UNIVERSAL
LIBRARY
OU_164193
UNIVERSAL
LIBRARY

OSMANIA UNIVERSITY LIBRARY

Call No. 530 E 965 Accession No. 25965

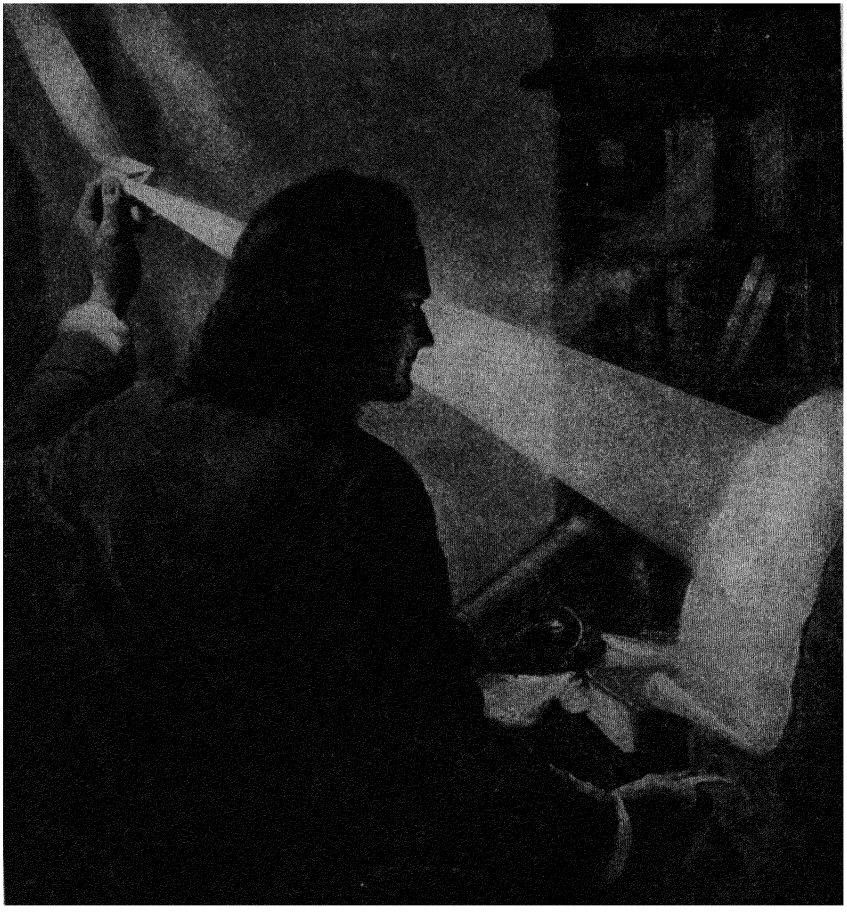
Author Eyring. C.F.

Title Survey Course in physics 1946

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

PRENTICE-HALL PHYSICS SERIES

A Survey Course in Physics



Copyrighted by Bausch & Lomb Optical Co.

PLATE I. Sir Isaac Newton (1642–1727) Separating White Light into Its Colors by Means of a Prism. Through his extensive experiments with light, Newton became one of the founders of modern optical science. (*Courtesy of Bausch and Lomb Optical Company, Rochester, N. Y.*)

A Survey Course

in

Physics

by

Carl F. Eyring, Ph.D.

Professor of Physics, and Dean of the
College of Arts and Sciences
Brigham Young University

NEW YORK
PRENTICE-HALL, INC.
1946

PRENTICE-HALL PHYSICS SERIES

COPYRIGHT, 1934, BY
CARL F. EYRING

COPYRIGHT, 1936, BY
PRENTICE-HALL, INC.
70 FIFTH AVENUE, NEW YORK

ALL RIGHTS RESERVED. NO PART OF THIS BOOK MAY BE REPRODUCED IN ANY FORM, BY MIMEOGRAPH OR ANY OTHER MEANS, WITHOUT PERMISSION IN WRITING FROM THE PUBLISHERS.

First Printing.....September, 1936
Second Printing.....October, 1937
Third Printing.....November, 1942
Fourth Printing.....January, 1945
Fifth Printing.....April, 1946
Sixth Printing.....July, 1946
Seventh Printing....September, 1946

PRINTED IN THE UNITED STATES OF AMERICA

Preface

THE orientation of a college student in the modern world cannot be considered complete without some profitable contact with science. Though probably not the most important, it is not too much to say that of all the sciences physics is the most fundamental. Yet with the crowded curriculum of the modern college, many students do not find time to elect a full year's course in this science. It seems, therefore that, a "Survey Course in Physics" extending through one quarter or one semester will meet a need felt especially by the non-science student.

This reconnoissance of the field of physics is expected in no way to replace the general course for which many splendid texts are written. It is to serve as an orientation course especially for those whose interests lie in the sphere of the humanities. A short course might skim through the field of physics, but this procedure could not possibly exemplify the true spirit of science. It has seemed best to limit the field, thus leaving many important topics untouched. The choice of subject matter, therefore, has involved the personal judgment of the author.

Eight years ago *A Survey Course in Physics* was first produced in trial form, and it has since been used as a text by more than a thousand students. Out of this experience has grown the conviction that the non-science student is intensely interested in the interpretation of his immediate physical environment. In the present revision, even more than formerly, the human body and its physical environment have been taken as the central theme. Thus, balance and force of gravity, force and change of motion, energy and hand-tools, heat in relation to body and household, the atmosphere of sounds, the world of light and

color, and electrical manifestations are the general topics considered.

In certain situations, the course as outlined may require a modification of emphasis. For example, if the majority of the class members are students of literature, speech, music, and art, parts of Chapters V, VI, and VII could be left out and more time be spent on the study of sound and color. On the other hand, if the majority of the class members are interested in the physics of the home, parts of Chapters VIII, IX, X, and XI could be omitted and emphasis be given to the chapters on heat and household physics. Thus, although not written as a textbook on sound and color, or as a household physics text, *A Survey Course in Physics*, under a proper selection and emphasis of subject matter, may serve these purposes in addition to that of orienting the non-science student.

The author is indebted to the scientists of the past and present for the fund of information from which he has freely drawn in building this text. He is grateful to the persons and institutions who have permitted the use of drawings and figures. It is hoped that due recognition and credit have been accorded those responsible for the material used. The author also expresses his gratitude to his colleagues for help and suggestions in the preparation of the manuscript; to Mr. Farrell Collett for the interest taken while executing the drawings; and to Professor Edward U. Condon, editor of the series of physics books published by Prentice-Hall, Inc., for his valuable suggestions and encouragement.

C. F. E.

Contents

PAGE

PREFACE v

ILLUSTRATIONS xi

CHAPTER

I. INTRODUCTION 1

Physics and human life. Some achievements of science. A deductive system of discovery. An inductive system of discovery—the "scientific" method: experimental facts; the hypothesis; a cyclic process; fact and hypothesis; instruments of science; the search for truth; co-operative science.

II. SIZE, WEIGHT, AND BALANCE 9

Units of length. Simple definition of force. Weight: the pound. The lever: the principle of the lever; three classes of levers. Center of gravity: position of the center of gravity. Parallel forces acting in a plane. Concurrent forces acting in a plane. Complete requirements of equilibrium for a balanced lever. Stability. Posture.

III. FORCE AND MOTION 25

Unit of time. Velocity. Acceleration. Falling bodies. Force and change of motion. Newton's first law of motion. Centrifugal and centripetal forces. The motion of projectiles. Inertia and mass: units of mass. Momentum. Newton's second law of motion. Gravitational attraction: falling bodies and air resistance. Newton's third law of motion.

IV. ENERGY AND WORK 50

Utilization of natural energy. The nature of energy. How is work measured?: a unit of work. The measure of how fast work is done—power. Work from moving bodies—the hammer. Kinetic energy. Conservation of energy. Potential energy. Machines: the law of machines; the simple lever; weighing machines; other more complex levers; the inclined plane. Friction: sliding friction; rolling friction; air resistance; friction reduction. The efficiency of a machine.

V. TOO SMALL TO BE SEEN 69

The atoms of which things are made. The motion of molecules: diffusion of gases; gas pressure; Boyle's law; causes of gas pressure; explanation of diffusion through porous cup. Molecular magnitudes. Liquids and solids.

CHAPTER PAGE

VI. THE NATURE OF HEAT 85

Is heat a fluid? Temperature and its measurement: simple air thermometer; mercury thermometer; thermostats; what is it that temperature measures? The measurement of heat: the calorie; specific heat. Heat and work: the equivalence of heat and work; Joule's experimental work; how may we get work from heat?; the human machine; the sun the source of energy.

VII. PHYSICS OF THE HOUSEHOLD. 106

Water supply. Water in three forms: ice; molecular forces in water; evaporation. Canning fruit. The heat of vaporization. The boiling point of solutions. Osmosis. Ventilation. Humidity. Heating the home: the transmission of heat; home heating devices; heat insulation.

VIII. MECHANICAL VIBRATIONS 138

The vibration of a mass attached to a spring: cause of the mechanical vibration; the conditions for simple harmonic motion; stiffness of a spring; energy of oscillations; amplitude of vibration; period and frequency of vibration; phase. Sympathetic vibrations—resonance: automobile on a washboard road. Forced vibrations. Transverse vibrations in rods and bars: factors controlling the period of an oscillating bar; a bar has many modes of vibration; the xylophone; the tuning fork; reeds. Transverse vibrations in plates. Transverse vibrations in strings: fundamental and overtones; complex vibrations; a traveling deformation. Longitudinal vibrations in rods. Vibrating air column.

IX. WAVES 158

What travels in a wave? Wave length: frequency of vibration, wave velocity, and wave length. Speed, reflection, and interference of waves on ropes: speed of waves on ropes; reflection of waves; interference of waves. Standing waves: standing waves in air columns. Waves in free air. Sound waves: sound, a wave motion in an elastic material medium; the intensity of sound; direction of sound travel; refraction of sound in the open air; reflection, transmission, and absorption of sound.

X. PITCH, LOUDNESS, AND TONE QUALITY. 175

Musical tone and noise. Pitch: pitch and frequency; standards of pitch. Loudness and intensity: limits of audition; intensity level and loudness level; a loudness scale. Limits of hearing: how many pure tones may be heard?; impaired hearing. The quality of a musical tone. Wave form. The human ear: structure; operation. Beats, consonance, and dissonance. Musical intervals. The musical scale.

CHAPTER PAGE

XI. MUSICAL INSTRUMENTS, SPEECH, AND AUDITORIUMS 205

Amplification: amplification by resonance; amplification by forced vibration. Stringed instruments: vibrating strings; the piano; the violin. Wind instruments: air columns; generator and amplifier of wind instruments; the pipe organ; the flute; the clarinet and oboe; brass instruments. The vibrating diaphragm: the telephone; the phonograph. The radio. The public address system. The human voice: the stimulator; the generator; the modifiers; speech power; frequency distribution of energy in speech; speech sounds; good singing voice quality; percentage articulation tests. Sound in auditoriums: speech in auditoriums; music studios and halls.

XII. ENERGY ON LIGHT WAVES 240

What is light?: source of light; intensity of a source of light; velocity of light. Diffraction. Reflection: reflected light; the law of reflection of light; regular reflection; diffuse reflection. The sky. Illumination: illumination and distance from the source; the foot-candle; brightness of a surface; foot-candles required for proper lighting; glare; shadows. Refraction.

XIII. COLOR 260

Colors in nature: the rainbow; colors of thin film; natural colors due to absorption and reflection. Dispersion. Mixing colored lights: complementary colors; mixing red, yellow, green, and blue light; Newton's color wheel; the law of mixing colored lights. Color vision: color zones of the retina; color blindness; Ladd-Franklin theory of color vision. Three characteristics of color: brightness; hue; saturation. Colors in art: hue, value, and intensity; mixing pigments; primaries, secondaries, and complements; color groups.

XIV. OPTICAL INSTRUMENTS. 284

The plane mirror. The convex lens: the focal length and principal focus; secondary foci; the nature of an image; series of images produced by the lens; relative size of object and image; numerical aperture of a lens. Defects of a convex lens: spherical aberration; distortion; astigmatism; chromatic aberration. The camera: operation; brightness of image; what controls the speed of a camera?; depth of focus; fixed focus; general suggestions. The eye: how the eye functions; defects of the eye; resolving power of the eye. Other instruments using lenses: the projection lantern; the simple magnifying glass; the compound microscope; the astronomical telescope.

XV. ELECTRICAL MANIFESTATIONS 305

Nature of an electric charge: two kinds of electricity; law of attraction and repulsion. Conductors and insulators: the

CHAPTER PAGE

condenser. How may we detect an electric current?: spark; heating effect; chemical effect; magnetic effect; right-hand rule. Magnets: electromagnets; permanent magnets; magnetic compass.

XVI. How Electricity Is Generated 322

Generation by intimate contact. Generation by induction: the electrophorus. Generation by chemical action: the dry cell; the lead storage battery. Generation by the action of light—photoelectric effect. Generation by the action of heat—thermoelectricity. Generation by cutting magnetic lines of force: the generator rule; source of the energy transferred to electricity in the process of electromagnetic induction; the generator; the telephone transmitter. Changing electrical energy into mechanical energy: the electric motor; the telephone receiver. Electrical measurements: the measure of the electric current—the ampere; the measure of the electrical potential—the volt; the rate at which electrical energy is used—the watt; electrical resistance—the ohm; Ohm's law. Electrical heating and lighting: the electric circuit; heating appliances; electric lamps.

XVII. Inside the Atom 343

Electric discharge in partial vacua. Cathode rays. Measuring the charge on the electron. Positive rays. X-rays: Coolidge tube; the nature of X-rays; crystal structure; general X-radiation; characteristic X-radiation; wave lengths of X-rays. Radioactivity: artificial disintegration of atoms; the structure of the nucleus. Atomic structure. Theory of magnetism. Electromagnetic waves. Cosmic rays. Corpuscles and waves: the Compton effect; the diffraction of electrons; uncertainty principle.

Appendix 367

Problem solving; completing sentences; procedure commands.

Index 369

Illustrations

Figures

PAGE

1. Growth of the Scientific Law 5
2. Measuring Length. 9
3. Lifting a Stone. 11
4. Spring Balance 11
5. Using a Lever 14
6. A Simple Lever 14
7. Three Classes of Levers 16
8. The Center of Gravity of a Meter Stick 18
9. Locating the Center of Gravity 18
10. Concurrent Forces 20
11. Types of Equilibrium. 21
12. Increasing Stability 22
13. Posture 23
14. A Stop Watch 26
15. Changes of Motion 26
16. A Falling Apple and Leaf 27
17. The Leaning Tower of Pisa 29
18. A Feather and a Penny in a Near-Vacuum . . . 30
19. The Pail Contains Water 34
20. Curves and Turns. 35
21. The Motion of Projectiles 37
22. Putting the Shot 39
23. Momentum 42
24. Changing the Momentum of a "Wagon" 43
25. Weight Due to Surface Gravity 46
26. Starting from the Mark 48
27. Doing Work 51
28. Work at Different Rates 53
29. Driving and Pulling Stakes 56
30. Potential Energy Changing to Kinetic Energy . . 58
31. The Lever 59

PAGE
32. The Windlass 61
33. Pulley Systems 62
34. The Inclined Plane 64
35. Reducing Friction. 66
36. Multiple Proportions 70
37. Friction and Motion of Molecules 71
38. Diffusion 72
39. Diffusion of Gases Through a Porous Cup. . . . 72
40. Air Pressure 73
41. A Mercury Barometer 73
42. An Aneroid Barometer 74
43. Illustrating Change of Pressure with Change of Volume 77
44. Diffusion in Liquids 82
45. Snow Crystals 83
46. Fire by Friction 86
47. Simple Air Thermometer. 88
48. Calibrating a Thermometer 89
49. Fahrenheit and Centigrade Scales Compared . . . 90
50. Thermostat Used in Homes 90
51. A Gas Thermometer 91
52. Kelvin and Centigrade Scales Compared 94
53. Comparing the Specific Heats of Water and Iron. . 97
54. Mechanical Energy Changed to Heat Energy . . 98
55. Joule's Apparatus 98
56. Food, the Source of Bodily Energy 100
57. Respiration Calorimeter 101
58. The Sun, Our Source of Energy 103
59. Home Water Supply 108
60. Expansion Due to Freezing 109
61. A Refrigerator 110
62. Regelation 110
63. Freezing Ice Cream 111
64. Water Surface Acting Like a Stretched Membrane . 111
65. Molecules at the Surface Are Attracted Toward the Interior 112
66. Capillary Action 113
67. Water Rising in Capillary Tubes 114
68. Drying by Evaporation 115
69. Illustrating the Process of Evaporation 116

PAGE
70. Evaporation and Saturation 116
71. Vapor Pressure 117
72. Boiling at Atmospheric and Reduced Pressures . . 119
73. Vapor at Surface Has the Liquid Temperature . . 120
74. Canning Fruit 121
75. Evaporation from Dilute, Condensation to Concentrated, Solution 124
76. Osmosis. 125
77. Ventilation 127
78. Hygrometer, Wet and Dry Bulb 129
79. Hygrometer, Hair Type 130
80. Illustrating Liquid Convection 131
81. The First Chimney 131
82. Hot Air Heating System 132
83. Hot Water Tank 132
84. Insulating Against Heat and Cold 135
85. Vibrating Spring 138
86. Amplitude of Vibration 141
87. Damped Oscillation 142
88. Out of Phase 180°. 144
89. Ready for a Dive 144
90. Illustrating Resonance 145
91. Resonance 146
92. Vibrating Bars. 148
93. Vibration Patterns on Bars 149
94. Tuning Fork 151
95. Chladni's Figures 152
96. Vibrating String 152
97. Comparing Time of Wave Travel with Time of Vibration 154
98. Vibrating Rods 155
99. Kundt's Apparatus 156
100. A Transverse Wave 159
101. Speed of Waves on Ropes 161
102. Reflection at Fixed End 162
103. Reflection at Free End 162
104. Destructive Interference 163
105. Interference Producing a Standing Wave 164
106. Standing Waves in Air Columns 167
107. Illustrating Waves in Air. 169

PAGE
108. Bell in Vacuum 171
109. Refraction of Sound 173
110. The Siren 176
111. Illustrating the Auditory Sensation Area 179
112a. Equal Loudness Level Curves 182
112b. Loudness Level of Common Sounds 183
113. A Loudness Scale for Pure Tones 184
114. Testing the Hearing of School Children 187
115. Demonstrating the Origin of Tone Quality 189
116. Sound Spectra of Tones 190
117. Wave Forms of Pure Tones 191
118. Compounding to Form a Complex Wave 191
119. Wave Forms of Musical Instruments 192
120. The Ear 193
121. Diagrammatic View of the Ear 194
122a. Pitch Variation Along Basilar Membrane 195
122b. Pitch Change Due to Loudness 195
123. Illustrating the Cause of Beats 200
124. Dissonance Between *Do* and the Other Notes of the Scale 202
125. Resonant Air Column 206
126. Strings and Soundboard of a Piano 208
127. Bowing the Strings of a Cello 210
128. Illustrating Vibrating Jets, Lips, and Reeds . . . 212
129. Pipes of an Organ. 212
130. The Horn 215
131. Hill-and-Dale Phonograph Reproducer. 217
132. The Loud-Speaker. 218
133. The Frequency Ranges of Instruments. 219
134. The Human Voice 220
135. Energy Distribution in Vowel Sounds 223
136. Oral Cavities for Certain Speech Sounds 226
137. Frequency and Intensity Level Ranges for Vowels and Consonants 228
138. Reflection of Sound 233
139. Auditorium with Good Design 234
140. Auditorium with Poor Design 234
141. Insulating Against Noise 236
142. Optimal Reverberation Times 237
143. A Candle 243

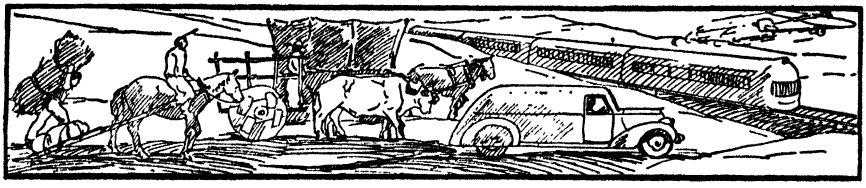
PAGE
144. Michelson's Velocity of Light Apparatus 244
145. The Blurring Effect of a Very Narrow Slit . . . 246
146. The Reflection of Light 247
147. Diffuse Reflection 248
148. Illumination Varying with Distance from Source . 250
149. The Foot-Candle 250
150. A Foot-Candle Meter. 253
151. Glare 254
152. Refraction 256
153. Refraction Through a Prism and a Lens 257
154. The Cause of Refraction; the "Broken" Oar . . . 258
155. Producing a Rainbow. 260
156. Interference Fringes 262
157. Diffraction Grating; Spectrum of Bronze 263
158. Selective Absorption and Reflection 266
159. Dispersion by a Prism 266
160. Recombining Spectrum Colors to Form White Light. 268
161. Mixing Colored Lights 272
162. Newton's Color Wheel 272
163. Color Zones of the Retina 273
164. Representing the Ladd-Franklin Color Theory . . 274
165. The Color Pyramid 277
166. Color Chart . . . . . *facing page* 280
167. Mixing Pigments Increases Selective Absorption . . 281
168. The Color Circle 281
169. The Plane Mirror 284
170. Images Formed by a Convex Lens 286
171. Spherical Aberration 287
172. Distortion 288
173. Astigmatism 289
174. Achromatic Lens 289
175. Camera; Depth of Focus Increases with Small Stop . 291
176. The Eye 295
177. Lines to Test Astigmatism 296
178. Eye Defects 296
179. Projection Lantern 297
180. Sound Picture Film (Variable Density) 299
181. Simple Magnifying Glass. 300
182. Simple Compound Microscope 301
183. Telescope 302

