proc. Nat. Acad.* **19** (1933), 222.
†† *Astrophys. J.* **77** (1933), 301.

coming from distant parts of the galaxy must always be reckoned with.

While the total optical thickness of the absorbing cloud of dust is clearly very large in the galactic plane, it is quite moderate in a direction perpendicular to this plane. Thus Hubble,† from his extensive counts of extra-galactic nebulae, estimated that the total optical thickness in a direction from pole to pole of the galaxy is of the order of 0·5 m. Other investigators find a somewhat larger value of the obscuration. The corresponding differential absorption, i.e. the reddening of light passing perpendicularly through the galaxy, is, according to Stebbins, 0·18 m., while Williams and Vyssotsky find the smaller value 0·13 m. It is obvious that the absorbing layer of dust forms a very flat disk in the galactic plane, much more flattened, in fact, than the galaxy itself. This result recalls the well-known fact that the equatorial dark markings in extra-galactic nebulae seen edge-wise are narrower than the nebulae themselves.

An intensive study of interstellar absorption is a necessary feature of stellar spectroscopy. For unless the reddening effect is taken properly into account, temperatures derived from colour indices may be very misleading for stars even at quite moderate distances. In fact, in some cases when the colour temperature appeared absurdly low, as for ζ Persei for example, a selective interstellar absorption was invoked as a probable cause even before Trumpler's work. But apart from this side of the problem, the study of cosmic dust appears to be of high theoretical importance for the study of the evolution of individual stars as well as for the galaxy itself, and by analogy, for all other galaxies.

† *Astrophys. J.* **79** (1934), 8.

APPENDIX

We give here some data which are useful in the application of the dissociation theory to stellar problems. These data are mostly taken from Jevons's *Report on Band Spectra of Diatomic Molecules*, with some additions collected by Russell (*Astrophys. J.* **79** (1934), 322). The relevant quantities are: 1. The fundamental vibration frequency w, given in wave number units. 2. The mean separation between the atoms, r_0, in angstroms. 3. The dissociation energy χ, in volts. 4. The ratio of the weights of the ground states involved, $G = G_A G_B / G_{AB}$. 5. The reduced molecular mass,

$$M = M_A M_B / (M_A + M_B).$$

The figures in ( ) are uncertain; those in [ ] are still more so.

TABLE 14

Molecular Constants

Mol.	w	r_0	χ	G	M	*Mol.*	w	r_0	χ	G	M
H_2	4,371	0·75	4·43	4	0·50	CO	2,167	1·15	(10·0)	81	6·86
BH	(2,230)	1·23	[3·45]	12	0·92	NO	1,906	1·15	[6·1]	6	7·46
CH	2,851	1·13	(4·0)	3	0·92	AlO	977	1·61	(4·1)	27	10·1
NH	(3,000)	1·08	[4·4]	8/3	0·93	SiO	1,240	1·50	(7·9)	81	10·2
OH	3,568	0·97	[5·0]	3	0·94	TiO	1,008	1·62	(6·74)	21	12·0
MgH	1,493	1·73	[1·41]	1	0·96	ZrO	937	(1·8)	(7·5)	21	13·6
AlH	1,681	1·64	3·06	12	0·96	C_2	1,641	1·31	[5·5]	9	6·00
SiH	(2,012)	1·53	(3·0)	3	0·97	CN	2,069	1·17	[7·1]	18	6·46
CaH	1,317	2·02	(1·9)	1	0·98	N_2	2,360	1·09	7·90	16	7·00
CuH	1,946	1·46	3·0	4	0·98	SiF	865	2·26	(5·0)	9	11·4
ZnH	1,552	1·61	0·91	1	0·99	SiN	1,152	1·53	(6·2)	18	9·36
O_2	1,556	1·20	5·09	27	8·00	Na_2	159	3·07	[0·76]	4	11·5
BO	1,885	1·20	(6·6)	27	6·50	HF	4,037	0·86	6·4	12	0·95

In the next table we give a survey of the ground configuration of astrophysically interesting atoms in their electrically neutral and singly ionized states. The ionization energy is expressed in volts, and the series relationship of the electron last bound is given in the simplest form, without indicating the state of other electrons. (From Bacher and Goudsmit, *Atomic Energy States*.)

TABLE 15

Ground Configurations and Ionization Potentials of the Atoms

El.	χ	*El.*	χ	*El.*	χ	*El.*	χ
1. H $1s$	13·53	20. Ca $4s$	6·09	1.		20. Ca^+ $4s$	11·82
2. He $1s$	24·47	21. Sc $4s$	6·7	2. He^+ $1s$	54·14	21. Sc^+ $4s$	12·80
3. Li $2s$	5·37	22. Ti $4s$	6·81	3. Li^+ $1s$	75·28	22. Ti^+ $4s$	13·60
4. Be $2s$	9·28	23. V $4s$	6·76	4. Be^+ $2s$	18·12	23. V^+ $3d$	14·1
5. B $2p$	8·28	24. Cr $4s$	6·74	5. B^+ $2s$	25·00	24. Cr^+ $3d$	16·6
6. C $2p$	11·22	25. Mn $4s$	7·41	6. C^+ $2p$	24·27	25. Mn^+ $4s$	15·70
7. N $2p$	14·48	26. Fe $4s$	7·83	7. N^+ $2p$	29·47	26. Fe^+ $4s$	16·5
8. O $2p$	13·55	27. Co $4s$	8·5	8. O^+ $2p$	34·93	27. Co^+ $3d$	17·3
9. F $2p$	18·6	28. Ni $4s$	7·61	9. F^+ $2p$	34·6	28. Ni^+ $3d$	18·2
10. Ne $2p$	21·47	29. Cu $4s$	7·68	10. Ne^+ $2p$	40·89	29. Cu^+ $3d$	20·2
11. Na $3s$	5·12	30. Zn $4s$	9·36	11. Na^+ $2p$	47·02	30. Zn^+ $4s$	17·89
12. Mg $3s$	7·61	31. Ga $4p$	5·97	12. Mg^+ $3s$	14·96	31. Ga^+ $4s$	20·43
13. Al $3p$	5·96	32. Ge $4p$	8·09	13. Al^+ $3s$	18·74	32. Ge^+ $4p$	15·9
14. Si $3p$	8·12	33. As $4p$	10	14. Si^+ $3p$	16·27		
15. P $3p$	10·5	34. Se $4p$	9·5	15. P^+ $3p$	19·8		
16. S $3p$	10·3	35. Br $4p$	11·8	16. S^+ $3p$	23·3		
17. Cl $3p$	12·96	36. Kr $4p$	13·94	17. Cl^+ $3p$	23·9		
18. A $3p$	15·69	37. Rb $5s$	41·6	18. A^+ $3p$	27·72		
19. K $4s$	4·32	38. Sr $5s$	5·67	19. K^+ $3p$	31·7		

TABLE 16

Universal Constants

Velocity of light	c	$= 2{\cdot}99796.10^{10}$ cm. sec.$^{-1}$
Electronic charge	e	$= 4{\cdot}770.10^{-10}$ e.s.u.
Electronic mass	μ	$= 9{\cdot}035.10^{-28}$ gm.
Mass of a hydrogen atom	H	$= 1{\cdot}662.10^{-24}$ gm.
Planck's constant	h	$= 6{\cdot}547.10^{-27}$ erg. sec.
Boltzmann's constant	k	$= 1{\cdot}3709.10^{-16}$ erg. deg.$^{-1}$
Constant of gravitation	G	$= 6{\cdot}66.10^{-8}$ cm.^{3} gm.$^{-1}$ sec.$^{-2}$
Avogadro's number	N	$= 6{\cdot}064.10^{23}$ mole^{-1}
Radiation density at 1° K.	a	$= 7{\cdot}64.10^{-15}$ gm. cm.$^{-1}$ sec.$^{-2}$ deg.$^{-4}$
Rydberg constant for infinite nuclear mass	R_∞	$= 109737{\cdot}42$ cm.$^{-1}$
Constant in Wien's displacement law $\lambda_m T = b$	b	$= 0{\cdot}289$ cm. deg.
One 'volt' in ergs.		$= 1{\cdot}59.10^{-12}$ erg.