PAGE
184. Pith Ball Electroscope 306
185. Gold Leaf Electroscope 307
186. Conductors and Insulators on a High Tension Line . 309
187. The Leyden Jar 310
188. The Terminals of a Battery Are Electrically Charged 311
189. Electrolysis of Water 314
190. Illustrating the Right-Hand Rule 316
191. The Electromagnet and Its Application 317
192. The Compass Needle 317
193. The Magnetic Field of a Bar Magnet 318
194. The Earth, a Great Magnet 320
195. Charging by Induction 323
196. The Electrophorus 324
197. A Galvanometer 325
198. A Simple Electric Cell 326
199. The Action of a Simple Cell 326
200. Testing a Storage Battery 328
201. A Photoelectric Cell 330
202. Thermoelectric Effect. 330
203. Electromagnetic Induction 331
204. Illustrating an Alternating Electric Current . . . 333
205. Moving-Coil Transmitter; Carbon Microphone . . 333
206. Motor Effect 334
207. Moving-Coil Head Receiver 335
208. Measurement of Amperes and Volts 336
209. A Fuse 339
210. Three "Heats" in an Electric Range 340
211. Bending Cathode Rays with a Magnetic Field . . 344
212. Oil Drop Apparatus 345
213. X-ray Picture 347
214. Coolidge X-ray Tube 348
215. Salt Crystal 349
216. K-Series 352
217. Collision of Alpha Particles with Oxygen Atoms . . 355
218. The Hydrogen Atom 361

Plates

I. Sir Isaac Newton Separating White Light into Its Colors . . . . . *frontispiece*

PAGE
II. Stratosphere Balloon *Explorer II* 75
III. The World's Greatest Dam 107
IV. An Oscillograph Record of a Sentence 224
V. The Disc Which Will Serve as the Concave Mirror of the 200-Inch Telescope. 303
VI. An Electric Discharge Produced by 800,000 Volts . 337



CHAPTER I

Introduction

Physics and Human Life

We live in the perennial presence of the force of gravity and the weight which it produces; we feel the warming influence of heat, sense the luminosity of light, bathe in an atmosphere of sounds; in fact, our senses are continuously besieged with outside stimuli. Through nerves terminating in the surface of the body and in special sense organs such as the eye and ear, these "messages of the outer world" are laid at the threshold of consciousness for interpretation. We distinguish ourselves from other persons, and persons from other things. We recognize the difference between the physical and the social environment. In the physical environment we list animals; plants; inanimate things such as automobiles, houses, and mountains; and forms of energy such as sound, heat, light, and electricity. We soon come to know that the rules of the physical environment also apply to the human body. Gravity acts on it as truly as it does on a stone; energy in the form of food must be fed to it as regularly as gasoline is furnished for an automobile. Thus, although the science of man must be much broader than the physics of the human body, yet a study of man's physical environment is a real and important part of the study of the whole man.

Some Achievements of Science

From the stone age to our present age of automatic machinery the race has struggled to find the means of using effectively its own physical energy and the energy derived

from the forces of nature. With the development of science has come an ever-increasing understanding of the forces of nature. One by one the hiding places of pent-up energy have been discovered, and their stores have been made available and safe for the use of mankind. Machines have evolved from crude hand tools of stone to more perfect tools of bronze, from bronze tools to tools wrought in iron, from hand tools of iron and steel to machine tools of our present generation.

The ancients and the men of the Middle Ages used chiefly the energy of men and animals, of winds, and to some extent of running water. They knew very little of the energy which has since been made available from heat and chemicals, and nothing at all of energy in the form of electricity. To them, power transmission meant treadmills, the moving of herds and slaves, or marching armies. The slave was the machine for all heavy labor. He was never a particularly cheap machine and furnished only a small amount of energy singly, but in mass he was powerful. He cultivated fields, constructed dams, and dug ditches. As a galley slave he propelled the so-called "ships" of war. By his toil, walls and towers, roads and bridges, temples and pyramids, were built. Urged by task-masters to give in muscular exertion all the energy available, these slaves, working side by side in co-operative effort, produced results that were remarkable, and in the case of the Pyramids, stupendous.

But men were not always to toil as slaves. The controlled accumulation, storage, and liberation of energy—the manufacture of power—pointed the way to the emancipation of man from the bondage of physical toil. Even in ancient times the manufacture of power from gravitational sources by the use of mechanical contrivances, such as sails, windmills, and waterwheels, enslaved the forces of nature and elevated a few men, at least, to the position of master. The intelligent task-master became the director of energy and ceased to be the driver of slaves.

Of man's material wealth, energy in its various forms has turned out to be the most real and enduring element. A knowledge of the laws of natural forces, an understanding of how these forces operate in the storage and release of energy, and an ability to control these operations have made of puny physical man a veritable giant of which Gulliver's tales have no equal. Using only small amounts of energy in the thought process, man has gathered knowledge by the scientific method, and now he wields physical powers akin to those attributed to some of the heathen gods. In our day a mere child has at his command the energy of a purring automobile motor; a slight pressure of the foot and a car weighing more than a ton moves along a highway at breakneck speed. A workman simply by closing an electric circuit sets loose the pent-up energy of explosives, and mountainsides crumble with the blast. By a mere push or pull an engineer moves his massive train over hills, through valleys, and across continents. A housewife by the turn of switches floods her home with light, washes the clothes of the family, sweeps the floors, and obtains heat for cooking.

A Deductive System of Discovery

The Greeks supported, as a cardinal hypothesis, the view that it is possible to construct the universe by deductive reasoning from a few assumptions—preferably only one. They diligently sought such broad generalizations and tested them in the fire of "reasonableness." Undoubtedly they used specific experiences as a basis for the "hop, skip, and jump" to generalizations; but they too easily forgot these basic experiences as they generalized. Starting from such "certain" (uncertain, as we think today) heights, they reasoned *down* to specific details. Having implicit confidence in their hypotheses and the method of their logic (some think their logic was not well founded), they felt no need of putting the specific conclusions to the test of experience.

An Inductive System of Discovery—the "Scientific" Method

Why not reverse the process? Why not *move up* to the general from the specific—follow the *inductive system* of discovery? Among the earlier men of science, Galileo, the sixteenth-century Italian who placed his faith in experimental methods, is given most credit for this reversal. He was pre-eminent in originating the modern scientific method. Such direct appeals to nature as his are responsible for our scientific age. Deductive reasoning may still be used in science *after* generalizations are established through the inductive system of discovery; but the specific details resulting from such reasoning must always be tested experimentally.

Experimental facts. The method of science harbors no mystery or magic; it is simply refined common sense. It demands that data be gathered without bias, without preconceived notions. Facts of science are obtained by observation and experimentation. The same experimental procedure and technique, whether under the hands of one individual or another, should bring to light the same facts. It takes this corroborative evidence to make facts out of observations, and to guarantee that the observers have normal senses.

The hypothesis. But experimentation, painstaking and difficult though it be, may not be the most difficult step in truth-seeking. The formulation of a fruitful hypothesis to serve as a guide to experimentation often taxes the inquiring mind more severely. The proposition to be tested experimentally comes sometimes as an intuitive guess, and at other times it matures during an inductive analysis. In the latter case, the study and classification of the facts of experience lead to the recognition of relationships and uniformities; out of such a background, an original and fertile mind builds the principle needed to correlate and explain the observed data; and finally, the new insight

gives guidance to new investigations. Thus the making of a preliminary guess, and then the formulation of a fruitful hypothesis, represent the first steps in the building of a scientific law or principle; but the scientific method demands that these first conclusions, reached through logical and intuitive processes, be checked and rechecked by further experimentation.

In the midst of puzzling situations and facts, the scientist imagines one or more explanations, and then puts his guesses to the test. If this view, this guess, this fond hope, actually proves to have merit on further genuine investigation, it will deserve the name of hypothesis. Later, usually after many alterations, it may be justly called theory, then law or principle if it can survive investigation. It follows, therefore, that a vigorous imagination is an asset in scientific work, provided it is given free rein when guesses and wild surmises are in order, but curbed when scientific facts fail to give assent (Fig. 1).

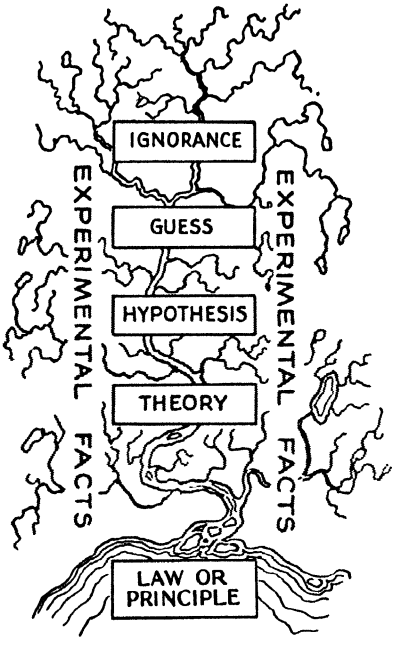


Fig. 1. Growth of Scientific Law.

A cyclic process. In summary, then, the scientific method is a cyclic process: hypotheses, built out of a background of experience, are tested by experiments; and these in turn yield results which enlarge the groundwork out of which new or modified hypotheses may emerge. By this cyclic procedure the scientific law or principle develops through the following steps: (*a*) ignorance, (*b*) preliminary guess, (*c*) hypothesis, (*d*) theory, (*e*) law or principle. The scientific attitude of mind, characterized by a willingness to abide by the facts and by an eagerness to see new rela-

tions, grows in the development of a new principle through these steps: (*a*) unbelief, (*b*) desire to believe, (*c*) belief, (*d*) faith, (*e*) knowledge. Many so-called principles and elements of knowledge change back to hypotheses and beliefs to be readjusted and even rejected as the store of scientific facts increases. Thus hypotheses and theories, unlike scientific facts, are changing things and are important insofar as they are useful in helping to interpret the facts in a simple, complete, and suggestive manner. It is this flexibility of hypotheses and theories that insures growth; this type of change gives strength to science—not weakness.

Fact and hypothesis. The scientist must always be on the alert to keep scientific facts free from scientific hypotheses. Is it a fact that the earth rotates on its axis once in a day and revolves about the sun once in a year? Or is this a hypothesis based on facts? About three hundred years ago honest and intelligent men, in defending the older and rival hypothesis that the earth was at the center of the universe, felt justified in mistreating their opponents. Yet night followed the day then as now; the seasons came and went as now they come and go. The new hypothesis won out because it gave a simpler, a more complete, and a more suggestive interpretation of the scientific facts which were being gathered then and have been gathered since.

Instruments of science. A scientist must be equipped with two important instruments of research. The first, an adequate mechanical instrument such as a microscope, a telescope, or a pair of balances, is so important that many discoveries have been delayed because of the crudeness of the apparatus used. The data-gathering aspect of the scientific method reduces essentially to the process of recording on some scale, dial, graph, or other metric device the workings of a piece of apparatus designed to measure a certain selected sequence of natural events while other phenomena, which might vitiate the results, are held constant. From inferences on how the apparatus works and

what the measurements mean, the required hypotheses are formulated. But were the scientific method to stop here, science would reduce in essence to beliefs concerning the working of machines. Much of theoretical science, especially mathematical science, accepts a more liberal interpretation of the scientific method and includes as a second important instrument of research the type of reflective thinking exemplified by mathematical analysis. Such an enlargement of the scientific field increases the fallibility of the scientific method, but gives to it freshness, vitality, and a fertile means of arriving at generalizations. This added source of error must be clearly kept in mind, especially since the success of science in contributing to human wants has caused the man of the street to shift his faith from the medicine man to the scientist as the source of infallible truth.

Thus we see that our scientific concepts, all our formal descriptions of reality, emerge from the patterns inferentially woven out of the data obtained from apparatus. The electron, proton, atom, or molecule, for example, is but a symbol which stands for a particular mosaic of measurements and inferences evolved in the scientific use of apparatus and formalized reflective thinking. But the scientist accepts as an unproved first principle that the external world exists in fact and that the pattern he gives to the universe has for the most part a one-to-one correspondence with reality. The contribution which science has made to human welfare is ample evidence that such a first principle has been a fruitful hypothesis.

The search for truth. The scientific method has helped man to discover that the happenings of nature occur in an orderly sequence and are rather simple of explanation. He finds he is not at the mercy of hobgoblins, sprites, or witches, but can, with intelligent co-operation, become the master, to a surprising extent, of the forces of nature. He is led to believe that the phenomena of nature have their explanation in natural, dependable laws which may be

formulated after a patient, open-minded, honest search after the truth.

The scientist is probably more interested in the search for truth than in truth itself. Many of his investigations are performed without the thought of practical applications. The casual observer is prone to ask, "What is the use of it all?" The scientist may well answer with the query, "Of what use is a newborn child?" History has shown that newborn theories of science, often unappreciated at first, on reaching maturity have often brought comfort, freedom, and wealth to a once uninterested world.

Co-operative science. Science must be co-operative endeavor. Newton once said, "If I have seen farther than Descartes, it is by standing on the shoulders of giants."

Questions and Problems

A suggested method of answering questions and solving problems is found in the Appendix.

1. The scientific principle develops through the following steps: (*a*) ....., (*b*) ....., (*c*) ....., (*d*) ....., and (*e*) .....

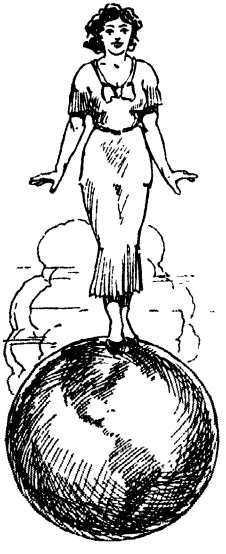
2. The theories of science may be expected to ..... from generation to generation.

3. The scientist must always be on the alert to keep scientific ..... free from scientific .....

4. List the characteristics of a scientific attitude of mind.

Suggested Readings

(1) Bell, Eric Temple, *The Search for Truth*, The Williams and Wilkins Company, Baltimore, 1934.
(2) Dampier-Whetham, Sir William C. D., *A History of Science*, The Macmillan Company, New York, 1932, Chaps. I and II.
(3) Jeans, Sir James, *The New Background of Science*, The Macmillan Company, New York, 1923, Chaps. I and II.
(4) Langdon-Davies, John, *Man and His Universe*, Harper and Brothers, New York, 1930, Chap. II.
(5) Loeb and Adams, *Development of Physical Thought*, John Wiley and Sons, Inc., New York, 1933, Part I.
(6) Planck, Max, *Where Is Science Going?* W. W. Norton and Company, Inc., New York, 1932, "Prologue," by Albert Einstein.
(7) Thomson, Sir J. Arthur, *Science for a New World*, Harper and Brothers, New York, 1934.



CHAPTER II

Size, Weight, and Balance

As children we soon become aware that other objects, such as persons, toys, furniture, trees, and stones, are encountered on every hand. We discover that our own bodies, although very different from lifeless forms, have much in common with physical things. Through touch and muscular effort we find that all bodies, including our own, have shape, size, and weight. But to say that an object is large or small is very indefinite, because a man would probably seem huge to an ant but small to an elephant. Some unit standard must be set up if sizes are to be accurately compared, for all measurements involve a comparison, the quantity to be measured being compared with a standard unit quantity (Fig. 2).

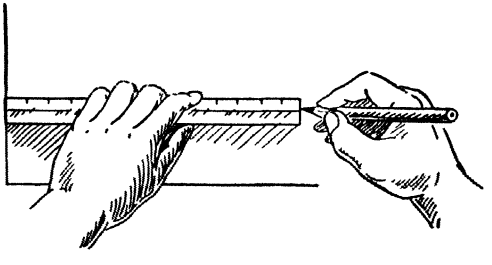


Fig. 2. Measuring Length.

Units of Length

The unit of length generally used among English-speaking people is the foot, supposed by some to have originated as the actual length of the foot of one of the English kings. The yard may have originated as one half the distance which a man can stretch both arms. But the length of feet and the stretch of arms vary from person to person. A more definite unit of measurement is required. Accordingly Great Britain, by an act of Parliament, defined the imperial yard as the distance measured at a temperature of 62°

Fahrenheit between two rulings on a certain bronze bar deposited and preserved in the Standards Office at London. In England a foot, to be of standard length, must have been at some time either directly or indirectly compared with and adjusted to the imperial yard.

During the French Revolution the units of length, as well as other standards which had been set up arbitrarily by generations of French kings, were discarded and a rational system was sought. In 1791 the National Assembly approved the meter as a length equal to one ten-millionth of the distance from the equator to the pole measured on a meridian through Paris. The necessary measurements were finished in 1799, and the length of the meter was determined. This length was marked off on a platinum-iridium bar manufactured especially for this purpose. The standard meter, which was thought to be susceptible of reproduction, was then defined as the distance at 0° centigrade between the two marks on this platinum-iridium bar now preserved in the International Bureau of Weights and Measures at Sèvres, near Paris. Later it was determined through more accurate geodetic measurements that the length recorded on the platinum-iridium bar is not exactly one ten-millionth of the distance from the equator to the pole and, therefore, not reproducible in the manner originally planned. Even so, the distance between the two lines on the metal bar has been accepted and designated as the international meter, and the yard in the United States is defined as $\frac{3600}{3937}$ of this standard. The unit of length usually employed by scientists is one one-hundredth part of the international meter and is called a *centimeter* (cm.). The unit of area is the square centimeter (sq. cm.); the unit of volume is the cubic centimeter (cc.). One inch (in.) is equal to 2.54 centimeters.

Simple Definition of Force

A muscular effort is required when one lifts a stone, throws a ball, or stretches a spring. We say that a force

is exerted in each action. Thus, by force a weight is lifted, a ball is put into motion, and a spring is stretched. Anything which accomplishes that which muscular exertion brings about exerts a force. Roughly then, a force is a push or a pull.



Fig. 3. Lifting a Stone.

Weight

The weight of a body is the force with which it is pulled toward the earth. The magnitude of this pull may be estimated by the muscular effort required to keep the body from falling, but as we all know from experience, such a method is inaccurate. If the precision demanded by science is to be obtained, an adequate technique of measurement must be discovered. A reliable machine which will register the weight of the body must be invented.

In our search for such an instrument let us turn to a study of the action of a spring. We know a force will stretch a spring. It is reasonable to suppose that if a given weight stretches the spring a certain distance, another weight of exactly the same magnitude will stretch it the same distance. At once this suggests an easy method of manufacturing a group of bodies of equal weight.

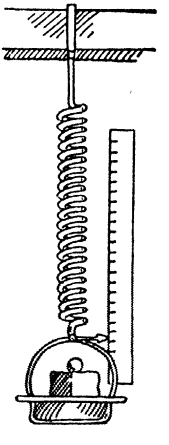


Fig. 4. Spring Balance.

The apparatus might take the form shown in Fig. 4 and consist essentially of a scale and a spring to which are attached a pointer and a weight pan. With the set of equal weights constructed as suggested above, a study of just how the spring stretches with increasing load becomes an interesting investigation. The results of such an experiment are shown on page 12.

The method of science warns against drawing general conclusions from a few instances. We may say, however,

STRETCHING SPRING

Number of the Equal Weights (W)	*Total Distance Stretched (S)*	*Stretch Divided by Weight (S/W)*
1	1.98	1.98
2	3.98	1.99
3	5.92	1.97
4	7.88	1.97
5	9.85	1.97

that the numbers of the third column at least suggest the conclusion that, within the limits of the precision of the apparatus and the so-called observational error of the experimenter, we may write,

$$\frac{\text{Stretch}}{\text{Weight}} = \text{Constant.}$$

In the shorthand language of algebra, the relationship becomes

$$\frac{S}{W} = k \qquad (2.1)^1$$

where S is the distance stretched, W is the weight applied (an integral number of equal weights in our case), and k is a constant. Notice that we have not designated a particular unit of length in measuring the distance, nor have we yet defined a unit of weight. We have simply made sure that the total distance measured or the total weight used is some definite part of a given magnitude of length or weight, respectively. With each particular choice of units, Equation (2.1) is modified only in the numerical value of the constant. When quantities are related as shown by Equation (2.1), they are said to be in direct proportion. Or more simply, for this specific case, doubling the weight

[1] Formulas in this and succeeding chapters are numbered to facilitate reference. The number preceding the decimal is the chapter number; the number following the decimal gives the sequence of the formulas within the chapter.

doubles the displacement; tripling the weight triples the displacement; and so on. This simple relationship was discovered and scientifically tested by Robert Hooke (1635–1703) and enunciated by him in 1676. Because of this simple law connecting the stretching of a spring and the force applied to it, a pointer reading, indicating the length of stretching, may be used to designate the weight of the suspended body. It is necessary, however, that the displacement produced by a standard weight be known, so that the proper numbers representing weights may be placed opposite the various marks of the displacement scale. This is the basis of the so-called spring balance.

The pound. In England the unit of force is the pull of the earth, at sea level and in the latitude of London, on a certain piece of platinum kept in the Standards Office and designated as the standard avoirdupois pound. It is necessary thus to specify where on the earth the force acts, since the pull of gravity changes with location, being greater on the surface of the earth at the poles than at the equator, and less above the earth than at its surface.

In the International Bureau of Weights and Measures at Sèvres, near Paris, is preserved a certain piece of platinum-iridium known as the *international kilogram.* One one-thousandth part of this standard is called the *gram.* The French Republic attempted to construct the standard kilogram so that it would represent the same amount of material as a liter (1,000 cubic centimeters) of water at a temperature of 4° centigrade. In the United States the pound mass is defined as a definite part of the platinum-iridium block preserved at Sèvres, specifically 0.45359243 kilograms. The pound force or simply the pound is the weight of this part of the standard block under a specified pull of the earth very nearly at sea level and 45° North latitude. Except in very precise measurements, the pound in England and the pound in the United States may be considered equivalent. An English pound to be a standard weight must have been at some time directly or indirectly

compared with and adjusted to the standard avoirdupois pound; the American, to the international kilogram.

With a set of standard weights on hand, a spring balance may have its scale divisions interpreted in terms of weights and be properly labeled so as to give direct readings in pounds.

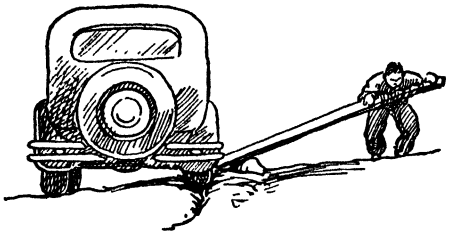


Fig. 5. Using a Lever.

The Lever

The principle of the lever. We usually elevate a small object such as a book, a chair, or a pail of water by lifting directly with the arms, legs, and body. But if our automobile goes into a ditch, direct lifting by human strength will usually be inadequate. We seek a long plank or pole and pry the heavy car out of the ditch and onto the road. Our actual experience has taught us that in this process the plank pivots at a point called the fulcrum, that the resisting force due to the weight of the car is applied at a position usually near the fulcrum, and that the acting force we exert is applied at another position which is as far from the fulcrum as possible. Such experience has probably suggested that the lever operates according to some definite law. Such a principle could best be investigated, not in the sweat and grime of a roadside accident, but in a laboratory, where apparatus and standards of measurement are available, and where conditions may be better controlled.

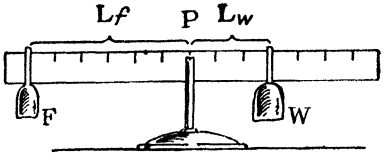


Fig. 6. A Simple Lever.

Consider the apparatus shown in Fig. 6. A long, thin bar, such as a meter stick, is pivoted at the fulcrum *P* and adjusted to right or left till it balances in a horizontal position. This balancing eliminates the need of considering the weight of the bar in subsequent measurements. Such a process serves as an example of a controlled experiment.

A simplification has been made by the elimination of what might be a troublesome variable. A weight W, which we shall call the resisting force, is suspended at a distance L_w from the fulcrum. If a balance is to be maintained, another weight F, designated as the acting force, must be applied at some distance L_f. Such a piece of apparatus yields data similar to the following:

RESULTS FROM USING A SIMPLE LEVER

W	L_w	F	L_f	WL_w	FL_f
100	40	100	40	4000	4000
100	30	100	30	3000	3000
100	30	50	60	3000	3000
200	15	100	30	3000	3000
200	10	500	4	2000	2000
300	40	150	80	12000	12000
300	30	200	45	9000	9000

A careful inspection of the last two columns leads to the conclusion that for the various combinations tried, the product of W and L_w is equal to the product of F and L_f. Thus within the limits of the experiment, we may set up the hypothesis that the lever acts according to the simple relation

$$WL_w = FL_f. \tag{2.2}$$

This equation has been tested over wide limits by various investigators, who have always found it to be true. In every case the distance of each force from the fulcrum—the lever arm of that force—is measured perpendicular to the line of action of the force. The product of a force and its lever arm is called its *moment.*

Thus it follows that the relation, which because of our limited experimentation we set down as a hypothesis, may now properly be considered by us as an established principle that may be expressed either in mathematical form (Equation 2.2) or in words as follows: *At equilibrium the moment*

of the acting force (the product of the acting force and its lever arm) *is equal to the moment of the resisting force* (the product of the resisting force and its lever arm).

If several forces act on each side of the fulcrum, it is necessary only to add the moments causing a rotation in one direction and put this quantity equal to the sum of the moments causing rotation in the other direction.

Three classes of levers. Levers are sometimes divided into three classes, depending on the relation between the positions of the fulcrum and the points of application of the acting force and the resisting force. The three classes of levers are illustrated in Fig. 7.

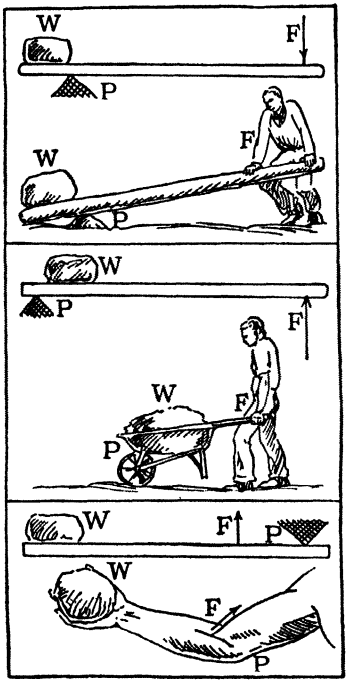


Fig. 7. Three Classes of Levers.

In the human body we find many levers. The joints are fulcra; contracting muscles produce the acting forces; the resisting forces are due to bodily weight, to the weight of attached objects, and at times to contracting muscles. The forearm is an example of a lever of the third class. As the forearm is raised, the elbow joint serves as the fulcrum; the pull due to contracting biceps is the acting force, and the weight of the forearm and attached objects is the resisting force. One may discover the three classes of levers in the foot by actually performing the following experiment:

Take a sitting position and cross the legs, thus causing one foot to be free from the floor. Move the ball of the foot up and down, observing that the ankle joint serves as a fulcrum. With the foot in its highest position grasp the toe with the hand and then try to lower the foot. The ankle joint, the fulcrum, is between the acting and resisting

force, and the lever is of the first class. With the foot in its lowest position push down on the toe with the hand and then try to raise the foot. The acting force is just in front of the ankle joint and much nearer to it than the resisting force. This is a lever of the third class. Now stand erect with both feet on the floor. Elevate the body by lifting the heels. The ball of the foot is now the fulcrum; the weight of the body is the resisting force applied at the ankle joint; the acting force is applied at the heel. This is a lever of the second class.

You will be interested to investigate the levers in other parts of the body, especially in the hands, arms, and legs. You will be impressed with the fact that the lever used is usually of the third class—a lever such that maximum motion is accomplished with the least muscular contraction. This permits the freedom of motion so beautifully displayed in a well-developed and agile body. It is true, however, that in such a lever arrangement the muscular forces must be many times greater in magnitude than the resisting forces, and hence the force that one may exert on some outside object will be much less than the actual pull of the muscles. We shall show in the next chapter how man has adjusted himself to this situation by the use of hand-tools and machines (many of them nothing more than levers of the first and second class) which permit the successful use of a small acting force against a larger resisting force.

Center of Gravity

By balancing the meter stick as shown in Fig. 6, we were able to neglect its weight. But often levers cannot be balanced when used in certain ways. How then may such a problem be solved?

Let a meter stick of uniform cross-section be selected and weighed on a spring balance. Then let one end be supported on a fulcrum and the other by a spring balance as shown in Fig. 8. At equilibrium $WL_w = FL_f$. The weight

W of the meter stick and the force F exerted at the end by the spring balance are known. L_f is equal to one meter. At once L_w may be calculated. Results of such an experiment show that for a uniform meter stick L_w is equal to 50 centimeters. This means that the sum of the forces acting on the particles which compose the bar may be considered as a single force acting at the bar's geometrical center, with a magnitude equal to the bar's total weight. The point at which this single force seems to act is called the center of gravity of the bar.

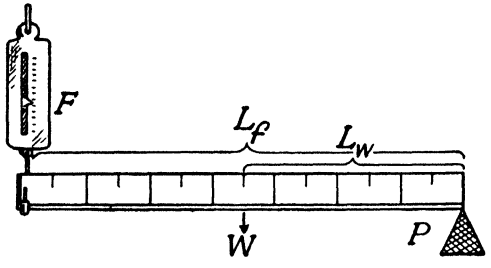


Fig. 8. The Center of Gravity of a Meter Stick.

Position of the center of gravity. When a body, such as a chair, is hung by a cord or suspended from the hand, the center of gravity must lie somewhere on a vertical line passing through the point of support (Fig. 9). If this were not true, the weight of the chair would be acting at a distance from the fulcrum and a downward rotation would result. The chair comes to rest at a position where the lever arm is zero, and this is at a position with the center of gravity directly beneath the point of support. If the chair is pivoted in a new position, the center of gravity will still be on a vertical line passing through the point of support. The intersection of the two lines thus obtained gives the position of the center of gravity, which may be at a point entirely outside the actual material of the chair.

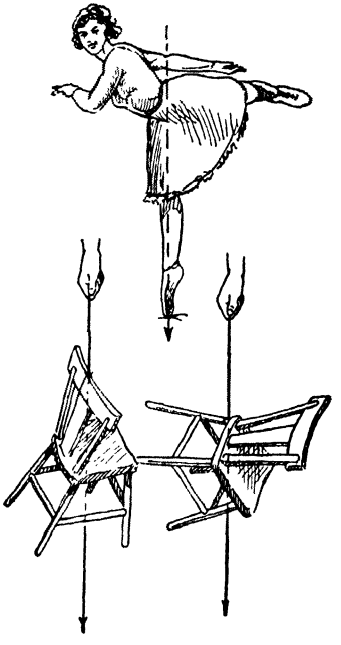


Fig. 9. Locating the Center of Gravity.

The center of gravity of the human body lies in the pelvic region and shifts about somewhat as various bodily positions are assumed. Teachers of gymnastics realize that, owing to the difference in build, the center of gravity is higher in men than in women, and that for this reason bar work is usually easier for the men. The ballet dancer takes the positions which bring the center of gravity above the point of support, the toe (Fig. 9).

Parallel Forces Acting in a Plane

If a spring balance is made to replace the fulcrum of a lever in equilibrium (Fig. 6), it is possible to determine just what upward force the fulcrum support has been exerting. When all weights are properly recorded, it is found that the weight registered by the spring balance is exactly equal to the weight of the meter stick plus the suspended weights F and W. One sees clearly from the figure that all these forces are acting in a plane and in a parallel direction. This simple experiment suggests the hypothesis that *when parallel forces are in equilibrium, the sum of those pulling in one direction is just equal to that of those pulling in the opposite direction.* In the study of the development of physics we find this hypothesis to be an experimentally tested principle.

Let us apply this principle to a simple case. Suppose that an object is at rest on a table top. The table is slightly depressed by the weight, and because of the deformation certain elastic forces arise just as in the case of the stretched spring. The deformation continues till the upward push of the table is equal to the pull of gravity.

Concurrent Forces Acting in a Plane

Let a weight be supported by two spring balances as shown in Fig. 10a. The three forces are in a plane and act out from a point and are thus concurrent. The sum of the weights recorded by the two spring balances is greater than the weight supported. A thoughtful consideration of the problem will soon yield the conviction that the

spring balances actually pull against each other as well as support the hanging weight. If this is the case, a balance pulls as if the force it exerts is made up of two parts, or, as we shall say, two components, one in the vertical direction, which helps to support the weight, and the other in the horizontal direction, which pulls against a similar and oppositely directed component of the force due to the other balance. In Fig. 10a we have let arrows represent these forces, and the simple geometrical figures should help the student to get a picture of the meaning of a force and the components which may be used in its stead.

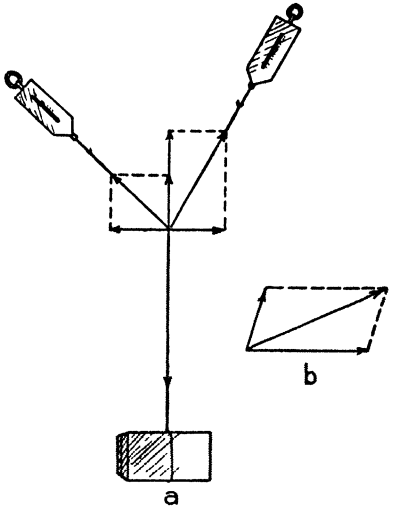


Fig. 10. Concurrent Forces.

Experimental investigations and theoretical studies have shown that, if a diagonal of any parallelogram is used to represent both the direction and size of a certain force, the sides (see Fig. 10b) represent the components of the force. In the above experiment we decided to imagine the components directed vertically upward and horizontally to the left or to the right, depending on the force pictured. Thus in this special case the parallelogram takes the form of a rectangle.

By the simple expedient of replacing a force by its components, it has been possible to transfer the problem of concurrent forces to one of parallel forces. At once we may apply the rule just discovered for parallel forces and conclude that, *when parallel forces or concurrent forces acting in a plane are in equilibrium, the components acting vertically upward balance those acting vertically downward, and the components acting horizontally to the right balance those acting horizontally to the left.*

Complete Requirements of Equilibrium for a Balanced Lever

Combining the principle of the lever with the law just stated, we may write down the following rule as the complete requirement for equilibrium in the case of the lever: *At equilibrium the sum of the force components acting vertically upward must be equal to the sum of the force components acting vertically downward; the sum of the force components acting horizontally to the right must be equal to the sum of the force components acting horizontally to the left; the sum of the moments tending to turn the lever about its fulcrum in one direction must be equal to the sum of the moments tending to turn the lever in the opposite direction.*

The principle is easily extended to the equilibrium of a body under forces not restricted to a plane, but this general case is beyond the scope of this text. We may state, however, that in general a balance of the forces and a balance of the moments are required for equilibrium.

Stability

The rule just stated gives the criteria for equilibrium, but it does not tell us explicitly the nature of the stability achieved with each balance. In general, bodies in equilibrium under the action of gravity may be said to be in either stable, unstable, or neutral equilibrium. If when a body is tipped it falls back into its original position, it is said to be in stable equilibrium; if when tipped it falls away from its original position, it is said to be in unstable equilibrium; finally, if it remains in equilibrium when displaced, it is said to be in neutral equilibrium. In the first case the center of gravity is

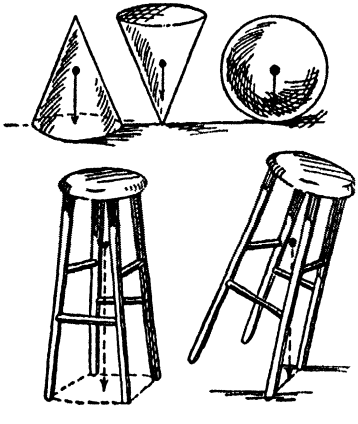


Fig. 11. Types of Equilibrium.

raised; in the second case it is lowered; and in the third case it is neither raised nor lowered when the body is tipped. The cones and ball (Fig. 11) illustrate the three types of equilibrium. From the definition just given try to classify these objects as to type.

If a body, such as a stool, is in stable equilibrium, the vertical line containing its center of gravity will pass through the area bounded by the straight lines connecting the points of support (Fig. 11). If the stool is now tipped so that the vertical center-of-gravity line touches a boundary of this area, the body will be in a condition of unstable equilibrium. A slight touch will cause it to move back to its original stable condition or to topple over.

Try the experiment with a chair. If the area of support is large and the center of gravity is low, considerable tipping (lifting of the center of gravity) will be required before the condition of unstable equilibrium is reached. Therefore, a chair with low center of gravity and large area of support is not easily upset. A tall vase often requires ballast such as sand in order that the center of gravity may be lowered and the stability increased. It is interesting to notice this principle operating in a person making ready for an attack. He lowers his center of gravity by assuming a crouching posture and increases the area of support by a separation of the feet. A charging football lineman may even use his hand as a point of support to increase his stability (Fig. 12).



Fig. 12. Increasing Stability.

Posture

The same laws which apply to the equilibrium of levers also apply to the human body. For example, as one moves the head forward and back, one is convinced that this part of the body is delicately balanced on a fulcrum, the top-

most vertebra, called the *atlas* (Fig. 13a). The head weighs from 12 to 15 pounds; hence, for equilibrium, the atlas must push upward with this same force. This the atlas is able to do, because it is supported on the framework making up the human skeleton. But for equilibrium the moments acting about this fulcrum must also balance. If the center of gravity of the head is brought directly above the atlas, the moments due to gravity will be zero, and equilibrium will be established without muscular effort. But the area of support furnished by the atlas is small, and a slight tipping of the head would cause the vertical center-of-gravity line to pass outside the area of support, and the head would fall unless checked by muscular forces. Actually then, equilibrium will probably never be established, at least over long periods, without the expenditure of muscular effort. But obviously this effort may be reduced to a minimum by a posture that maintains the center of gravity more or less directly over the atlas.

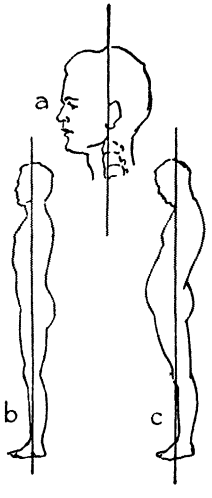


Fig. 13. Posture.

On the other hand, if the head is allowed habitually to tip forward, there will be a need of continued muscular exertion to balance the moment now produced by gravity, because the center of gravity seldom if ever takes a position over the point of support. Such a head position obviously occasions a waste of muscular effort.

When a further study of the body is made, one is led to the conclusion that a correct posture is attained when the moments due to the weights of the various body parts are balanced with the least expenditure of muscular effort. Such a posture is one in which the long axis of the body including the neck and head is in a vertical line. The head is straight above the chest, hips, and feet; the chest is up and forward; the abdomen is flat, and the curves of the back are not pronounced (Fig. 13b). In an incorrect

posture the head hangs forward; the chest is flat; the abdomen protrudes; and the curves of the back are exaggerated (Fig. 13c). In this last case the weights of the body are definitely out of balance, and counteracting moments must be set up by muscular exertion.

Questions and Problems

1. How many centimeters are there in a foot?

2. If a one-pound weight stretches a spring 1.5 inches, how far will a three-pound weight stretch it?

3 Assuming that a frictionless teeter is balanced when not in use, what is the weight of a child sitting 10 feet from the fulcrum if balance is re-established by a 100-pound child sitting 8 feet from the fulcrum?

4. (*a*) List the types of levers. (*b*) Classify the following articles as to type: scissors, nutcrackers, tweezers, claw hammer, monkey wrench, and tin snips.

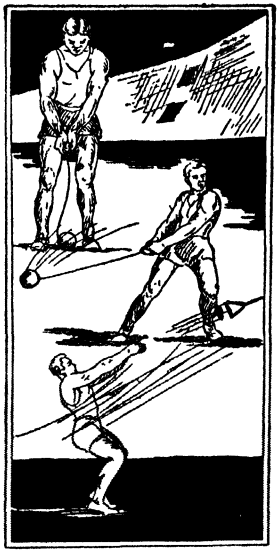
5. Make a simple drawing illustrating a child pulling a sled over a level surface. By means of arrows sketch the force exerted by the child and the horizontal and vertical force components. Indicate the force which moves the sled over the snow.

6. Explain why a circus performer is able to walk a tight-rope.

7. Classify the shifts (upward, downward, forward, backward, to the right, to the left) of the center of gravity of the human body when a person (*a*) lifts the arms, (*b*) leans forward, (*c*) stands on the right foot, (*d*) changes from low- to high-heeled shoes.

Suggested Readings

(1) Kimball, A. L., *College Textbook of Physics*, Henry Holt and Company, Inc., New York, 1923, pp. 29–38.
(2) Kirkpatrick and Huettner, *Fundamentals of Health*, Ginn and Company, Boston, 1931, pp. 228–236.



CHAPTER III

Force and Motion

Through experience we have learned that extra muscular exertion is required when we start running to catch a train or when we stop suddenly to avoid collision with a passing automobile. Then, too, we have discovered that we push or pull whenever we start or stop the motion of the many objects we encounter on every hand. This means that we exert forces whenever we change the motion of ourselves or the objects of our environment. We shall be interested to determine just how force is related to motion as we extend the study of forces begun in the last chapter. But as we proceed, we must equip ourselves with clear ideas concerning the words used and with precise definitions of the units of measurements involved.

With a little thought we become aware that motion is composed of two elements, ***distance*** and *time*, because to say a body moves is to imply that after a lapse of time it is found in a new position. In Chapter II we introduced the units of length; we shall now define the unit of time, and then certain types of motion.