TABLE 17

Astronomical Data

The sun:	
Mass	$1{\cdot}985.10^{33}$ gm.
Radius	$6{\cdot}951.10^{10}$ cm.
Mean density	$1{\cdot}4109$ gm. cm.$^{-3}$
Surface gravity	$2{\cdot}736.10^4$ cm. sec.$^{-2}$
Luminosity	$3{\cdot}780.10^{33}$ erg. sec.$^{-1}$
Surface brightness	$2{\cdot}08$ erg. sec.$^{-1}$ cm.$^{-2}$
Absolute bolometric magnitude	4·85
Mean distance from the earth	$1{\cdot}494.10^{13}$ cm.
Number of seconds in one year	$3{\cdot}156.10^7$ sec.
One light year	$9{\cdot}461.10^{17}$ cm.
One parsec = 3·26 light years	$3{\cdot}08.10^{18}$ cm.

INDEX

Abel, solution of an integral equation, 205, 263.
Abney, on the rotation of stars, 202.
Absorption, classical theory, 94; quantum theory, 99; line profiles, 120, 134 et seq.
Abundance of elements, in the sun, 169, 183; in the stars, 182, 183, 251.
Adams, calibration of Rowland intensities, 167; determination of the solar rotation, 202; planetary spectroscopy, 259.
Adel, identification of planetary bands, 260.
Alkali spectra, 57.
Ambarzumian, ionization in nebulae, 318.
Angular momentum, 44.
Antisymmetry of wave functions, 60.
Auroral transitions, 303.
d'Azambuja, on the profiles of magnesium *b*-lines, 138.
Azimuthal quantum number, 36.

Bahr, vibration bands in H_2O, 225.
Baldet, on the spectra of comets, 244.
Balmer's formula, 44, 45.
Bands, vibration, 224; electronic, 225; branches, 227; multiplet structure, 228; in stellar spectra, 230.
Bartlett, theory of forbidden transitions, 306.
Baxandall, identification of AlO in the spectrum of Mira, 230; discovery of CH in stars, 231.
Beals, expansion of the atmospheres of *WR* stars and novae, 291; diagram of *WR* profiles, 291; intensity of interstellar lines, 343.
Becker, intensity of the corona, 264
Beer, the eclipse of ζ Aurigae, 285; forbidden O I lines in Nova Herculis, 313.
Bergstrand, intensity of the corona, 265.
Bethe, 192.
Birge, interpretation of TiO bands in stellar spectra, 230; discovery of carbon isotopes, 232; solar temperature from band intensities, 235.
Birkeland, corpuscular theory of aurora borealis, 268.
Blaton, theory of quadripole radiation, 306.
Bobrownikoff, spectra of Halley's comet, 244.
Bohr, theory of the Rydberg constant, 45; ionization by swift particles passing through matter, 320.
Bok, spectrum of η Carinae, 315.
Boltzmann, molecular statistics, 10; the gas constant, 12; radiation law, 108.
Born, intensity of Zeeman components, 75.
Bowen, spin doublets, 58; identification of nebular lines, 311; excitation of permitted lines of O III and N III in nebulae, 333, 334; abundance of elements in nebulae, 336.
Boyce, metastable states, 303, 314.
Brackett, infra-red hydrogen series, 186, 188.
Breit, dispersion theory, 152.

Cambresier, theory of stellar bands, 245, 248, 251.
Campbell, spectrum of comet Daniel, 244.
Canonical equations of motion, 5; distribution of Gibbs, 12, 23.
Carroll, rotational deformation of line profiles, 206, 212; theory of a spherically expanding star, 292.
Cauchy's theorem, 16, 85, 89.
Central fields, in atoms, 35.
Chandrasekhar, influence of opacity on line intensities, 194, 200; chromospheric theory, 269, 274; profiles of emission lines originating in an expanding envelope, 292, 296, 297; three-state cycle theory of a hydrogen envelope, 328.
Chapman, corpuscular theory of aurorae and magnetic storms, 268.
Characteristic differentials, 191.
Chartier, convergence test, 85.
Christie, eclipse of ζ Aurigae, 285.
Christy, identification of TiO bands in stellar spectra, 230; molecular dissociation in stars, 245.
Cillié, intensities of Balmer lines in nebulae, 331.
Comets, 244.
Condon, forbidden lines and the theory of magnetic dipole radiation, 304 et seq.

Cowling, influence of magnetic forces on the distribution of matter in the chromosphere, 279.
Curtiss, frequency of bright-line *B* stars, 286.
Cyclic transitions, 122, 140; in dilute temperature radiation, 145; in nebulae, 324, 325.
Czerny, vibration bands in HCl, 225.

Damping constant of spectral lines, 86.
Darwin, on the use of Cauchy's theorem in statistical mechanics, 16; influence of the electronic spin on the hydrogen terms, 55; intensity of Zeeman components, 75.
Davidson, chromospheric spectra, 262.
Debye, interaction between charged particles in a gas, 156.
Decay, time of, 93, 95.
Degenerate systems, 22.
Dirac, quantum statistics, 13; theory of quantum mechanics, 19; the relativistic wave equation, 58; wave-mechanical formulation of Pauli's principle, 60; dissociative equilibrium in the presence of a temperature gradient, 158.
Dirichlet's integral theorem, 85.
Dispersion, electrons, 95, 99; theory, 89.
Displaced terms, 74.
Dissociation theory, 13; of stellar compounds, 245.
Doppler effect, 90, 100, 163.
Dorgelo, sum rule, 72.
Dunham, discovery of CO_2 in the spectrum of Venus, 259; on the identification of the previously unidentified bands in the spectra of the major planets, 260.

Eddington, theory of the fine structure constant, 53; approximate solution of the equation of transfer, 112; theory of mixed scattering and absorption, 126, 129, 165; origin of long period variability, 223; ionization in nebulae, 316; interstellar absorption lines, 339, 343.
Effective quantum number, 57.
Effective temperature, 114.
Einstein coefficients, 23, 29, 85, 144.
Electric dipole, 30; moment, 30, 40.
Electron, spin, 53; density in stars, 180; density in nebulae, 319.
Ellerman, spectra of *N* stars, 243.
Elvey, deformation of lines in the spectrum of eclipsing binaries, 210, 211; rotational speed of stars, 212, 213, 214.
Emission lines, theory of, 83; in stellar spectra, 286.
Emissivity, definition, 106.
Epstein, quantum theory of the Stark effect, 220.

Fabry, effective temperature of the sun, 118; temperature of interstellar gases, 321.
Fermi, quantum statistics, 13.
Ferraro, influence of magnetic forces on the equilibrium of the chromosphere, 281, 282.
Ferrers's associated Legendre polynomials, 37.
Flux of radiation, 106, 113.
Forbidden lines, in nebulae, 311; in novae 313; excitation of 335.
Fowler, A., spectra of stripped atoms, 311.
Fowler, R. H., formulation of statistical mechanics, 16; intensity maxima of spectral lines as a function of effective temperature, 175; the rate of fading-out of the absorption lines, 179.
Fraunhofer, solar absorption lines, 125.
Fues, normalization of continuous wave functions, 190.

Galactic rotation, 338.
Gaunt, theory of the photoelectric effect, 192.
Genard, profiles of emission lines in the spectra of novae, 292, 298, 299, 300.
Generalized momenta, 4.
Gerasimovic, on ultra-violet emission in stars, 189; influence of light pressure on the equilibrium of envelopes in *Be* stars; hypothesis of expansion of *Be* envelopes, 290, 292.
Giauque, discovery of the oxygen isotopes in earth's atmosphere, 232.
Gibbs, statistical theory, 10, 13.
Gordon, theory of the photoelectric effect, 192.
Goudsmit, hypothesis of the spinning electron, 54.
Groups, electronic, in the atoms, 61 et seq.