Unit of time. The time needed for the apparent revolution of the sun above the earth, actually the time of one revolution of the earth, is one of the most constant quantities which we meet in nature. The time, averaged throughout the year, between two successive passages of the sun overhead (across the meridian) is known as the mean solar day. The mean solar second is defined as $\frac{1}{60} \times \frac{1}{60} \times \frac{1}{24}$

of the mean solar day. It is one of the oldest standards of measurement, not having been changed since the time of the ancient Chaldeans. It is probably trite to say that clocks and watches are used to measure time. Here you see a stop watch which measures time intervals to fifths of seconds (Fig. 14).

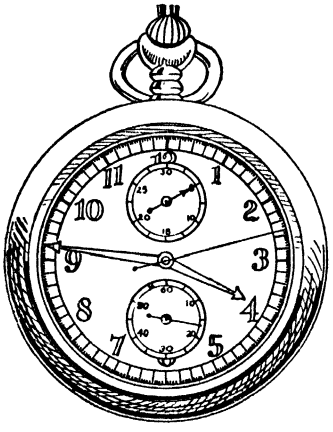


Fig. 14. A Stop Watch.

Velocity

If you recall the danger of skidding attendant upon the rapid turning of a corner, you will be convinced that forces are as intimately associated with direction changes as with speed changes. For this reason it is desirable to include both items in the precise definition of motion. The precise statement of the motion of a body is designated as the body's *velocity*. If a certain object, such as an automobile, moves equal distances in equal times on a smooth road with no curves, elevations, or depressions, it is said to have a uniform velocity. The velocity is uniform, first, because the speed of the motion is uniform, that is, the car moves equal distances in equal times; second, because the direction of the motion does not change. In Fig. 15a you see a car decreasing its velocity by decreasing its speed. The car of Fig. 15b may show a steady and constant speedometer reading (uniform speed) on the winding road, but even so it experiences a change in velocity because of the changes in the direction of the motion. The car illustrated in Fig. 15c decreases its speed and also changes its direction of travel as it goes over the

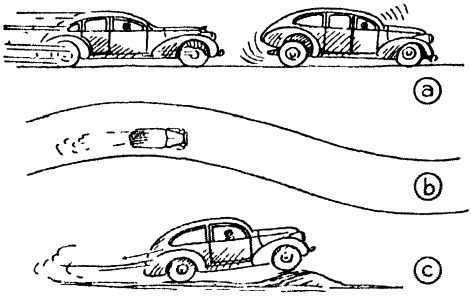


Fig. 15. Changes of Motion.

bump in the road. As defined, then, velocity has two aspects: the speed aspect which determines the rate of travel, and the direction aspect which has to do with the direction of travel.

Acceleration

Whenever the velocity of a body changes because of a change in speed or direction, or both, the body is said to have been accelerated. Thus the cars illustrated in Fig. 15 experience acceleration—the first a negative acceleration with a decreasing speed, the second an acceleration with a changing direction of motion, the third an acceleration with both a changing speed and a changing direction. If the change in velocity comes in equal amounts during equal periods of time, the acceleration is classed as uniform acceleration. Thus, when a *circular* motion (Fig. 19) shows a uniform speed (equal distances traveled in equal times), the direction aspect of the velocity *changes* equal amounts in equal times. Here the acceleration is constant in amount but swings around as the particle moves, being directed always toward the center. Or again, when an automobile moving in a *straight* line gains speed at a steady rate, say a speed of one mile per hour each second, the speed aspect of the velocity increases equal amounts in equal times—an example of uniform acceleration.

Falling Bodies

When an apple and a leaf become detached from a twig, they begin to fall (Fig. 16). Any unsupported body, we know, falls toward the earth under the unbalanced force of gravity, its weight no longer being opposed by an equal and opposite force. The apple descends with a swift motion, seeming to go faster and faster as it falls. The leaf settles more gently, or at most flutters a bit, as it moves

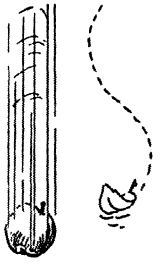


Fig. 16. A Falling Apple and Leaf.

toward the ground with a downward motion not difficult to measure. The casual observer is likely to conclude that the apple falls more *rapidly* than the leaf *simply because* it is attracted to the earth with a greater force. Then, too, the same observer remembers that a heavy stone falls so swiftly that the eye has difficulty in following it; that snow flakes and feathers, if in still air, fall gently downward with small, uniform speeds; that the tiny water droplets of a fog may not fall at all, and that smoke usually rises. Undoubtedly Aristotle (384–322 B. C.) had observed such simple phenomena when about 350 B. C. he wrote in the fourth book of his work *On the Heavens* the statement: "That body is heavier than another which, in equal bulk, moves downward quicker." The unquestioning mind, or the mind which places "reasonableness" ahead of experimentation, would easily jump to the conclusion: "Of course, of course—the 'heavier' the 'quicker'; this follows without doubt!" But if you balk at such a mental leap, try the experiment with a feather and a coin, in equal bulk (volume), and you will see where Aristotle got his idea. We hope you will not close your mind after this simple test, because you may get around to trying the experiment with an iron ball and a wooden ball of the same diameter (same bulk). If you do, you will find that Aristotle was wrong, that he generalized toward that which seemed "reasonable" with too great haste. The Greek thinkers insisted that it is possible to build the universe by *deductive reasoning* from one, or at most a few, "reasonable or self-evident" assumptions. This attitude stimulated Aristotle to leap to an unwarranted generalization, to a doctrine accepted on his authority through the Middle Ages and into the Renaissance. By allowing pieces of wood and lead to fall from a tower, Leonardo da Vinci (1452–1519) recognized that falling bodies take on accelerated motion, but we have no record that anyone seriously questioned the doctrine of Aristotle until the day of Galileo Galilei (1564–1642), a period of time covering eighteen centuries.

Does it not seem obvious that a knowledge of just how a body falls owing to its weight must be ascertained in a situation where *no force* except *gravity* acts, or where the other forces are so *small* as compared to the body's weight as to be negligible? At least, the critical mind would be willing to accept such a conclusion as a guide to experimentation. Using this suggestion and the materials at hand, the experimenter (and this could have been Galileo) allows a sheet of paper (or parchment) to fall. Something tends to buoy it up, especially when the sheet is in a horizontal position. When the paper is crushed, the buoyant effect is greatly reduced. Our experimenter hastens to find a feather and a coin. They fall in accordance with the doctrine of Aristotle. He crushes the feather. The objects fall with nearly the same speed. At once he concludes: to render negligible the buoyant action of the air, make use of objects with large weight and small bulk. In excitement he searches for two lead balls —a heavy one and a light one!

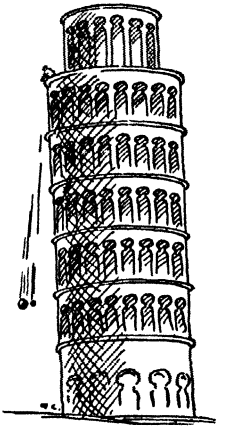


Fig. 17. The Leaning Tower of Pisa.

As a young professor at Pisa, Galileo came face to face with the cherished Aristotelian doctrine that the rate at which a body falls depends upon its weight—a cherished doctrine, not because the people cared much how bodies fall, but because the idea could be traced back to Aristotle. Prior to Galileo it did not occur to anyone actually to test the doctrine by experiment. But one morning, undoubtedly after a good deal of private experimentation, Galileo ascended the leaning tower of Pisa (Fig. 17) and before the assembled audience allowed a "one-pound shot and a one-hundred-pound shot" to drop at the same instant. The balls started together, fell together, and struck the ground together. Thus he demonstrated to the satisfaction of some present, but not all (for it was difficult for some to give up the teachings of Aristotle), that a heavy shot falls

at the same rate as a lighter one. His studies finally convinced him "that in a medium *totally devoid of resistance* all bodies would fall with the same speed." With the invention of the air pump a few years after his death came the complete verification that all bodies, both light and heavy, fall at the same rate when free from air resistance. This may be vividly shown by causing a penny and a feather to fall in a long tube from which the air has been pumped (Fig. 18). Shortly we shall discuss how air resistance modifies the motion due to gravity alone.

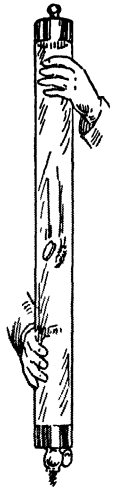


Fig. 18. A Feather and a Penny in a Near-Vacuum.

Force and Change of Motion

The demonstration at the leaning tower of Pisa was indeed an excellent qualitative experiment, but Galileo sensed the need of accurate measurements. Accordingly, he conceived the idea of letting a perfectly round and polished brass ball roll down a smooth incline; thus frictional forces were reduced to a minimum. In this manner he used a constant component of the force of gravity to give him speeds he could measure. Accurate clocks were not available, so he compared times by weighing the amounts of water escaping from a small hole in a bucket during the intervals to be compared. His measurements led to the discovery that freely falling bodies, bodies acted on by the constant force of gravity and unaffected by air resistance or some other external force, have a *uniformly accelerated motion.* Referring to this discovery, Joseph Louis Lagrange (1736–1813) said: "It required an extraordinary genius to unravel laws of nature from phenomena which were always before our eyes, but the understanding of which had escaped philosophical inquiry."

Even casual observations convince one that force and motion are intimately related, but these experiments performed by Galileo were the first to reveal that a body not

balanced by the upward pressure of its support will fall under the force of gravity (neglecting air resistance) with a *uniformly accelerated motion.* Thus out of this monumental scientific research, marking as it does the dawn of the scientific method, there developed the new concept, *acceleration,* and its association with the idea of force. This means that a force may be *detected* by the rate of change of velocity (acceleration) which it imparts to an object, as well as by the push or pull needed to oppose its action.

Galileo found that a body near the earth falls with a uniform acceleration of approximately 32 feet per second each second, unless its motion is affected by air resistance or some other external force. This means that the force of gravity acting on any freely falling body originally at rest is sufficient to change its velocity from 0 feet per second to 32 feet per second during the first second, from 32 feet per second to 64 feet per second during the next second, and so on. If thrown vertically upward, an object will lose velocity at the rate of 32 feet per second each second, finally come to rest, and then begin a uniformly accelerated motion downward.

Newton's First Law of Motion

Sir Isaac Newton (1642–1727) erected a magnificent scientific structure upon the experimental results of Galileo. Just what inferences may be drawn from such experimental data? First let us remember that the great pioneer experimenter demonstrated that when a force acts on a body *its velocity is changed,* and further, that a constant force imparts uniform changes in velocity, or uniform acceleration. Suppose an object is at rest and a constant force is applied to it. The object moves faster and faster and faster; its velocity increases at a steady rate. But suppose the force causing the acceleration is removed; what happens? The velocity ceases to increase, the acceleration stops (there is now no force to change the velocity), and the body maintains the velocity which it acquired during the action of

the force. This velocity the body *must keep*, because to rid itself of a velocity requires the action of a force, and there is now no force acting. Thus, through careful and logical inferences, Newton offers us a picture of how bodies perform when no force acts, and formulates his First Law, which may be stated thus:

All bodies remain at rest if at rest, and in uniform motion in a straight line if in motion, unless some external force is applied to compel a change.

Does such a conclusion agree with our experience? Can it be possible that it takes no force to keep an object at a uniform velocity? To see bodies remaining at rest is a common experience, and it is thus easy to agree with this first aspect of the law. But what casual observer has seen a body which is in motion continue to have that motion without change? Rocks roll down mountainsides and come to rest; water plunges over cliffs but finally comes to comparative rest in a placid lake; machines without the application of power run down. Is not *rest* the natural thing instead of *rest* and *uniform motion in a straight line?*

To find the answer let us search with Newton behind nature's seeming camouflage. When objects are in motion it is so difficult to free them from external forces, especially those of friction, that rarely do we observe the ideal situation stated in the second part of the law. We usually identify the force we apply to overcome friction with the one we suppose is required to keep the object in motion; when, as a matter of fact, we simply make the force we apply equal and opposite to the force which would otherwise bring the object to rest—we simply reduce the *resultant force* to *zero*. If from childhood we had performed all our acts on the surface of a very slippery ice pond, the second part of the law would have seemed just as reasonable as the first.

Continuing our search, let us approach some of our common experiences in the light of Newton's statement. Why

do we sink deeply into the seat cushions as a careless driver lets the car go forward with a lunge? Because bodies (our own bodies) tend to remain at rest or to maintain a uniform velocity. Why do we lurch forward when brakes are very suddenly applied? Because bodies having a given velocity tend to keep that velocity. As we find ourselves moving forward, or more precisely, as we discover the car slowing down, we muster sufficient physical exertion to check our motion. In automobile accidents why does a person crash head-first through the windshield, or strike the steering wheel with perhaps fatal force? Because the telephone pole which is struck while the car is traveling 60 or 70 miles per hour rapidly changes the velocity of the car, and all the parts of the structure, including the occupants, tend to maintain the excessive speed and direction of travel.

Thus with Newton's help we are able to interpret certain of our simplest experiences. We see that his law is an idealization. It illustrates how the scientific method strips away the odds and ends that so often confuse the casual observer as he interprets his physical environment. Newton recognized that this idealization is achieved in planetary motions, where frictional forces are infinitesimally small. His laws received their verification and triumph in the explanation (to which his laws led him) of why our earth and moon and the planets move in accordance with the laws of John Kepler (1571–1630). This was the greatest scientific problem of the sixteenth century. Because of the tendency of all moving bodies, and therefore of a planet, to continue in uniform speed in a straight line, a planetary system, when once set going, needs no force to keep the planets moving, but a force will be required to explain the continued deviation from a straight path which takes place as a planet swings about the sun. Newton argued that the same type of force, the force of gravitation, existing between an apple and the earth and causing the apple to fall, also exists between the sun and a planet, causing the planet to fall toward the sun (the motion needed if the former is to

deviate continually from a straight path). Before we close this chapter, we shall discuss briefly what Newton discovered concerning the nature of gravitational attraction.

Centrifugal and Centripetal Forces

The tendency of a moving body to preserve its speed and direction of travel (its velocity) accounts for the fact that mud flies tangentially from a rapidly moving wheel, for the fact that a stone may be thrown from a sling, and for the fact that the hammer flies away as the whirling athlete lets it go. This tendency is felt also as one swings a pail of water rapidly in a circle (Fig. 19). Try it. Unless you are careless, the water will not fall out and wet you, since the pail and the water tend to fly away from the center of the circle with a force which increases rapidly as the speed of turning is increased. Actually, the water when overhead not only remains in the bottom (or should we say the top?) of the pail, but it may press upward on the bottom with a force greater than that with which it normally presses downward (owing to gravity) when the pail is not being swung around. Thus the pail keeps the water from flying away and the arm keeps the pail from flying away. This means that the "out-from-the-center" or *centrifugal* forces are balanced by the "toward-the-center" or *centripetal* forces. Do away with the centripetal force by letting go of the pail, and, were it not for the force of gravity and air resistance, the pail would hurtle away forever in a straight line in the direction it was going at the time it was let go, a direction tangent to the circular path in which it was made to travel by the arm.

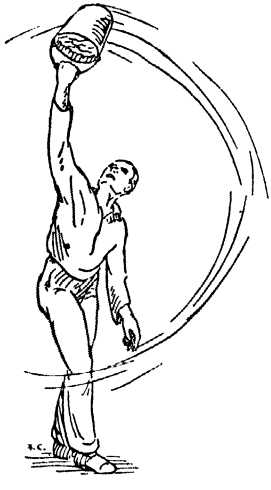


Fig. 19. The Pail Contains Water.

The tendency of a moving body to continue in a straight line is felt also as one rides rapidly in a car around a curve

(Fig. 20). One finds oneself struggling with the seat cushions and the side of the car for the right of way. But let us turn our attention to the car; its safety is also our safety. The centripetal force, which urges the car from a straight line and into a curved path as the steering wheel is turned, comes from the ability of the rubber of the tires to grip the pavement (Fig. 20a). If the rubber fails to grip, as often it does on an icy road or in loose gravel, the car does not make the turn—it *skids*. The centrifugal force, representing the tendency of the car to move in a straight line, acts at the center of mass (gravity) of the automobile, not at the surface of the road where the centripetal force acts. Since the center of mass is some distance above the road (it is nearer the ground in recent automobile models than in old models), a force moment results which tends to turn the car over (Fig. 20b). The higher the center of mass, the larger is the moment and the greater the danger of tipping over.

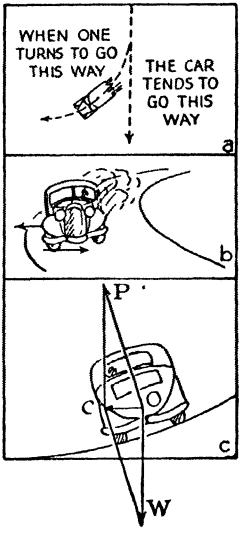


Fig. 20. Curves and Turns.

Thus, under the action of a large centrifugal force caused by high speeds on sharp curves, a car is in danger of skidding or tipping over (Fig. 20b). But by banking a road (Fig. 20c) such a danger may be avoided. In the figure the weight of the car is represented by OW, the push up of the road by OP, and the resultant of these forces by OC. When the car is not moving, the resultant force will tend to tip the car toward the inside of the curve; but as the curve is taken at increasing speeds, an outward (centrifugal) force will be reached which will exactly balance the inward force. At this speed the car will be in no more danger of tipping over or skidding than if it were on a level, straight road. If the speed is increased beyond this point, of course the car will tend to turn over.

At excessive speeds emery stones and flywheels may burst, the centrifugal forces exceeding the cohesive forces of the material. Very large centrifugal forces are also experienced in the testing of airplanes. For example, the pilot climbs to a height of 18,000 feet, turns the plane into a steep dive, drops 10,000 feet, reaching a speed of more than 300 miles per hour, and then "pulls out" rapidly into a horizontal direction. The centrifugal force in such a case may be as great as from 8 to 14 times the force of gravity. The pilot presses down on the seat as if he weighed a ton; the liquids of his body move forcibly downward; blood leaves his head; often he experiences temporary blindness; and in extreme cases, when he accidentally "pulls out" too rapidly, blood vessels of the head are broken and intestines are ruptured. The centrifugal force is balanced by a centripetal force arising from the resistance between the air and the surfaces of the plane. If the structure of the airplane can withstand the strain, the airplane passes the test. But under this ordeal planes have been known to lose their wings, and pilots their lives.

The dairy separator in which the heavier milk moves out farther from the axis of rotation than the lighter cream, the centrifugal drier used to remove water from clothes, and the governor placed on machinery to maintain uniform speeds of rotation, are a few of the many practical uses of centrifugal force. Whirling rotors which develop a centrifugal force equivalent to 7,000,000 times the force of gravity are being used by scientists in an attempt to separate the atoms of different weight (isotopes) of the same chemical element. Such a rotor turns so fast that a point on its periphery travels 16,000 miles per hour—approximately 4.5 miles per second. Under such a force the spinning metal may fly apart; therefore the scientists barricade themselves behind a bank of sand supported by wooden planks.

The Motion of Projectiles

To illustrate the motion of a projectile, let the apparatus as shown in Fig. 21 be set up on a horizontal table and adjusted until the balls *A* and *B* are at exactly the same height. When released, the rod *C* moves rapidly to the right. The ball *A* is released to fall freely under the action of gravity, and at the same instant the ball *B* is projected horizontally. As nearly as one may judge from the sound of impact, the balls strike the table top at exactly the same instant. Can this be possible? Let the experiment be tried again. Yes, the balls do fly through their paths of different length in the same time. As we seek to understand this interesting performance, undoubtedly we shall find helpful the principles established by Galileo and Newton. First, we remember that the starting or initial velocities of the balls are not the same. Ball *B* starts off with an initial horizontal velocity but with no vertical velocity. Ball *A* has no initial velocity whatsoever. Gravity acts only in the *vertical* direction, so it cannot be counted on to change the initial horizontal velocity of ball *B*. Then too, the air resistance encountered by a metal ball moving at slow speed is very small. Thus there is no perceptible horizontal force to compel a change and, therefore, the starting horizontal velocity tends to persist (Newton's First Law). This velocity carries ball *B* steadily to the right. Yet all the while the horizontal component of velocity is persisting, gravity is doing its bit in the vertical direction, accelerating the ball just as if no horizontal velocity were present. Since both balls, in terms of the vertical, start from rest and fall from the same height, both will be compelled by gravity to make the vertical distance

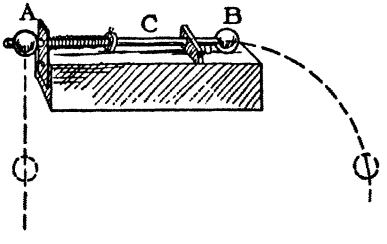


Fig. 21. The Motion of Projectiles.

in exactly the same time (Galileo's discovery). Thus both should strike the table at the same instant. All this means that ball *B* acts as if it had *two* distinct motions taking place simultaneously. It falls vertically with a uniformly accelerated motion, maintaining vertical positions exactly identical with those of *A*, and it has an initial horizontal velocity acquired during the time of impact with the rod *C*, which it *maintains* during the flight. The net result of these two motions is the curved path taken by the ball.