Hale, discovery of solar magnetism, 218, 279; observation of the spectra of *N* stars, 243.

Halley's comet, 243, 244.
Hamilton, equations of motion, 5; variation principle, 3; integration theory, 5; partial differential equation, 6, 8, 9; function, 6, 8, 9; operator, 19, 20.
Harmonic analysis of line profiles for the detection of rotation, 207.
Hartmann, discovery of interstellar lines, 337.
Hartree, invention of the method of the self-consistent field, 35.
Heaviside layer in the atmosphere, 184, 185.
Heger, discovery of interstellar sodium lines, 337.
Heisenberg, separation of spin doublets, 55; incorporation of Pauli's principle in wave mechanics, 60; the intensity of Zeeman components, 75.
Heitler, theory of the reaction isochore, 248.
Higgs, the electron density in stellar atmospheres, 181.
Hogg, profiles of Ca^+ lines in stellar spectra, 181.
Holtsmark, theory of pressure broadening, 221.
Hönl, relative intensities of multiplet lines, 76, 77.
Hückel, interaction between charged particles in a gas, 156.
Huggins, absorption bands of ozone, 258; discovery of bands in the spectrum of Jupiter, 259.
Hutchisson, on the theory of molecular vibrations, 238, 242.

Ingram, on the analysis of Mg^+ 4,481 in the spectrum of Algol, and other stars, 211, 213.
Innes, variation of η Carinae, 314.
Integral equation, for the total intensity of J, 130; for the intensity of a rotationally deformed line profile 207; for the radiation from an envelope, 263.
Intensity, total for coincident lines in degenerate systems, 42; of Fraunhofer lines, 160; of radiation, 106; of arc lines, 175; of spark lines, 177; maxima of lines, 175.

Jackson, solution of the equation of transfer by integral equations, 130.
Jacobi, integration theory of the equations of motion, 1.
Jeans, on the equilibrium of planetary nebulae, 289.
Johnson, M. C., on the emission of hydrogen and helium from a star by radiation pressure, 291.
Johnston, H. L., discovery of the oxygen isotopes, 232.
Jordan, on the spin separation of doublets, 55; on the intensity of Zeeman components, 75.
Joy, the influence of band absorption on the variability of Mira stars, 223; on the spectrum of Mira, 230.

Kaplan, on the origin of the auroral spectrum, 258.
Kennelly, the Heaviside-Kennelly layer, 184, 185.
Kienle, on the continuous radiation from the stars, 189.
Kiepenheuer, on the theory of the Corona, 282, 283.
Kinetic potential, 2.
King, A., discovery of the carbon isotopes, 232; derivation of the solar temperature from the intensity distribution in bands, 235.
King, L. V., solution of the transfer equation by the use of integral equations, 130.
Kirchhoff, radiation law, 109.
Klein, wave-mechanical theory of the Zeeman effect, 217.
v. Klüber, on the profiles of Fraunhofer lines, 134.
Kramers, theory of series-limit absorption, 194.
Kratzer, detection of isotopes from the analysis of band spectra, 232.
Kronig, theory of multiplet intensities, 77; theory of multiplet structure in bands, 228.

Ladenburg, test of the influence of thermal Doppler effect on line width, 102.
Lagrange, equations of motion, 2.
Laguerre polynomials, 51.
Landé, spin separation of alkali doublets, 58; on the intensity of Zeeman components, 76.
Laplacean operator, in polar coordinates, 36; in cylindrical coordinates, 217.
Laporte, on multiplet spectra in *Handbuch der Astrophysik*, 72, 77.

Legendre polynomials, 37; associated, 38, 217.
Light pressure, defined by the radiation tensor, 107; isotropic light pressure, 108.
Limb darkening, coefficient of, defined, 119; negative coefficient of in the theory of *Be* envelopes, 289, 290.
Limb effect, progressive red shift of lines on passing from the centre to the limb of the sun, 268; in eclipsing binaries, meaning a progressive deformation of line profiles during the eclipse, 212.
Lindblad, on the theory of the continuous spectrum from a star, 117.
Line profiles, of a pure emission line, 86; of pure absorption lines, 100 et seq.; in the spectra of the sun and the stars, 134 et seq.; of *H* and *K* in stellar spectra, 181; in spectra of rotating stars, 202; profiles of lines originating in an expanding envelope, 291–300; of interstellar lines, 340, 341.
Liouville's theorem, 11.
Loomis, isotope effects in band spectra, 232.
Lorentz, theory of the Zeeman effect, 216, 218.
Lowen, quantum theory of dispersion, 153.

v. Maanen's white dwarf, 222.
Magnetic quantum number, definition, 36.
Mass formula of the relativity theory, 7.
Matrix, of perturbation, 81.
Maxwell, on molecular statistics, 10; velocity distribution, 13, 100; electromagnetic equations, 94.
McCrea, differential effects of opacity, 194; density of hydrogen in the chromosphere, 263, 264; theory of the limb effect, 268; on the electron density in the chromosphere, 273.
McLaughlin, rotational limb effect studied in Algol and β Lyrae, 212.
McLennan, identification of the green auroral line with a forbidden O I line, 258.
Mellin, inversion formulae for some integrals, 205.
Menzel, on metastable states of astrophysical significance, 303, 314.
Merrill, identification of ZrO bands in the spectra of *S* stars, 230; typical bright-line profiles in *Be* stars, 287; a case of periodic variations of a *Be* bright line, 288; discovery of forbidden lines in the spectrum of *R* Aquarii, 314; forbidden lines in H.D. 45,677, 315; discovery of new interstellar lines, 327.
Metastability, definition of, etc., 301, 302; astrophysical significance of, 303; theory of transition from metastable states, 304–7.
Mie, theory of light absorption by small metallic spheres, 344.
Millikan, spin separation of alkali doublets, 58; spectra of stripped atoms, 311.
Milne, theory of the continuous spectrum of the sun, 117; theory of combined absorption and scattering, 126; the method of maxima, 175, 181, 182; differential effects of opacity, 194, 200; a calcium chromosphere supported by light pressure, 269; expulsion of atoms by light pressure, 274.
Minnaert, theory of line profiles, 132; observation of line profiles, 134, 138, 151; theory and observations of total line intensities, 160, 164, 165, 166, 167; search for absorption edges in the sun, 188; eclipse observations, 262.
Mitchell, eclipse observations, 262.
Mitra, on the theory of the limb effect, 268.
Molecular compounds, theory of, 223–9, 234–41, 245–51; in stars, 230.
Momentum, four-dimensional, 9.
Moore, G. E., calibration of Rowland intensities, 167.
Moore, J. E., radial velocity of the corona, 267; width of Fraunhofer lines in the corona, 264.
Morgan, difference between giants and dwarfs, 196.
Mulders, profile of a Mg *b*-line, 138; total line intensities, 160, 164, 166, 167.
Multiplet spectra, nomenclature of, 70.

Nebula, the Orion, 307; the Trifid, 307, 319, 322.
Nebulae, planetary, emission, reflection, 308; ionization in, 316; temperature in, 320.

Newall, identification of the *G* band with CH, 231.
Newton, equations of motion, 1.
Nomenclature, spectroscopic, 62.
Normalization, of discrete wave functions, 25; of continuous wave functions, 190.
Novae, lines widened by Doppler effect, 288, 291; diagrams of line profiles in, 288, 289, 290.

Occultation effect, 290, 294.
Öhman, observation of a white dwarf, 222.
Oldenberg, on intensities in band spectra, 244.
Oort, on differential galactic rotation, 338.
Opacity, 184–9, 194–201.
Operators, 19.
Optical depth, definition, 111.
Ornstein, intensity rules, 76; on eclipse, observations, 262.
Orthogonality, 22.