The same analysis holds for the flight of a bullet shot horizontally from a rifle. As unbelievable as the statement may seem to the uninitiated, we must conclude that the time of flight (time before striking the ground) of such a missile is the same as the time it takes a piece of lead to fall from the end of the gun barrel vertically downward to the ground. As one aims for the mark in target practice, the barrel points up at a very small angle above the straight line to the target. At a range of 300 yards, for example, the bullet from a high-powered rifle, as it moves swiftly on its curved path to the target, may rise two feet above this straight line (the distance depending on the weight of the bullet and the power of the gun). In long-range shooting, guns fire at rather large angles above the horizontal. Were it not for air resistance, which at the high speeds of bullets must be taken into account, a maximum range would be achieved by firing at an angle of 45°. But in practice, angles greater than this are used to obtain a maximum range, the angle used depending upon atmospheric conditions and on the size and shape of the projectiles. A simple experiment with a jet of water from a garden hose will make this discussion more clear. Try reaching out for a maximum distance with the jet, and you will see that the greatest range is obtained at some definite angle.

In all the examples given above, we have discovered a very interesting property of all bodies: The tendency to remain at rest if at rest, or to resist a change in velocity

(speed and direction) if in motion. *Inertia* is the name given to this property of all bodies.

Inertia and Mass

How may inertia be measured? Continuing our investigation, let us search out an athlete who is putting a 16-pound shot. Ask him to modify his usual form and heave the weight with all his might (Fig. 22) in a horizontal direction. Next ask him to put the shot in a horizontal direction with less muscular exertion. The first throw is greater than the second. Since in both cases the shot falls through the same vertical distance, the times of flight are the same—simply the time required for the shot to drop from the man's shoulder to the ground. The discussion of the motion of projectiles has prepared us for this statement. Interpreting our observations in the light of this fact, we see clearly that in the first throw the shot must have had a greater horizontal velocity than in the second; and further that the greater muscular exertion of the first throw produced, during the heave, a greater acceleration than in the second. At once we conclude that the larger the force acting on a certain body, the greater the acceleration.

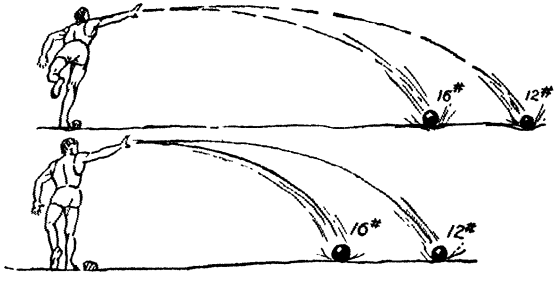


Fig. 22. Putting the Shot.

Now supply the athlete with a 12-pound shot and ask him to heave this new weight with all his might in a horizontal direction. The distance of the throw is found to be greater than that with the 16-pound shot. The weights fall through the same vertical distance and the times of flight are thus the same. The horizontal velocity of the 12-pound shot, therefore, must have been greater than that of the 16-pound shot. This means that during the process of throwing, the smaller weight was given the greater

acceleration. We may assume the muscular efforts, and therefore the forces, in the two mighty heaves to be approximately equal. This being true, the *lighter* shot must have opposed a change of motion the *least. Thus it follows that the 12-pound shot has less inertia than the 16-pound shot.* This suggests the guess that the inertia of a body is directly proportional to its weight.

With this in mind let us study the results obtained by Galileo at the Tower of Pisa. The 100-pound weight is attracted to the earth with a force one hundred times greater than that which attracts the one-pound weight. But, according to our hypothesis, the inertia of a body is directly proportional to its weight; hence, the 100-pound weight should oppose a change in motion with an intensity 100 times greater than that exerted by the one-pound weight. That is, although the heavier weight is urged to the ground with a force which is 100 times greater, the opposition to a change of motion, according to our hypothesis, is also 100 times greater. Only if the acceleration is the same for both balls will the force due to gravity and the opposition due to inertia in each case be balanced in a similar manner. And this common acceleration is exactly what Galileo found to be true. Hence, we conclude that our assumption is correct: the *inertia* of a body is *directly proportional* to its *weight.* Thus we may easily compare the inertias of a series of bodies by comparing the weights obtained by the use of a spring balance. We remember, of course, that the measurements must be made with the bodies at the *same locality,* as were those which fell at the tower of Pisa. A body's inertia, it seems, is a property of the object itself. It is an index of the quantity of matter in the body and does not depend, as does weight, upon factors outside itself, as, for example, the nonuniformity of the pull of gravity over the surface of the earth.

Units of mass. Although we have found a simple means of comparing the inertias of a group of bodies, a standard of inertia must be established before quantitative measure-

ments may be made. The quantitative measure of the inertia of a body is called the *mass* of the body. Thus it follows that the masses of a series of bodies bear the same numerical relationship as the weights of these same bodies if determined at the same locality. In the interest of simplicity, therefore, it would be natural to designate as the standard unit of inertia the mass of the object whose weight is the standard unit of weight. With this in mind, the pound mass has been defined as the mass of the piece of platinum kept in the Standards Office, London, the same piece of metal used to define the pound weight. Scientists have designated the international kilogram as the standard of mass. This standard kilogram is the mass of a certain piece of platinum-iridium preserved in the International Bureau of Weights and Measures at Sèvres, near Paris. One one-thousandth of the mass of this block is the standard gram.

Early writers, such as Galileo and Descartes (1596–1650), not having a clear concept of mass, did not distinguish between the mass and the weight of an object (the distinction is often overlooked even today). The real need for a distinction came in 1671, when it was found that the same object has different weights on different parts of the earth's surface. Mass and weight were clearly perceived as distinct by Newton.

Momentum

Recall the experiment of the athlete putting a shot. So long as he heaved on the same mass, on the 16-pound shot, for example, the accelerations produced were proportional to the forces acting. However, when he used the same muscular exertion on the 16-pound and 12-pound masses, the smaller mass was given the greater acceleration. Thus, although in the first case we could have accurately judged the forces by the accelerations produced, in the second case, if this had been the criterion, we would have judged the

smaller mass to have been heaved with the greater force. The measurement of acceleration alone, then, is not sufficient to determine the force acting. The mass of the body is obviously a very important element in the measurement. Newton was the first to recognize these facts and, in order to present the problem clearly, multiplied the mass of a body by its velocity, calling the new concept "quantity of motion," now known by the single word *momentum.* This new quantity is very helpful in describing the action of moving bodies. For example, we know that it is easier to catch a baseball than a medicine ball falling at the same speed (Fig. 23); and we clearly recognize that the problem of stopping a mosquito and the problem of stopping an automobile, though they are going at the same speed, are very different indeed. These moving objects differ very greatly in mass and, hence, very greatly in momentum, even though possessing the same speed.



Fig. 23. Momentum.

Applying this new concept to the putting of the shot, we find that the rate of change of momentum not only gives an accurate index of the force acting in the first case, where the rate of change of velocity also served well, but gives this index also in the second case, where the rate of change of velocity failed. But as a final check let us test this conclusion with the experimental work of Galileo. In the same locality a ten-pound mass and a five-pound mass fall freely with the same acceleration. Thus the rates of change of velocity are the same. But since the masses are different, the first being twice the second, the rates of change of momentum are not equal. *The first is twice as great as the second. But the force on the first weight is twice that acting on the second. Hence, the rate of change of momentum is directly proportional to the weight and, therefore, may be used to measure force.* This leads to another law enunciated by Newton.

Newton's Second Law of Motion

The rate at which the momentum of a body changes is proportional to the force producing the change and takes place in the direction of the straight line in which the force acts.

This law is well illustrated by the action of the apparatus shown in Fig. 24. The momentum of the "wagon" may be changed rapidly or slowly by a jerk or by a gentle pull. The more rapid the change of momentum, the more the spring stretches and the greater is the force acting. When once the "wagon" is given the desired velocity, the accelerating force (stretch) disappears, and the spring shows only a very small stretching, just enough to overcome friction and thus reduce the resultant force to zero.

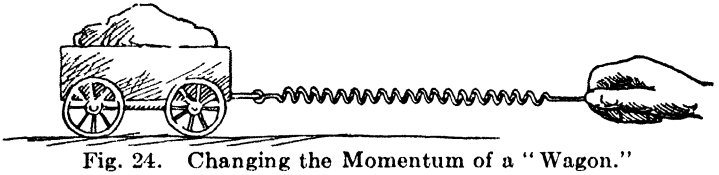


Fig. 24. Changing the Momentum of a "Wagon."

Since the mass of *any given body* is constant, the rate of change of momentum is registered only by the rate of change of velocity, that is, by acceleration. Hence, for *a given body*, the rate of change of momentum is the mass times the acceleration, and in the language of algebra the second law takes this form:

$$F = kma, \tag{3.1}$$

where F is the force, m is the mass, a is the acceleration, and k is a constant which depends on the kind of units of measurement used.

We are now prepared to define the new unit of force, a unit which is the same on all parts of the earth and, hence, independent of the force of gravity, a unit which is freed from the sense of muscular effort. If in the Equation (3.1) the mass m is measured in grams and the acceleration

a in centimeters per second, per second, then we may by agreement put the constant k equal to unity and call the units found by multiplying m and a simply force units or designate them by some special name. The name selected for this force unit is the Greek word *dyne*, meaning *force*. The dyne, then, is the force required to give one gram of mass an acceleration of one centimeter per second, per second. It is a very small quantity. Nearly a half million of them are needed to make a pound force. When the force is measured in dynes, the mass in grams, and the acceleration in centimeters per second, per second, Equation (3.1) takes the simplified form

$$F = ma. \tag{3.2}$$

Since the acceleration under the action of gravity is 980 centimeters per second, per second (the value changes with the locality on the earth's surface), the pull of gravity on a gram of mass (gram force) is 980 dynes.

Gravitational Attraction

We are always in the presence of gravitational attraction, in the presence of that which gives *weight to all objects*. Already we have learned that the weights of a group of objects at a certain locality are proportional to their masses, and that the weight of any one of these changes from place to place. Thus it appears that the gravitational attraction, commonly known as weight, is the joint action of the earth and the particular object under consideration. The mass of the earth and the mass of the object undoubtedly are important factors controlling the mutual attraction. Then, too, distance from the center of the earth may be a factor. So far, we have tacitly assumed that a freely falling body, such as an apple falling from a tree, descends under a *constant* force. Does this mean that an apple, taken in the gondola of a stratosphere balloon to a height of some 73,000 feet, would not change in its weight? A scientist on board could easily find out by the use of a sensitive

spring balance. Newton had no stratosphere balloon in which to reach such high altitudes, but he did have the astronomical observations and conclusions of Kepler, and a genius for seeing broad and general principles. Kepler had set forth three empirical laws describing the motion of the planets about the sun, but had given no explanation of why one might expect these celestial bodies to behave as they do. "Standing on the shoulders of giants," Newton saw, with extraordinary insight, an interpretation of Kepler's laws in terms of his own laws of motion, already discussed, and of a gravitational force described thus: *The gravitational attraction between any two bodies varies directly as the product of their masses and inversely as the square of the distance between their centers of gravity, and acts in the direction of the straight line connecting these centers.*

Thus the weight of an apple may be interpreted as the gravitational attraction between two masses, the apple and the earth, thought of as concentrated at points approximately 4,000 miles apart—the distance from the earth's surface to its center. This means that if an apple could be taken to an elevation of 4,000 miles, its distance from the earth's center thus being doubled, its weight would decrease to one-fourth of its normal value. Yet at the height of man's highest ascent, the weight of an apple is reduced by less than one per cent, and the difference in weights of an apple when hanging on a tree and then when lying in the grass is so very small as to be measurable only by very precise and delicate instruments. In discussing freely falling bodies, therefore, we have been justified in assuming that they fall under a constant force of gravity.

But does an apple attract another apple according to the same law of attraction found for a planet and the sun, and for an apple and the earth? Newton concluded from his profound investigation that gravitational attraction is universal, every particle of matter in the universe attracting every other particle according to the law he found true for the sun, moon, and planets. Henry Cavendish (1731–

1810) initiated the experimental technique used in the measurement of gravitational attraction between objects of ordinary size. Such measurements were brought to high perfection in 1927 by Dr. Paul R. Heyl, of the United States Bureau of Standards. Newton's law of universal gravitation tells just how the attraction depends on mass and distance, but it does not give in dynes the force between two one-gram masses placed with centers of gravity one centimeter apart. The latter is the contribution of Doctor Heyl, and now one may calculate to a high degree of accuracy the attraction between any two bodies by knowing their masses and the distance between their centers of gravity. Such a calculation reveals that the weight of an apple (200 grams) is approximately 60,000,000,000 times the attraction between two such apples placed one foot apart. It is easy to see why such a force is unnoticed in our everyday experiences.

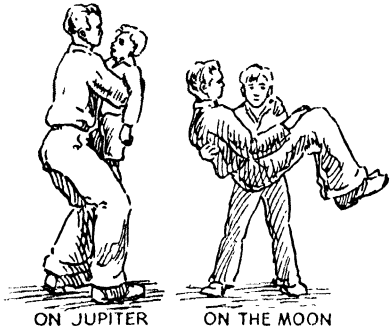


Fig. 25. Weight Due to Surface Gravity.

Applying Newton's law of universal gravitation to the moon, the planets, and the sun, we may calculate the surface gravity, and thus the weight of a given object, on these our celestial neighbors. For example, on the moon a 150-pound man (earth weight) would weigh 26 pounds, an easy lift for a small child (Fig. 25); on Mars this man would weigh 57 pounds (a boy's weight on a man's legs—think of the athletic possibilities!); on Jupiter he would weigh 396 pounds, a great burden to himself; and on the sun (of course, he would burn up first) his skeleton would collapse and his body become crushed under a bodily weight of more than two tons!

Falling bodies and air resistance. We are now in a position to study the effect of air resistance upon falling bodies. When a feather is dropped, it tends to *accelerate*

under the force of gravity. But as its speed increases the air resistance also increases, quadrupling with each doubling of speed. Thus the force of gravity which makes the feather go faster and faster is soon completely overcome by the air resistance, which has been increasing rapidly with the increasing speed. A limiting or terminal velocity is reached when the resultant force becomes zero. Thus we see that under the action of *no* force, the inertia of the feather keeps it traveling with a uniform velocity—an excellent example of the operation of Newton's First Law. As you recall seeing a man jump from an airplane, you will remember that his motion was at first greatly accelerated; but, with the unfolding of the parachute, he soon reached a rather moderate terminal velocity and settled to the ground safely. Had this same man jumped without a parachute from an airplane some two or three miles high, he still would have reached a terminal velocity with the air resistance equal and opposite to the force of gravity. But he would have reached the dangerous speed of more than 100 miles per hour. In an airplane dive from 18,000 to 9,000 feet, the late Jimmy Collins reported having reached a terminal velocity of 395 miles per hour. At this speed, acceleration ceased; gravity was completely overcome by air resistance.

Newton's Third Law of Motion

In our study of balance in the last chapter we noticed that equilibrium is obtained by the action of forces which are always equal and opposite. Newton recognized this pairing off of forces and gave us his third law:

To every action (force) there is always an equal and opposite reaction.

This is illustrated by the action and reaction present when a batter strikes a baseball and by the action and reac-

tion as a sprinter leaps from the mark (Fig. 26). Or, again, this law may be vividly illustrated by the action of a set of collision balls.

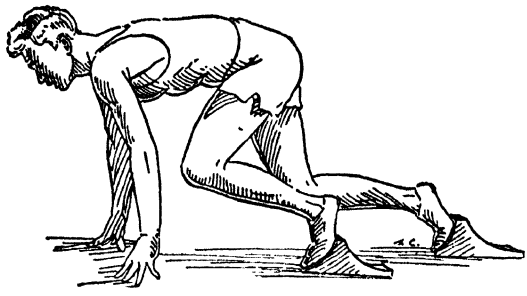


Fig. 26. Starting from the Mark.

Questions and Problems

1. A mean solar day is made up of how many seconds?

2. A car travels 20 miles in half an hour. What is its average speed?

3. A car starts from rest and in one minute is traveling 45 miles per hour. What is its average acceleration (*a*) expressed in miles per hour per minute, (*b*) expressed in miles per hour per second, (*c*) expressed in feet per second per second?

4. A car makes a turn at a uniform speedometer reading. Does the velocity of the car change? Explain. By the use of a drawing explain why it is desirable to bank curves.

5. If a body falls freely from rest, what will be its speed after five seconds of fall? In working the problem, disregard air resistance.

6. Why was the experiment performed by Galileo at the Leaning Tower of Pisa so important?

7. A body will move with constant acceleration if .... is impressed upon it.

8. (*a*) Why is a car traveling at 60 miles per hour wrecked if it strikes a telephone pole?

(*b*) Why is a car traveling at 10 miles per hour only slightly damaged?

(*c*) Why is it that a car traveling 5 miles per hour may not be damaged at all?

9. Molecules travel approximately one-fourth of a mile per second. Why are their impacts on your face not detected?

10. A body remains at rest .... and in uniform motion in a straight line .... unless some .... is applied to it to compel it to change its condition. This is Newton's ..... Law.

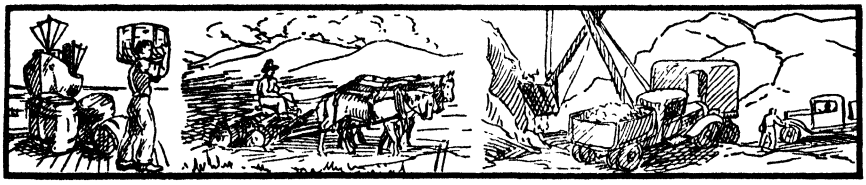
11. Why is it that a person with a massive body and weak legs cannot make rapid changes of motion?

12. A force is measured by the rate of change of .....

13. If a person carelessly jumps from a rowboat, expecting to land on a dock, he may come short of his destination and fall into the water. The explanation is given by Newton's ..... Law.

Suggested Readings

(1) Cajori, Florian, *A History of Physics*, The Macmillan Company, New York, 1914, pp. 31–37.
(2) Dampier-Whetham, Sir William C. D., *A History of Science*, The Macmillan Company, New York, 1931, Chap. IV.
(3) Dietz, David, *The Story of Science*, Sears Publishing Company, New York, 1931, Part I.
(4) Langdon-Davies, John, *Man and His Universe*, Harper and Brothers, New York, 1930, Chaps. III and IV.
(5) Lenard, Philipp, *Great Men of Science*, The Macmillan Company, New York, 1933, pp. 24–39, 67–83.
(6) Northrop, F. S. C., *Science and First Principles*, The Macmillan Company, New York, 1931, Chap. I.



CHAPTER IV

Energy and Work

Utilization of Natural Energy

The energy of the wind was the first to be harnessed. Sails were placed on vessels by the ancient Egyptian mariners and by the Vikings of the North; yet warships driven by the oars of galley slaves were considered by Caesar to be superior to sailing vessels because of the unreliable nature of the winds. During the Middle Ages, windmills were used extensively for grinding grain. But winds may not blow, and if they do they may vary greatly in speed. The available energy, therefore, is uncertain and decidedly variable, since a doubling of the wind velocity makes the available power eight times as great. Although free and inexhaustible, the wind is one of the least economical sources of energy, because, to be economical, energy must be concentrated in form, certain, and reasonably constant in supply.

The energy of falling water has been harnessed by the use of water wheels and turbines. As long as 2,000 years ago water power was used to grind grain. Those early wheels consisted of a set of paddles which dipped into a running stream. The water pushed on the paddles and the wheel turned around. Such low-efficiency current wheels depending on a flowing stream have been superseded by the more efficient wheels and turbines run by falling water, a more concentrated form of energy.

The Nature of Energy

We have just indicated how man has been able to harness natural energy and make it do his work. Was it the air and the water that did the work? Still air and quiet water

cannot do work. Wind, then, is more than air; a flowing stream, more than water. The difference is energy, and *energy manifests itself as a capacity to do work.* Thus, *energy* and *work* are synonymous terms.

How Is Work Measured?

Air pushes against the windmill; the machinery turns, and work is done. Water strikes the paddle of the water wheel; the machinery turns, and work is done. A man lifts dirt from a cellar, and he does work. A woman sweeps a carpet, and she does work. What common elements do these processes have? First, there is the equivalent of a push or a pull, and second, there is motion. Or simply, when a force moves an object from one position to another, work is done.

From our experience in lifting weights, we know that lifting a 50-pound object *two* feet requires just twice the work needed to lift it *one* foot. Also, lifting a 100-pound object one foot requires twice the work needed to lift a 50-pound object the same distance (Fig. 27). Thus, it would appear that work might be measured by multiplying the force acting by the distance moved. But care must be taken to measure the distance in the direction of the acting force.

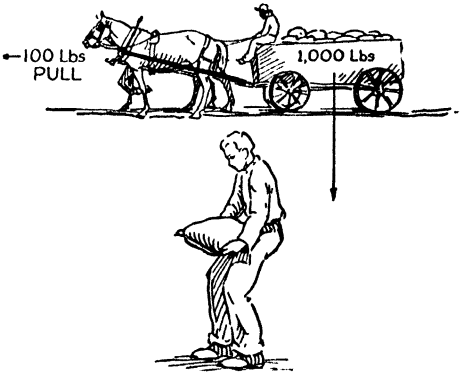


Fig. 27. Doing Work.

If in either of the above cases the weights had been moved in a horizontal direction, there would have been no motion in the direction of the force acting and, therefore, no work done. When a body is slid along a table, the work done is in opposition not to the pull of gravity, but to the forces of friction which oppose the sliding. The same is true in moving a loaded wagon

(Fig. 27). In the illustration, the horses are working against a force of 100 pounds, not against a force of 1,000 pounds, the weight of the wagon and load. Thus, we are led to the scientist's definition of work: When an object is moved by the action of a force, *the work done is equal to the magnitude of the force multiplied by the distance moved along the line of action of the force.*

A unit of work. We are now prepared to define a unit of work. The unit commonly used is called the *foot-pound.* It is the work done when the force of one pound moves a body a distance of one foot in the direction of the straight line in which the force acts. We may now write,

(foot-pounds) = (pounds of force) × (distance measured in feet).

Stated in the abbreviated language of algebra, this is

$$W = Fd \tag{4.1}$$

where W stands for the work, F the force, and d the distance moved in the direction of the straight line in which the force acts. If the force is measured in dynes and the distance in centimeters, the work will be measured in *ergs*, a word derived from the Greek *ergon*, meaning *work*. This is a very small unit of work, more appropriate for application to insects than to men. Therefore, the *joule*, equal to 10,000,000 ergs, is used in practical measurements.

The Measure of How Fast Work Is Done—Power

About 1780, James Watt (1736–1819) improved the steam engine to such an extent that steam power would have been cheaper for pumping water from English coal mines than horse power. But in order to convince the mine operators of this fact, Watt was forced to rate his steam engine in terms of the horses it would replace. Accordingly, he performed this simple experiment. He found that the ordinary draft horse used at the mines to pump water would walk, on the average, two and one-half miles per hour while pulling with a force of 150 pounds,

and that it was able to work at this rate for a full working day. From these data he calculated that this kind of horse could work for a day and deliver each minute of that time 33,000 foot-pounds of work. This has since become the practical unit of the rate of doing work, and is called *one horsepower* (H. P.). Thus, a 20 H. P. engine can do work equivalent to twenty horses of this type. As steam replaced horseflesh, the comparison of engines with horses ceased, but still one engine continued to be checked against another by comparison of the respective horsepower ratings.

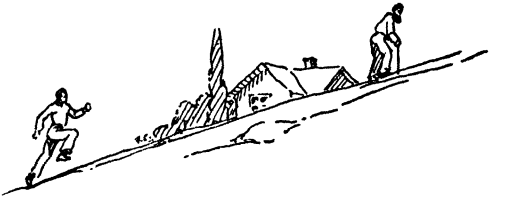


Fig. 28. Work at Different Rates.

A weak man might do as much work as a strong man if given sufficient time. When the youth and the old man, of equal weight (Fig. 28), reach the top of the hill, they will have done approximately the same amount of work. But obviously, the man requires more time than the boy, because, being old, he works at a much slower rate. Very often when one works by the day, the rate of doing work is less than when one works by the job. If this is true, then when working by the job, one manifests a greater power of accomplishment than when working by the day. This *rate of doing work* is called *power* and is measured commonly by the unit defined above, the horsepower. The *foot-pound per second*, meaning a foot-pound of work accomplished in a second, could be used, but the larger unit, the *horsepower* (33,000 foot-pounds per minute, or 550 foot-pounds per second), is more practical. In science the joule per second, called the *watt*, in honor of James Watt, is the unit of power used.

Work from Moving Bodies—The Hammer

A stone used to drive a stake into the ground is indeed a very simple tool. Yet the operation of the ten or more varieties of our present-day hammers is explained by an

analysis of how a stone may be used to pound. The stone is grasped firmly in the hand, lifted a distance above the head of the stake, and forced rapidly downward, gravity helping the change of motion (Fig. 29). As the stone strikes, the stake is forced into the ground, the resistance offered by the earth during the penetration being the force which changes the momentum of the stone. If the ground is hard, the penetration resistance will be large, the momentum change rapid, the time of action short, and the penetration small. When the earth is soft, the penetration resistance will be small, the change of momentum slow, the time of action long, and the distance which the stake penetrates large.

The work done on the stake is equal to the average force acting times the distance the stake is forced into the ground. A larger stone or a greater velocity, or both, would have driven the stake deeper, and more work would have been accomplished. Thus, it appears that the ability to do work, which so far has been measured in terms of the force acting and the distance moved, may also in some manner be related to mass and velocity—a fact which was hinted at in the consideration of the water wheel.

Kinetic Energy

A stone in motion, we have seen, has an ability to do work. This is evidence that when in motion it possesses energy—of a type we call *kinetic*. Energy often manifests itself to the senses as a sort of "vitalized action" of such things as stones, air, water, and other forms of matter. Some think that the early writers used more descriptive words than we now use when they termed kinetic energy "living force" or "*vis viva*." Can we, by knowing the mass of the stone and its velocity, calculate its "*vis viva*" or kinetic energy? Yes—rather simply, as follows: Assume that the stone loses its velocity at a uniform rate, and take v to represent the velocity which it possesses when it first hits the head of the stake. Its final velocity will, of course,

be zero. Because of the assumed uniform change, the average velocity during the drive will be $v/2$. If t is the time required to complete the process, the stick will penetrate the ground a distance $vt/2$, because the average rate of travel times the time of travel gives the distance traveled, as we all know. Thus, we obtain the *distance quantity* needed to calculate the work done.

The other needed quantity, the force acting, is obtained as follows: The momentum of the stone as it strikes is equal to its mass times its velocity, or simply mv. At the end of the operation, this momentum is zero because the impact has brought the stone to rest. The total change in momentum is, therefore, mv. But this change took place uniformly in the time we have called t; hence, the rate of change of momentum, how fast the momentum changed, is simply mv/t. The rate of change of momentum, as we pointed out in the last chapter, is always proportional to the acting force, and equal to it, if proper units are used. If grams are used to measure the stone's mass, centimeters and seconds to measure velocity, seconds to measure time, and dynes to measure force, the equality rather than the proportionality is assured, and we may write the relations just found for the force and the distance as follows:

$$F = \frac{mv}{t}, \tag{4.2}$$

$$d = \frac{vt}{2}. \tag{4.3}$$

We have already shown in this chapter that the formula for work is

$$W = Fd. \tag{4.4}$$

Hence, using the above equations for force and distance, we have at once

$$W = \frac{mv}{t} \times \frac{vt}{2} = \frac{1}{2}mv^2, \tag{4.5}$$

as the work done, measured in ergs, by the stone owing to its kinetic energy. Identifying this work *accomplished* with the kinetic energy *possessed* by the stone by virtue of its motion, we have this important statement: *The kinetic energy (KE) of a body is equal to one-half its mass times the square of its velocity;* or, in the language of algebra,

$$KE = \tfrac{1}{2}mv^2. \tag{4.6}$$

As an aid in simplifying the above argument, were we justified in setting up the ideal condition—that the stone should lose its velocity at a uniform rate? One would expect that the same amount of energy would be given up by the stone in coming to rest quite independently of what brought it to rest and how it changed its speed in losing its energy. But for the sake of argument, let us suppose that one could find a unique way of bringing the stone to rest so that it would give out a maximum of energy, and also another way of bringing it to rest so that it would give out a minimum of energy. If this second method could be reversed by starting with the stone at rest on the head of the stake (Fig. 29) and giving it a velocity by pulling up the stake, there would be a means of giving the stone an initial velocity by expending less energy than it would return in being brought to rest on the head of another stake by the first method. Thus, by giving the stone a velocity after one fashion (*pulling* stakes in a certain manner) and bringing it to rest after another (*driving* stakes in a certain manner), energy could be accumulated, and a machine could be built that would *run* by using the energy it *created!* Men have dreamed of perpetual-motion machines which could do just this thing; they have designed and built devices which at first seemed to give promise; but

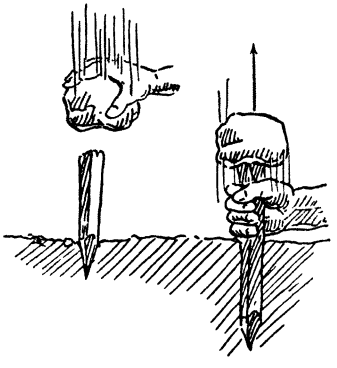


Fig. 29. Driving and Pulling Stakes.

not one attempt has been successful. The failure of mankind to build a perpetual-motion machine, strange as this may seem, gives us the right to assume that the stone discussed above loses its velocity at a uniform rate even though it may actually decrease in velocity in a very complex manner.

Undoubtedly you will wish a summary of this involved argument. To begin with, we questioned the right to use a certain simplification. Then, to make a case, we set up the hypothesis that the use of the simplification was *not* justified, and were led to the possibility of making a machine which could run on the energy it created. Finally, remembering the experience of the race with so-called perpetual-motion machines, we denied the possibility of such a machine, rejected our hypothesis, and thus established the right to use the simplifying assumption. Then, too, our denial of the possibility of producing a machine which can run on the energy it creates has prepared us for one of the greatest generalizations of science.

Conservation of Energy

We are thus led to the principle: *Energy can neither be created nor be destroyed.* This is the "Principle of the Conservation of Energy." Although arising out of man's fruitless attempt to achieve perpetual motion (doing work without the expenditure of energy), today the principle is spoken of by scientists without any particular reference to men or their machines. The comprehensiveness of this generalization will be more apparent as we move through the chapters which follow.

Potential Energy

As a hammer is lifted and brought to rest in the act of pounding, work is done by the artisan; energy is transferred from him to the hammer. At its highest position, the hammer has no motion, and thus no kinetic energy. But it does have the ability to do work. It possesses a kind of

energy called *potential* energy, which is made available for the pounding by being transformed into kinetic energy during the downward stroke. The artisan may let the hammer fall, simply using the stored potential energy; or he may urge it downward, thus adding energy to the hammer. Similarly, a stone resting on the brink of a precipice possesses potential energy (Fig. 30). When the support gives way, the stone falls, and its potential energy is progressively changed into kinetic energy. In mountain climbing, one becomes vividly aware of the fact that the potential energy stored in the human body during the climb and in the supporting stones during the long ages of mountain building is all too willing to be converted into kinetic energy. In general, all forms of energy which cannot be classed as kinetic are called potential. Some of the more common forms are:



Fig. 30. Potential Energy Changing to Kinetic Energy.

1. Potential energy due to position with respect to the ground, such as the hammer in the above illustration, a man on the top of a cliff, or water in an elevated reservoir.

2. Potential energy due to the distortion of an elastic body, such as a springboard just ready to throw the diver upward, a stretched or compressed spring ready to close a screen door as a person passes in or out of a home, or a bent spring in the coil which is ready to set the table of a phonograph in motion.

3. Potential energy possessed by substances that can do work by a chemical reaction, such as the pent-up energy of coal, dynamite, gasoline, or food.

A body may possess at the same time both kinetic and potential energy. The action of a simple pendulum is a good example of this condition. As the bob sweeps past the center position in either direction, its energy is all

kinetic; as it reaches its position of rest on either side, its energy is all potential; at intermediate positions, the energy is part potential and part kinetic, the potential being transformed into kinetic as the speed increases, the kinetic being transformed into potential as the speed decreases. Were it not for the loss of energy due to air resistance and the bending of the string at the point of support, according to the principle of the conservation of energy the pendulum would continue to vibrate indefinitely, the sum of the potential energy and kinetic energy of the bob at any instant being constant.

Machines

The law of machines. The application of the principle of the conservation of energy to machines gives rise to the law: *The work done upon any machine by the agents or forces which operate it is exactly equal to the total work done by the machine against all the forces which oppose this operation.*

The simple lever. Another simple tool used early in the development of the race is the lever or pry, already discussed in Chapter II. As pointed out in the consideration of the hammer, a tool does not *create* energy, but transforms it, being both contrived and used to extend the force which a human being can normally exert. A person drives a stake because he finds he cannot push it in; he moves a hammer through a large distance in order that he may drive a nail through a short distance. Thus, in effect, the force exerted directly on the hammer is magnified into a force large enough to drive the nail. It is often the need for this extending of man's normal force that leads to the use of the lever.

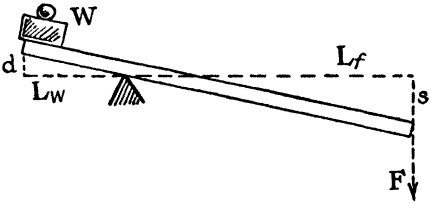


Fig. 31. The Lever.

Fig. 31 shows a lever which has been moved from the horizontal position so that a resisting force W has been lifted a distance d. Neglecting the work done against frictional

forces, such as air resistance and the rubbing of moving parts, and forces due to a shift of the position of the center of gravity of the bar itself, we have, according to the law of machines,

$$Wd = Fs. \tag{4.7}$$

The two triangles (Fig. 31) are similar; hence,

$$\frac{s}{d} = \frac{L_f}{L_w}, \tag{4.8}$$

and at once

$$WL_w = FL_f. \tag{4.9}$$

which we recognize as the law of the lever (Equation 2.2), already shown to hold true for equilibrium.

Let us restate the procedure just finished in order that we may be impressed with the worth of the principle of the conservation of energy in problem solving. First we imagine the bar displaced through a very small distance without the dissipation of energy. This idealized, arbitrary displacement is usually called a *virtual* displacement. If we remember that the law of the lever applies to a condition of equilibrium, to a condition of no motion and, thus, no frictional forces, it is clear to us why we have carefully emphasized that no energy is to be dissipated. Next we take an inventory of the work done due to the virtual displacement. Then to this so-called virtual work we apply the law of machines. Finally, making use of the geometry of the apparatus, we arrive at the law of the lever, which we met first in the laboratory. Thus we have illustrated one of the fascinating aspects of science, the inferring of specific rules from great generalizations. Such inferences should be tested in the laboratory—and this work is just as fascinating.

Weighing machines. The simple beam balance is a form of lever with the arms made accurately the same length. The object to be weighed is placed on one pan and standard weights are placed on the other. When equilibrium is established by the proper selection of weights, the pull of

gravity on the object will be the same as on the weights since the lever arms are equal, and the masses will also be the same. Thus, to find the mass of the object, one needs only to add up the weights.

Other more complex levers. *The windlass.* The force F is applied at the crank handle (Fig. 32). Since only the force component acting perpendicular to the crank is effective in turning the windlass, we shall assume that the workman is careful to exert his force in this effective direction. The weight W is supported by a rope which wraps on the drum. The crank and the drum act as a lever. The force has a lever arm R equal to the length of the crank, and the weight a lever arm r equal to the radius of the drum. At equilibrium the moments balance (or give the windlass a virtual displacement and apply the law of machines, as in the case of the simple lever), and



Fig. 32. The Windlass.

$$FR = Wr. \tag{4.10}$$

Solving,

$$F = \frac{r}{R}W. \tag{4.11}$$

Since the ratio r/R is a fraction less than unity, F is smaller than W, and a person is able by the aid of this machine to lift a weight which he could not move by direct action. However, the acting force F moves through a distance R/r times the distance which the weight is lifted. Thus, there is no violation of the principle of the conservation of energy.

In actual practice, frictional forces, caused by the rubbing of the rope on the drum and the shaft on its bearings, must be opposed. Thus all the work done by the acting force may not be used to lift the resisting weight. For example,

a certain drum has a radius of five inches and the crank a length of fifteen inches. Judging from these dimensions, if no friction were present, F would be one-third W; but actually, 60 pounds instead of 50 pounds are required to lift a load of 150 pounds.

Mechanical advantage. The ratio of the resisting force to the acting force is called the *mechanical advantage* of a machine. Thus, in the example of the windlass just described, the mechanical advantage is actually 2.5; yet, had there been no friction, it would have reached the ideal value of 3. This means that the actual mechanical advantage is always less, often much less, than the ideal mechanical advantage achieved only under the ideal condition of no friction. The ideal mechanical advantage of a windlass is the ratio of the length of the crank to the radius of the drum. The actual mechanical advantage may approach, but never reach, this ideal maximum value— thus, a knowledge of the ideal value is desirable.

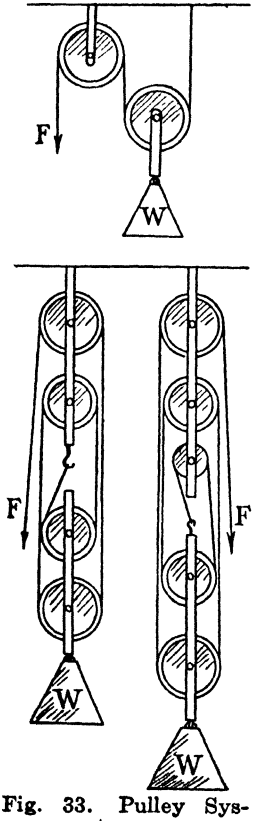


Fig. 33. Pulley Systems.

A combination of pulleys. A single fixed pulley is a form of lever with equal arms. Its main use is in changing the direction of the force. For example, it permits one to pull *down* and yet lift a bucket *up* out of a well. But the acting force moves the same distance as the resisting force, and therefore the former can never be smaller than the latter. Actually, in the case of lifting a bucket of water from a well, the person pulls down on the rope with a force greater than the weight because of the friction between the moving parts.

In the case of the movable pulley shown in Fig. 33, top, the attached weight is lifted only one foot when the rope to which the force is attached is shortened two feet.

Neglecting the work due to frictional forces, we have, by the law of machines, the equation

$$2 \times F = 1 \times W. \tag{4.12}$$

This means that, for the ideal case of no friction, F would be equal to one-half of W. The ideal mechanical advantage is *two*, since W would be twice as large as F.

By similar reasoning, it may be shown that, for the ideal condition of no friction, F is one-fourth of W in Fig. 33, lower left, and one-fifth of W in Fig. 33, lower right. In general, F is $1/n$ of W, where n is the number of ropes supporting W (the weight of the movable pulley block must be included in the weight of W). Under ideal conditions, therefore, W is n times as large as F, and the ideal mechanical advantage of a combination of pulleys is simply equal to n, *the number of ropes supporting the movable pulley block*. In commercial pulleys, the friction is usually so great that the ideal mechanical advantage is never closely approximated.

The inclined plane. A mountainside is essentially a series of inclined planes; the paths up its slopes may be steep or gentle. One raises himself with greater effort, yet more rapidly, along the steep trail; the longer trail is climbed with greater ease and comfort. Why is this the case? It is not because a person is more heavy on a steep trail, but probably because the same speed of travel will be attempted on both paths. For example, on the one, ten steps may mean a lifting of the body through the vertical height of four feet; on the other, ten steps may mean a lifting of the body through a height of two feet. If these ten steps are attempted at the same rate, twice as much work will be done in the first case as will be done during the same time in the second. This accounts, in part at least, for the fatiguing effect of the steep trail; yet, no doubt, the average person climbs more efficiently when the action is more nearly like the regular process of walking.

Mountain roads also are composed of series of inclined planes. A person with driving experience knows that a steep grade cannot be taken with the same speed as a gentle one. Some grades are impossible! A study of the action of a car on a simple inclined plane (Fig. 34) should aid in the interpretation of mountain driving. Let us call the weight of the car W and the force exerted through the action of the engine, transmission, etc., F. Then, under the ideal condition of no friction, the force F can work through the distance s and lift the weight W (the car) through a vertical height d. According to the law of machines,

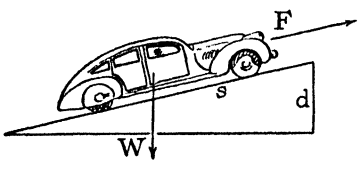


Fig. 34. The Inclined Plane.

$$Fs = Wd. \tag{4.13}$$

Solving,

$$F = \frac{d}{s}W. \tag{4.14}$$

Thus, neglecting friction, the acting force is the d/s part of the resisting force. This fraction is always less than unity, since the height of an inclined plane is shorter than its length. The ideal mechanical advantage, on the other hand, is always greater than unity, being equal to the length of the plane divided by its height, or simply s/d. If a road rises one foot in 100 feet of road distance traveled, the ideal mechanical advantage is 100 and the force needed to propel the car is $\frac{1}{100}$ part of the car's weight plus the force needed to overcome friction. If the road rises two feet in this distance, the required force (not including friction) doubles, and so on. An engine has a certain horsepower, which means that it can do work no faster than a certain rate. If small forces are to be overcome, the engine has the power to move the car in a given time farther than if large forces are opposed. Hence, a hill must be taken at a slower speed than a road with an easier grade. If one could be content to climb two feet in twice the time required

to climb one foot, then the power requirement would be the same. On steep hills it is just as easy to overwork cars as men! Permit your car to work efficiently by shifting to a lower gear even before it labors under the load.

A keen-edged tool is a very interesting inclined plane. When man learned to chip flint so that it would have a keen edge, he hastened the appearance of many tools so necessary in the building of a civilization. A knife with a thin blade and a keen edge may be easily forced through wood, leather, and flesh. By the use of a mallet, a stronger, blunter edge can be used to cut metals and chip stone. Tin snips and scissors are examples of the combination of levers and keen edges.

Friction

In the study just made of machines, we have noted that friction always enters as a hindrance because the work expended in overcoming it cannot be used for useful purposes. To minimize friction in machinery, therefore, is an important task. Yet friction is not always a hindrance. Were it not for friction, we could not walk, run, pick up objects, or start and stop trains and automobiles. Recall your experience on an icy road, or your attempt to pick up a very slippery fish, and you will be impressed with the value of friction.