Pannekoek, theory of line profiles, 132, 136, 147, 165; ionization theory, 158; damping constant from observations of α Cygni, 167; solar eclipse observations, 262.
Parkhurst, observation of *N* spectra, 243.
Partition function, 14, 155.
Paschen, series, 186, 188; spectra of strongly ionized atoms, 311.
Pauli, on the formulation of the Fermi-Dirac quantum statistics, 13; reference to an article in *Handbuch der Physik*, 32; exclusion principle, 59, 60.
Payne, profiles of *H* and *K*, 181; on the abundance of elements, 182; differential effects of opacity, 200; constitution of *WR* stars, 288; on forbidden lines, 303, 314.
Periodic system, theory of, 59; diagram, 65.
Pertner-Exner, *Meteorologische Optik*, 185.
Perturbation, theory, 26, 50; matrix, 82.
Phase space, 10.
Photoelectric effect, 189.
Piccardi, dissociation of molecules, 245.
Pickering, discovery of He II lines, 288.
Planck, density of thermal radiation, 24, 109, 117.
Plaskett, H. H., fine structure of Balmer lines, 46; the temperature of the sun, 118; observations of line profiles, 134, 139, 147.
Plaskett, J. S., on the rotation of 21 Cas, 212; fine structure of lines in the spectra of γ Cas, β Lyr, and v Sgr, 288.
Poynting's vector, 96, 99.
Probability coefficients, 11; in multiplets, 75; in hydrogen, 148.
Profile, *see* Line profiles.

Quadripole radiation, 305.
Quantum conditions, 21; numbers, 36, 56.

Radiation, dilute temperature, 145; transfer of, 105; tensor, 107; thermal, 24; field theory of, 73 et seq.
Rare earths, 68.
Rayleigh, radiation law, 115, 123; scattering, 185.
Reaction isochore, 248.
Recurrence formulae, of Legendre polynomials, 38; of Ferrers's associated Legendre polynomials, 38; of Hermite polynomials, 34.
Refraction index, theory of, 94.
Reiche, test of thermal Doppler effect, 102.
Resonance lines, theory of, 83; wave lengths of, 176; theory of the radiative transfer within, in a nebula, 328.
Richardson, discovery of AlH in sun spots, 321; solar temperature from band intensities, 235.
Righini, a study of a Mg *b*-line in the sun, 138.
Ritz, correction to Rydberg's formula, 49.
Rodrigues's formula, 37.
Rosenberg, on the cross-section for atomic collisions, 276.
Rosenfeld, dissociation of stellar compounds, 245, 248, 251.
Rosseland, on the effect of cyclic transitions on the residual line intensities, 152; on solar magnetic fields, 219; on the intensities of auroral bands, 244; on a chromosphere supported by corpuscular rays, 268, 275.
Rossiter, rotation effects in β Lyr and β Per, 212.
Rotating stars, 202, 289.
Rowland, intensities of Fraunhofer lines, 160, 167, 169.

Rubens, the spectrum of H_2O, 225.
Rubinowicz, theory of quadripole radiation, 306.
Rufus, intensity of the *G* band, 231.
Russell, nomenclature, 70; Russell-Saunders coupling, 72; multiplet intensities, 77; abundance of elements in the sun, 160, 167, 169, 182; opacity formulae, 194; state of ionization in stellar atmospheres, diagrams, 199, 201; state of dissociation, diagrams, 253, 254, 255, 256, calculations, 257; interpretation of interstellar lines, 337.
Rydberg, constant, 45; formula, 49.