Sliding friction. There is always friction when two surfaces rub. If these surfaces are magnified, they will be seen to have a hill-and-dale appearance (Fig. 35a). As a matter of fact, friction is caused by the interlocking of these small irregularities. *Approximate* rules governing sliding friction may be summarized thus:

(1) Friction is nearly independent of velocity. Yet, starting friction is greater than sliding friction, as one may discover by shoving a heavy box across a floor; and it decreases somewhat with increasing speed, causing the friction of automobile brakes to be less at 60 miles per hour than at 20 miles per hour.

(2) If surfaces are dry, the friction does not depend much on the area of surface contact, as may be verified by pulling a brick across a table with its three different-sized surfaces in turn making contact with the table top.

(3) Lubrication greatly decreases friction, as may be discovered by placing soap or paraffin on a drawer that sticks, paraffin on skis, or oil on dry bearings. But with well lubricated surfaces, the friction is nearly proportional to the area of contact—very different indeed from the action of dry surfaces.

(4) The force needed to overcome friction, though varying greatly with the condition and kind of surface, is nearly proportional to the total force which presses the sliding surfaces together. Thus it is twice as difficult to shove a box across a floor if its weight is doubled. The constant ratio between the frictional force and the pressing force is called the *coefficient of friction* for these particular surfaces.

Rolling friction. When man learned to put his vehicles on wheels, he greatly reduced the friction involved in moving them, because rolling over the small irregularities which cause friction is easier than sliding over them (Fig. 35c). In deep snow, a bobsleigh (Fig. 35b) is easier to pull than a wagon, but this is far from being true on ground not covered with snow. The common wheel bearing of the wagon has been superseded by the roller-bearing in the wheel of the automobile, and thus sliding friction is replaced by rolling friction.

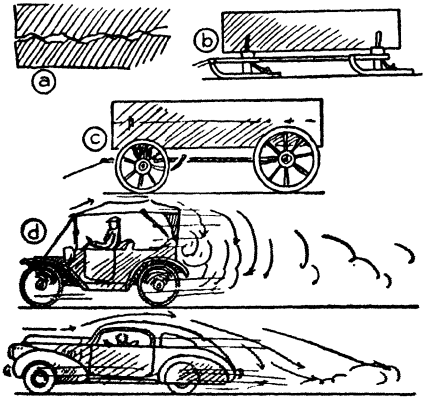


Fig. 35. Reducing Friction.

Air resistance. At high speeds, air resistance is by far the greatest resistance which an automobile or train encounters. This resistance is not independent of the velocity, as in the case of sliding friction; the power needed

to overcome it is nearly proportional to the cube of the speed. The doubling of the speed means an eight-times increase of power. This explains why it is so expensive to run a train or automobile at a high speed.

Friction reduction. The most effective means of overcoming friction, therefore, is found in the proper selection of bearing materials, lubrication, the replacing of sliding by rolling contacts, and the reduction of air resistance by proper streamline design. Air turbulence always follows in the wake of the old-style automobile (Fig. 35d). Energy is wasted by this vigorous churning of the air, and resistance is thus increased. The modern car, with a proper streamline design, slips through the air with little churning effect. Thus the air resistance is reduced.

The Efficiency of a Machine

Thus it becomes clear that not all the energy put into a machine can be used by it in doing useful work. Some of this supply must be used to overcome friction and is wasted. The efficiency of the machine is therefore the ratio of the useful work *done by it* to the total work *done upon it.* Thus, in the windlass described above, the *useful* work done in lifting the 150-pound weight one foot is simply 150 foot-pounds, while the *total* work (3 ft. $\times$ 60 lbs.) done upon the machine is 180 foot-pounds. The efficiency of this machine is $\frac{5}{6}$, or 83.3 per cent.

What has become of this wasted energy? Why is the nature of things such that this leak is ever a menace to the operator of machines? These questions will be answered in the next chapters.

Questions and Problems

1. When we find a body that can do work, we have evidence that it possesses .....

2. Work is defined as the ..... of the force ..... by the ..... along the line of action .....

3. Power is the .....

4. If your weight is 150 pounds and you climb a stairway with a vertical height of 15 feet, how much work do you do against gravity? One day the climb is made in 10 seconds; another day in 15 seconds. Contrast the work and the power of these performances.

5. Energy can neither be ....

6. Assume that you can exert a force of only 75 pounds. Describe how you might successfully lift a 150-pound weight a distance of 1 foot.

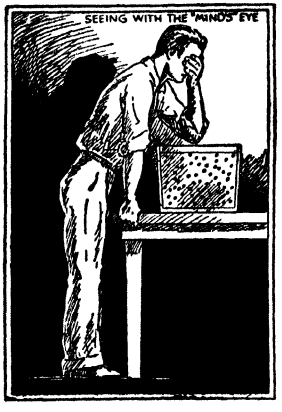
7. How might the Egyptians have elevated the heavy stones found at the top of the pyramids?

8. Explain why oil is an important requirement of an automobile motor.

9. List the reasons why present-day automobiles are stream-lined.

Suggested Readings

(1) Crew, Henry, *General Physics*, The Macmillan Company, 1927, pp. 136–149.
(2) Farnham, *et al.*, *Profitable Science in Industry*, The Macmillan Company, 1925, Chaps. IX and X.
(3) Lenard, Philipp, *Great Men of Science*, The Macmillan Company, 1933, pp. 126–136.



CHAPTER V

Too Small to Be Seen

The Atoms of Which Things Are Made

When we observe stones, sand, dust, and drops of water, do we see what the world is made of? On division, stones may become sand, and sand, dust. But what of dust, when the division is endlessly continued? On division, drops of water may become spray, and spray mist. But what happens when a mist disappears?

Two thousand years ago Titus Lucretius (98–55 B. C.) wrote his treatise *De Rerum Natura*, in which he attempts to explain these phenomena by assuming that air, earth, and water—in fact, all things—are composed of tiny bodies, singly too small to be seen. Continued subdivision, according to this view, would lead at last to atoms, minute corpuscles which cannot be cut. This view makes the attempt to study the invisible a fascinating adventure. On the other hand, suppose that a substance such as water were homogeneous in the truest sense. Then a continued division would reveal no new properties; the minutest portion would still show the same characteristics as the largest portion; the things of human size would be but an enlarged replica of the things of the microcosm; there would be nothing worth searching for beyond the visible. Fortunately for those who are fired with the spirit of adventure, the picture given by Lucretius turns out to be nearer the truth.

At the beginning of the nineteenth century, John Dalton (1766–1844) formulated the atomic theory in its classical form. It is based upon the experimental facts summarized as follows:

(*a*) *The law of the conservation of matter.* A candle burns and is consumed; yet, if we weigh the candle, the oxygen it uses in burning, and the products of combustion, the chemical balance shows that nothing has been lost. The matter has changed form, but it has not been annihilated. A careful check on many other chemical changes leads to this conclusion: Matter changes its form, but not its total amount, during chemical reactions.

(*b*) *The law of definite proportions.* Pure water, no matter what its source, can always be decomposed into oxygen and hydrogen. Any sample always produces a weight of oxygen 7.94 times the weight of the hydrogen obtained. This illustrates the law of definite proportions.

(*c*) *The law of multiple proportions.* Hydrogen peroxide (an antiseptic) is another compound of oxygen and hydrogen. On being decomposed, results show that there is always a weight of oxygen 15.87 times the weight of hydrogen. Hence, we see that a given amount of hydrogen unites with *exactly twice* as much oxygen to form hydrogen peroxide as it does to form water. This illustrates the law of multiple proportions (Fig. 36).

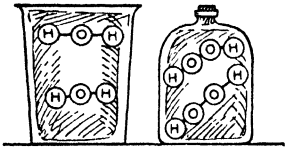


Fig. 36. Multiple Proportions.

Dalton saw that he could interpret these facts of science by accepting the hypothesis that *atoms* are the elementary constituents of the so-called elements, the ultimate particles of a given element all having the same definite mass but singly being so tiny as to be far beyond the visibility obtained in the most powerful microscope. These atoms unite in definite patterns to form the molecules of substances. The study of how this happens is the major province of chemistry.

Since the time of Dalton, this hypothesis has been successfully checked time and time again, and scientists are agreed that the modern atomic theory gives a rather reliable picture of the microcosmic reality. This modern theory

requires the atom, originally meaning something which cannot be cut, to take on a new aspect and become a configuration of the primordial electrons and protons, the negative and positive corpuscles of electricity—a sort of planetary system with electrons (the number characterizes the chemical element) revolving about a positively charged nucleus, which in turn is made up of protons and neutral particles called *neutrons*. This is a dependable, yet oversimplified, picture, the latest "model" of modern science being still too complicated for the beginner. Modern theory still makes use of the *atoms* of Dalton; but through the help of laboratory techniques unknown to him, certain of the so-called chemical elements are found to be made up of different-weight atoms (*isotopes*), presumably the result of atoms with the same number of "planetary" electrons (of the same chemical element) taking into their nuclei a slightly different number of neutrons—thus adding to atomic weight without a change in chemical properties.

A further discussion of the structure of atoms is given in Chapter XVII. We shall now study the behavior of atom-groups, atoms bound together in the tiny, invisible (even with the very best microscope) units of substances called *molecules*. Because of the historical development of their study, the structure of atoms and the action of molecules are topics included in the study of physics.

The Motion of Molecules

If we can find evidence that molecules move, then we may make a preliminary guess that the energy of machines which seemed to be lost owing to friction may have been taken on as kinetic energy by these invisible corpuscles (Fig. 37). If

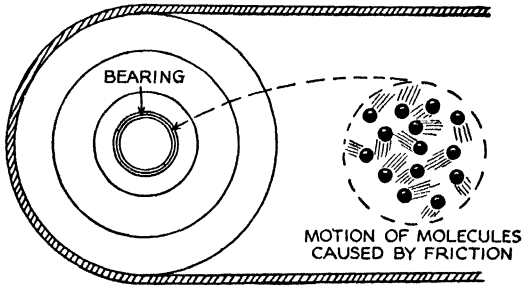


Fig. 37. Friction and Motion of Molecules.

this is true, then the astonishing thing is, not how much mechanical energy is rendered useless by friction, but rather how little.

Diffusion of gases. If a little ether, ammonia, chlorine, or any volatile substance of powerful odor is introduced into a room, in a very short time the scent may be detected in all parts of the enclosure. A part of this movement is ordinarily caused by air currents, but if precautions are taken to avoid such drafts, the mixing will still take place rather rapidly.

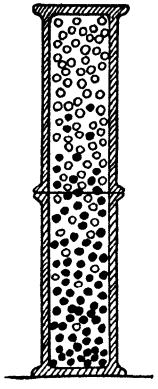


Fig. 38. Diffusion.

Let a small rubber balloon be filled with illuminating gas. It rises to the ceiling. This is evidence that the gas is lighter than air, just as it is known that a floating block is lighter than water. Let a jar of illuminating gas be placed over a jar of air. Even though the gas is lighter than the air, after a time the gas is found mixed with the air and the air with the gas, as is evidenced by the fact that *both* jars contain an explosive mixture. This diffusion certainly supports the idea that the molecules of a gas are in motion.

Continuing the investigation, let a porous cup of unglazed earthenware be closed with a rubber stopper through which passes a glass tube (Fig. 39). By means of a support, let the open end of the tube dip into a dish of colored water, and let a glass jar be placed over the cup. By means of a rubber tube connected to a gas jet, let illuminating gas fill the jar. Bubbles pass out through the colored water, showing that the pressure inside the porous cup has increased. Now let the jar be removed, so that the gas on the outside of the cup may escape. The rise of water in the tube indicates that the pressure inside the cup is decreasing. Finally the water drops, and the pressure is the same inside and outside the cup. How can these facts

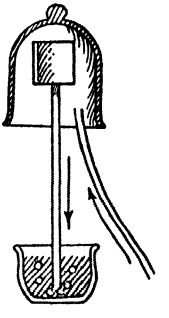


Fig. 39. The Diffusion of Gas Through a Porous Cup.

be explained? We must first discover the cause of the pressure which a gas exerts.

Gas pressure. If a rubber membrane is stretched over a glass jar and the air is exhausted by the aid of a pump, the membrane will be more and more depressed until finally it is broken by the outside air pressure (Fig. 40). If the lips of Magdeburg hemispheres are placed in contact, the halves may easily be pulled apart, but if while together the air is exhausted from between them, they may be pulled apart only with great effort. These experiments demonstrate that air exerts pressure. When lemonade is sucked through a straw, the air pressure is made less on the inside of the mouth than on the outside, and the liquid is forced up the straw.

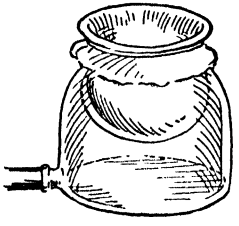


Fig. 40. Air Pressure.

Is the air limited in the total pressure it can exert? If so, how can this maximum pressure be measured?

Evangelista Torricelli (1608–1647), a friend of Galileo, sealed one end of a four-foot glass tube and completely filled it with mercury. He closed the open end with his thumb, inverted the tube, and immersed the end in a dish of mercury (Fig. 41). On removing the thumb, he observed the mercury fall from the upper end, leaving a so-called vacuum there. But the surface *did not fall lower than about 30 inches. This experiment revealed two important facts.* First, nature does not abhor a vacuum, as the Greeks and Romans had taught as an explanation of the drawing of liquid through a straw. Second, the air has a measurable maximum pressure. Air pressure, it seems, is the cause of the phenomena formerly attributed to a mysterious sucking power of a vacuum. Blaise Pascal (1623–1662), desiring further evidence, carried a "Torricelli tube" to the top of a church

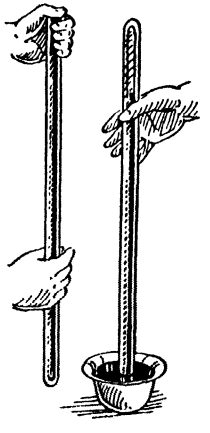


Fig. 41. A Mercury Barometer.

steeple. He argued that, if the instrument actually measured the air pressure, it should indicate less pressure at higher altitudes. He detected a slight drop of the mercury, "but desiring more decisive results he wrote to his brother-in-law to try the experiment on the Puy de Dome," a high mountain in the south of France. The brother-in-law observed a difference of three inches in the height of the mercury and wrote back that he was "ravished with admiration and astonishment." Pascal concluded "that the vacuum is not impossible in nature, and that she does not shun it with so great a horror as many imagine."

The modern mercurial barometer is essentially a "Torricelli tube." It measures the atmospheric pressure in inches or centimeters of mercury—which means that the air has sufficient pressure to hold a mercury column that high. (One *atmosphere* is taken equal to 76 centimeters of mercury.)

The *aneroid barometer* is used regularly by geologists and aviators to determine differences of altitude by observing differences of barometric pressure (Fig. 42). It contains a hollow, airtight metal box in which there is a partial vacuum. An increase in atmospheric pressure causes the top of the box to bend in; a decrease of pressure causes it to spring out. By a system of levers this motion is recorded by a pointer which moves back and forth over the face of a dial calibrated to read air pressure in inches of mercury and, often, the elevation in feet above sea level. Some of these instruments are so sensitive as to be able to record a change in pressure when moved from a table to the floor.

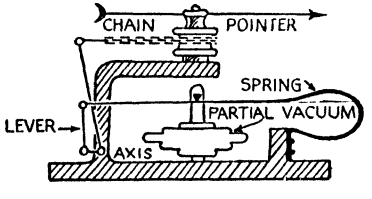
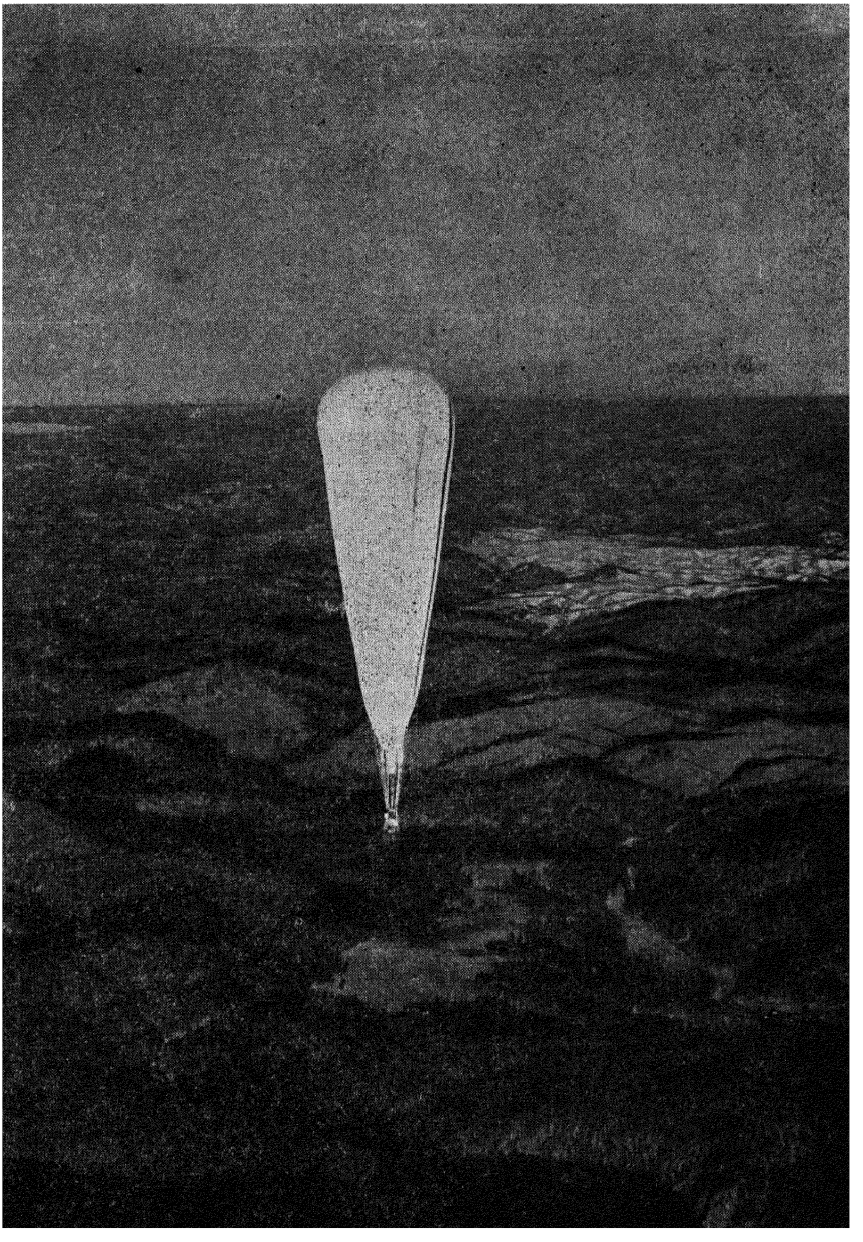


Fig. 42. An Aneroid Barometer.

Boyle's law. In the so-called sucking process, the decrease of pressure in the mouth is brought about, not by discarding air, but by increasing the volume which it occupies. Inhaling during breathing depends also on this



©*N. G. S.*

Photograph by Master Sergeant G. B. Gilbert and Captain H. K. Baisley.

PLATE II. One of the official photographs of The National Geographic Society-United States Army Air Corps Stratosphere Balloon "Explorer II," as it rose above the Black Hills of South Dakota November 11, 1935, to a record height of 72,395 feet above sea level. Initially the bag, composed of 2⅔ acres of fabric, was just partly filled with helium, but it became distended and spherical as the outside air pressure diminished with the increasing altitude. The gondola carried more than a ton of scientific instruments which were in charge of Captain Albert W. Stevens. Captain Orvil A. Anderson was pilot. (*Photograph reproduced by special permission of the National Geographic Magazine.*)

process. About 1662 Robert Boyle (1627–1691) found that when he kept the temperature of a gas constant but allowed it to expand into twice its original volume, the pressure was halved. He also discovered that when it was crowded into half its original volume, its pressure was doubled. He found this inverse relationship between volume and pressure when he tripled, quadrupled, or changed the volume in any manner.

Boyle's law may be stated thus: *If the temperature is kept constant, the volume of a given mass of gas varies inversely as the pressure.* The law is often written in algebra as follows:

$$P_1V_1 = P_2V_2 = \text{a constant (temperature and mass constant)}, \tag{5.1}$$

where P_1 and V_1 represent the pressure and volume at one time and P_2 and V_2 the quantities at another time. As an example, suppose that an automobile tire is pumped to three atmospheres (approximately 45 pounds per square inch); the air is crowded into one-third the space it previously occupied as free air. The work needed in this compression is performed by a compression pump installed at the service station. Boyle's law does not hold for very high pressures or very low temperatures but works well under average conditions.

Causes of gas pressure. Any preliminary guess as to what causes gas pressure must account for Boyle's law, the pressure noted in the porous cup experiment, and the diffusion of gases. Diffusion has been interpreted by picturing gases made up of *molecules in motion.* This may be our cue. If for any reason the velocity of a molecule is changed, as, for example, when the molecule strikes a rigid surface (Fig. 43), forces which might be the source of gas pressure will arise. Let us see.