Saha, theory of spectral classes, 175.
St. John, the limb effect, 268.
Sandford, discovery of the carbon isotopes, 232.
Saunders, the Russell-Saunders coupling, 70, 72.
Scattering, coherent, 122.
Schalén, theory of interstellar absorption, 344.
Schlesinger, discovery of stellar rotation, 211.
Schrödinger, wave equation, 19; perturbations in line series due to resonance, 53.
Schütz, test of line profiles by total absorption, 102.
Schwarzschild, quantum conditions, 21; on the use of integral equations in problems of radiative transfer, 130; on line profiles, 134; on the theory of the Stark effect, 220.
Screening, 64.
Selection principles, for the harmonic oscillator, 34; for central fields, 39, 40, 41, 42.
Self-adjoint operator, 20.
Self-consistent field, 35.
Shajn, on the rotation effect in eclipsing binaries, 210, 213.
Shane, intensity of the *G* band, 231.
Slater, theory of valency, 63.
Slipher, on comets, 244; on the interpretation of planetary bands, 260; on Hubble's variable nebula, 314.
Slob, on total line intensities, 160, 164, 166.
Sommerfeld, quantum conditions, 21; the fine-structure formula, 53; on multiplet intensities, 76, 77.
Spin, electronic, 53, 54.
Stark effect, 220.
Stebbins, on TiO and C_2 in α Herculis, 255; on interstellar absorption, 345.
Steensholt, on the relative intensities of bands, 244.
Stevenson, theory of forbidden transitions, 306.
Stobbe, on the photoelectric effect, 190, 192, 193.
Störmer, on the corpuscular theory of aurorae and magnetic storms, 268.
Stratton, on eclipse spectra, 262.
Strömgren, B., on line profiles, 152.
Struve, on the profiles of *H* and *K*, 181; on the rotation effect in eclipsing binaries, 210, 211, 212, 213, 215; on Stark effect, 221; on the abundance of H and N in late-type stars, 251; on 17 Lep, 284; rotation hypothesis of *Be* stars, 289.
Sum-rules, 42.
Swan bands, 231, 235, 242.
Swings, 230, 245; on the abundance of H and N in late-type stars, 251; on the rotation hypothesis for *Be* stars, 289, 295.

Temperature, effective, 114; in giants and dwarfs, 181; from intensity maxima of lines, 182; from bands, 234 et seq.
Thackeray, on line profiles, 151.
Thomas precession, 55.
Thomas-Kuhn sum rule, 39, 99.
Thomson, scattering formula, 100.
Transauroral transitions, 303.
Transfer, equation of radiative, 102.
Trumpler, discovery of interstellar absorption, 344.
Turner, intensity distribution in the corona, 264.

Uhlenbeck, discovery of the electronic spin, 54.
Unsöld, on line profiles, 134, 135, 136, 147, 151; on the Stark effect, 221.

Vector potential, 7.
Vegard, on upper air temperature, 236, 237; an auroral theory, 268.
Verwey, on Stark effect in white dwarfs, 222.
Vyssotsky, on interstellar absorption, 344.

Waller, on the Rydberg-Ritz correction, 51.

Wave equation, 18, 25.
Weisskopf and Wigner, probability distribution law, 87 et seq., 122, 141.
Westgate, rotation of the stars, 214.
Weyl, normalization of characteristic differentials, 190.
Wien, approximation, 116; displacement law, 118.
Wigner, see Weisskopf.
Wildt, dissociation theory, 245; discovery of methane and ammonia in the major planets, 260.
Wiles, analysis of line profiles, 140, 141.
Williams, E. T. R., on interstellar absorption, 344.
Wilson, on the eclipse of ζ Aurigae, 285.
Wolf, spectra of planetary nebulae, 310.
Wolf-Rayet stars, 288, 291, 310.
Woltjer, on an outward drift of Ca in the sun, 274, 275; on reabsorption in an expanding envelope, 331.
Woolley, effect of cyclical transitions on the residual line intensities, 140, 146, 147; on the equation of ionization, 158; on the effective temperature of the sun, 159; smoothness of the continuous spectra due to cyclic processes, 188.
Wurm, on an isotope effect in *N* stars, 233; on the temperature of late-type stars, 238, 242; on the temperature of comets, 244.

Yü, on 'ultra-violet appendages', 189.

Zanstra, on comets, 244; light pressure in planetary nebulae, 269, 331, 333; on reflection nebulae, 308; on the temperature of nebulae, 310.
Zeeman effect, 75, 216; in the sun, 218.
Zwaan, intensity of *H* and *K* lines, 273.

PRINTED IN
GREAT BRITAIN
AT THE
UNIVERSITY PRESS
OXFORD
BY
JOHN JOHNSON
PRINTER
TO THE
UNIVERSITY