If a molecule hits a wall and stops and starts off again, it exerts a force because its momentum is changed. The total pressure, the force on a square centimeter, is the sum of all the forces that act on this unit area. If the molecules

are crowded into half the original space, they will be twice as close, on the average; there will be twice as many to strike the square centimeter; and, if the average speed is constant, the pressure will be doubled (Fig. 43). If the gas is allowed to occupy double its volume, the molecules will be twice as far apart, on the average; there will be only half as many to strike the square centimeter; and if the average speed is constant, the pressure will be halved. This agrees with Boyle's experiments. Then, too, it seems that Newton's laws of motion, which hold for objects of human size, also operate in the realm of the infinitely small. Gas pressure, then, whether on the surface of a lemonade or in an automobile tire, is due to molecular bombardment.

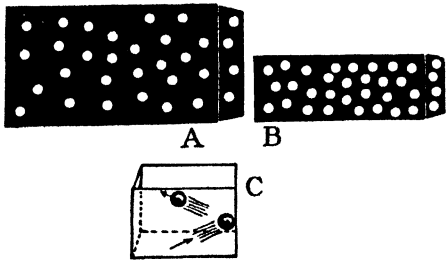


Fig. 43. Illustrating Change of Pressure with Change of Volume.

We shall now make a more careful analysis illustrating how mathematics may be used in the development of a scientific theory. A cubic centimeter of hydrogen at the temperature of melting ice and weighing 0.000089 grams is found to exert a pressure of 76 centimeters of mercury, while a cubic centimeter of oxygen at the same temperature and weighing 0.00143 grams exerts the same pressure. How is it that the hydrogen can exert the same pressure as the oxygen, which weighs 16 times as much?

A simple calculation will answer this question. Suppose that there are in the cubic centimeter of hydrogen N molecules, each of which has a definite mass, M. Let C be an average speed which will properly represent the many different speeds which no doubt are present, owing to the haphazard jostling about of the molecules. When one of these little bodies hits the wall of the cubic centimeter container, it rebounds with the same speed as the speed with which it strikes. First, it comes to rest, losing MC momentum; then it gains the same amount as it bounds off with

the same speed. The wall had first the task of stopping the molecule and then that of giving it momentum, the total change of momentum being $2MC$. The force which the hydrogen molecule exerts upon one side of the cube is equal to the change of momentum it experiences at that particular surface during one second. A molecule may hit more than once in a second. On the average, it needs only to go *across* the cube and *back* before it hits again. But the distance of this trip is 2 centimeters, and the speed of travel is C; hence, the number of times that C can be divided by 2 gives the number of round trips, and thus the number of impacts, during a second. The change of momentum per second at the surface due to a single molecule, that is, the force exerted by a single molecule as it bombards, is, then,

$$2MC \times \frac{C}{2} = MC^2. \tag{5.2}$$

Not all molecules may be expected to contribute their bombardment to a particular surface; the forward-and-back motion which we have described could take place equally well in three directions. Therefore, we take $\frac{1}{3}N$ as this side's share of molecules and get for the force on the square centimeter, that is, the pressure,

$$P = \tfrac{1}{3}NMC^2. \tag{5.3}$$

If for oxygen we use the corresponding small letters, we get for its pressure,

$$p = \tfrac{1}{3}nmc^2. \tag{5.4}$$

Now the product of the mass of one molecule times the number of molecules per cubic centimeter is the density; hence, for hydrogen $NM = 0.000089$ grams per cubic centimeter, and for oxygen $nm = 0.00143$ grams per cubic centimeter. Since the two gases are exerting the same pressure, $p = P$, and we can write,

$$\tfrac{1}{3}NMC^2 = \tfrac{1}{3}nmc^2,$$

or

$$0.000089C^2 = 0.00143c^2,$$

and

$$\frac{C^2}{c^2} = \frac{0.00143}{0.000089} = 16.$$

Extracting the square root,

$$\frac{C}{c} = 4.$$

This means that the hydrogen molecule moves with four times the speed of the oxygen molecule when the gases are at the same temperature and pressure. In general, it means that the molecules of a less dense gas, under the same conditions of temperature and pressure, always move faster than those of a more dense gas, the squares of the speeds varying inversely with the densities.

By the use of the above formula for pressure and the substitution of the proper values for density and pressure, the average speed (the root mean square speed) of a molecule of oxygen at room temperature is calculated to be about 1,000 miles per hour, and that of hydrogen about 4,000 miles per hour!

Explanation of diffusion through porous cup. Now reread the details of the experiment performed with the porous cup. Illuminating gas is less dense than air, which means that its molecules must move faster than air molecules when the gases are at the same temperature and pressure. Therefore, the illuminating gas molecules strike the porous cup oftener than do the air molecules, and in one second more of them pass into the cup than air molecules pass out. This clearly is the cause of the increase of inside pressure. When the jar is removed and air replaces the gas which surrounded the outside of the cup, the gas molecules which previously came in through the porous wall now pass out faster than air molecules come in. This causes the decrease in pressure. As a further verification of this

explanation, in the same experiment let us substitute compressed air for illuminating gas. No change in pressure is observed. This result agrees perfectly with the explanation we have just made.

Molecular Magnitudes

So far we have pointed out the experimental basis for the guess that all things are made up of molecules and atoms, that the molecules of gases are in motion with very high speeds, and that gas pressure is due to molecular bombardment. The data we have submitted would not warrant the calling of this guess an hypothesis, perhaps, but many experiments, which we cannot record for lack of space, and long, careful mathematical studies, which are too advanced to be here included, give so much support to this preliminary guess and others as to justify the building of the so-called "kinetic" theory—a well-established theory of the action of molecules.

By the use of experimental data and the methods of the kinetic theory, we are able to determine the following molecular magnitudes:

(*a*) Air molecules are striking our faces with velocities of about one-fourth of a mile per second—velocities exceeding the speed of bullets. We are thankful that molecular momentum and not speed causes damage to the face. (The smallness of mass is the cause of the smallness of momentum.)

(*b*) We cannot conceive the smallness of the mass of an air molecule (oxygen 5.3×10^{-23} grams, nitrogen 4.6×10^{-23} grams). The mass of a baseball is certainly very small as compared with the mass of the earth; the mass of the air molecule is just as small as compared with the baseball. When an air molecule strikes our face, we take no more notice of it than does mother earth when her cheek is struck with a baseball.

(*c*) The diameter of an air molecule (oxygen) is 3×10^{-8} centimeters. Particles must have a diameter a thousand

times larger than this to be visible under the most powerful microscope. If one could imagine an enormous enlargement that would make it possible for a man to hold the sun in his hand, as he does normally a large orange, and if he could then shrink up again to normal size, he would experience the same degree of shrinking necessary to become the size of a molecule. Or again, if a thimble full of air were enlarged to the size of the earth, the molecules would have the size of basketballs—although not the same shape, as we shall discuss in a later chapter.

(*d*) There are about two billion persons on earth. If each of these were able to count 100 air molecules per minute, it would take the entire population three hundred thousand years to count the number of air molecules that go into a quart bottle as the milk is poured out! If placed side by side in a single row, there would be a billion molecules in every foot—a lifetime would be needed to count a row a foot long!

(*e*) Yet, in a gas the molecules are not crowded. They move, on the average, 300 times their diameter between impacts. They have as much room as 50 basketballs would have in an average-sized auditorium 100 feet by 50 feet by 15 feet.

Liquids and Solids

Liquids may partly fill a container; but gases, no matter how rarefied, make use of the whole volume. Liquids are also denser than gases, a fact suggesting that their molecules are much closer together. Then, too, these molecules in large numbers cling together in drops, giving evidence of a molecular attraction not manifest in a gas under ordinary conditions. In consideration of these differences, are we justified in assuming that the molecules of a liquid are in motion?

If a tall jar is nearly filled with water colored with blue litmus (or some other acid indicator), made slightly alkaline with ammonium hydroxide, and a little concentrated

sulphuric acid is very carefully introduced into the bottom by means of a pipette, a pink layer will appear where the acid comes in contact with the litmus (Fig. 44). Sulphuric acid is 1.8 times as heavy as water; yet, after a few hours, the layer of contact will have moved some distance up the tube, showing that the acid molecules gradually found their way up into the water. This *diffusion* is evidence that liquid molecules are in motion.

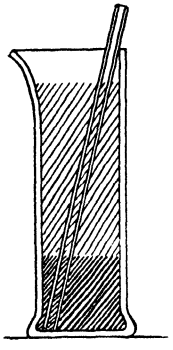


Fig. 44. Diffusion in Liquids.

A piece of ice, quartz, or steel is held together with strong cohesive forces—the presence of large molecular attractions is thus indicated. All solid things are either crystalline or amorphous. Most substances tend to form into crystalline structures as they pass from the liquid to the solid state. Often, under unfavorable conditions, the crystals are so small as to be invisible to the naked eye. However, an amorphous substance shows no crystallizing tendency as it passes from a liquid to a solid form; in fact, it often becomes difficult to determine whether it should be called a very viscous liquid or a soft and pliable solid. An amorphous substance, such as glass, sealing wax, or asphalt is, then, not very different in structure from a liquid. But crystalline substances such as ice and snow have a definite, orderly structure not characteristic of the liquid state. Your experience of seeing ice form in spectacular designs on cold window panes will convince you of this fact; or you may try this interesting experiment: Select a clear slab of ice, one-quarter of an inch thick, from the surface of a pond and place it as you would a slide in a projection lantern. The heat from the light rays causes the ice to melt and fantastic shapes (ice crystals) appear outlined on the screen. A careful study of snowflakes also reveals interesting crystalline forms (Fig. 45).

The permanency of crystal structures is evidence that molecules, if in motion at all, are restricted in the main to limited regions of activity. Marshaled in an orderly

array within a crystal, they probably vibrate back and forth about a fixed position. Yet there seems to be a very slight diffusion of such molecules. Sir Roberts-Austin,

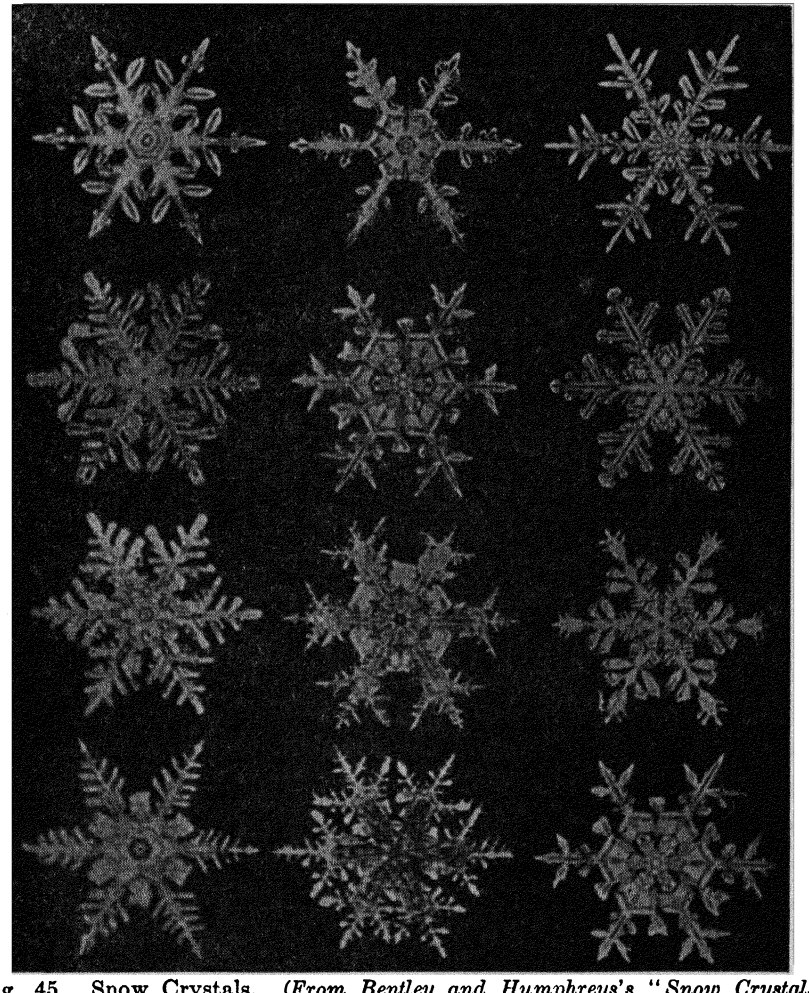


Fig. 45. Snow Crystals. (*From Bentley and Humphreys's "Snow Crystals." Courtesy of McGraw-Hill Book Company, Inc., New York.*)

an English scientist, clamped a block of lead and a block of gold firmly together so that their surfaces made intimate contact, and let them remain thus for a considerable length of time. Analysis revealed that molecules of each metal had very slowly diffused into the other metal.

Thus, we are led to the conclusion that all things are made up of molecules, which seem ever to be on the move. Then, too, we have observed how man has been able to see beyond the visible by using the "mind's eye." But we still have not answered this question: How may the average speed of molecules be increased, and how does this increase affect our senses? You will be interested in finding the answer in the next chapter.

Questions and Problems

1. Justify the statement, "The molecule is but a symbol standing for a particular mosaic of measurements and inferences evolved in the scientific use of apparatus and formalized reflective thinking," making use of the measurements and inferences introduced in this chapter.

2. List all the evidences you can which support the kinetic theory of gases.

3. From your experiences cite evidences, if you can find any, against the kinetic theory of gases.

4. If the temperature is kept constant, the volume of a given mass of gas varies .....

5. List the advantages and disadvantages of the mathematical analysis used in the discussion "Causes of gas pressure."

6. What viewpoint should one take toward the statement "All things are made up of atoms and molecules, and these tiny bodies are in constant motion"?

Suggested Readings

(1) Fisk, D. M., *Exploring the Upper Atmosphere*, Oxford University Press, New York, 1934.
(2) Jaffe, B., *Crucibles*, Simon and Schuster, New York, 1930, Chap. VII, "Dalton."
(3) Langdon-Davies, John, *Man and His Universe*, Harper and Brothers, New York, 1930, Chap. V.
(4) Lenard, Philipp, *Great Men of Science*, The Macmillan Company, New York, 1933, pp. 48–51, 62–66, 176–184.



CHAPTER VI

The Nature of Heat

Is Heat a Fluid?

By direct observation we come to know that the sun warms the earth, that fires are hot, and that boiling water is hotter than melting ice. We discover early in our experience that cold objects may be warmed by bringing them in contact with hot objects; that fires may be built in houses to warm them; and that ice may be put in a refrigerator to cool food. Simple tests show that the conduction to the skin of what we call "heat" arouses a sense datum of warmth, and its conduction away in excess of a certain rate, a sense datum of cold. Persons often speak of "heat" coming from a stove and "cold" from ice, not having clearly in mind, perhaps, the fact that the heat comes from the stove and goes to the ice.

From these and other common experiences, one is impressed with the "flow" of heat. This fact and the method of thinking in the eighteenth century led to the preliminary guess that heat was a fluid. But since a body was found to weigh the same when hot and when cold, the fluid had to become one of the imponderable fluids which were the fashion in science in that century of materialism. This imponderable was *caloric*, a derivation from the Latin word *calor*, meaning *heat*. All things were supposed to have this fluid soaked up in their pores, a hot body containing more than a cold body. The heat produced by squeezing, hammering, and rubbing (Fig. 46)—the heating effect of friction—was accounted for by assuming that these mechanical processes would bring some of the fluid

to the surface, where, although it could not be seen, it might be recognized by the sense of warmth. Heat conduction was simply the flow of this "caloric" from the hot to the cold body.

As the evidence that all things are made up of moving atoms and molecules increased, the so-called "caloric" theory was adjusted to the new facts. The rubbing caused the molecules to vibrate, this motion increased the effective space occupied by the particles, and the caloric was forced out. It was an easy step to the modern theory of heat: just to recognize that kinetic energy of molecules is heat energy and eliminate the caloric. But there were many who held to the hypothesis of the imponderable caloric, and why not? It had been of great service in helping scientists to co-ordinate facts.



Fig. 46. Fire by Friction.

In 1798 Benjamin Thompson, Count Rumford (1753–1814), published the results of experiments on the boring of a cannon. He was surprised at the large amount of heat generated. He placed a water jacket around the brass cannon and used a blunt steel drill. "At the end of two hours and thirty minutes the water actually boiled!" he reported, and he concluded: "It is hardly necessary to add that anything which any insulated body, or system of bodies, can continue to furnish *without* limitation cannot possibly be a *material substance;* and it appears to me extremely difficult, if not quite impossible, to form any distinct idea of anything capable of being excited and communicated in the manner in which heat was excited and communicated in these experiments, except it be *motion.*"

The following year, Sir Humphry Davy (1778–1829) used a clockwork to rub two pieces of ice together in a

vacuum. Part of the ice was melted, although the chamber was kept below the freezing point. The evidence seemed conclusive, but the idea was so new that it was not until 13 years later that he felt sure that he was justified in stating that "the immediate cause of the phenomenon of heat is motion, and the laws of its communication are precisely the same as the laws of the communication of motion."

The experiments of Rumford and Davy normally would have rendered a death blow to the caloric theory, but the materialistic explanation was hard to kill, and it prevailed through the first quarter of the nineteenth century. Even so, these experiments mark the birth of our modern theory of heat which has been so successful in correlating heat and work and in extending the principle of the conservation of energy to heat problems. Again we see that hypotheses and theories, unlike scientific facts, are changing things and are important only insofar as they are useful in helping to interpret the facts in a simple, complete, and suggestive manner.

Temperature and Its Measurement

Galileo was probably the first to use an instrument to measure temperature. Even today, most of us depend to a very large extent on certain dermal senses to determine how hot or how cold an object is. If it feels hot, its temperature is high; if it feels cold, its temperature is low. Although very convenient and useful and exceedingly important as a protection, these sense data are not impersonal or accurate. We must look, therefore, into the outer world for measurable changes which go hand in hand with the changes in these sense experiences.

Simple air thermometer. At the University of Padua, Italy, in 1592 Galileo constructed, as far as we know, the first thermometer. He had observed that most solids, liquids, and gases expand as their temperature increases, the phenomenon being very noticeable for gases. Accord-

ingly, he confined air in a bulb, as shown in Fig. 47, and detected its expansion and contraction with a change in temperature by observing the liquid fall and rise. This instrument has serious limitations. To measure temperature as a function of air expansion or contraction, all other influencing factors must be either eliminated or held constant—this is very important and characterizes a scientific measurement. The simple air thermometer has at least two uncontrolled factors, *viz.*, air pressure and changes in liquid temperature, which affect the thermometer readings. Although this thermometer marks the introduction of measurements into the study of heat, it is crude and inaccurate and was soon superseded by more precise instruments.

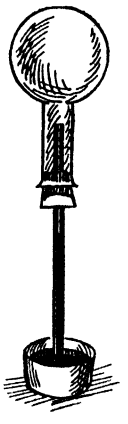


Fig. 47. Simple Air Thermometer.

Mercury thermometer. A mercury thermometer is constructed by blowing a thin-walled bulb on the end of a thick-walled glass tube of very small, *uniform* bore, in diameter often less than that of a human hair. The bulb and tube are completely filled with mercury at a temperature slightly above the highest temperature for which the instrument is to be used, and the tube is sealed off. As the thermometer cools, the mercury contracts, leaving a near-vacuum above it. To make it a measuring instrument, it is necessary to calibrate the thermometer.

Calibration of a thermometer. Certain processes, such as freezing and melting under standardized conditions, are always found to take place at the same temperature. For example, under one standard atmosphere, ice melts and pure water boils at fixed temperatures. One may discover the validity of these statements even with an uncalibrated thermometer as follows: Let the uncalibrated thermometer be placed in a bath of ice and pure water and the mercury level marked on the stem. Let the position be noted in another similar bath. No change is detected. Let the

same experiment be carried out with boiling water. Similar results are obtained (Fig. 48).

Let the stem be scratched deeply at the marks recorded when the thermometer is in melting ice and boiling water. These scratches may be known as the temperatures at which ice freezes and water boils under standard conditions and may be designated by certain numbers agreed upon.

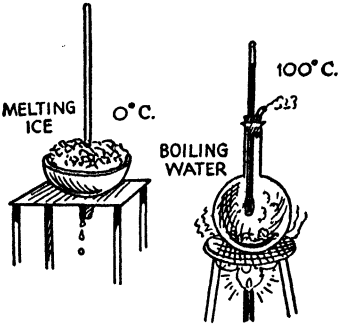


Fig. 48. Calibrating a Thermometer.

The Fahrenheit scale. Gabriel Daniel Fahrenheit (1686–1736) was very successful in making mercury thermometers by reason of his ability to clean mercury. He called the water freezing temperature 32°, the temperature of the mouth of a healthy person 96°, and the lowest temperature he could obtain by mixing ice, water, and sal-ammoniac zero. On this scale the boiling point of water happened to come at 212°. He later used this temperature instead of the temperature of the human body as a fixed point. Thus, there are 180 Fahrenheit degree divisions between the melting point of ice and the boiling point of water.

The centigrade scale. This scale is characterized, as its name implies, by having 100 divisions between the two fixed points, those of melting ice and boiling water under one standard atmosphere. Andreas Celsius (1701–1744) called the melting ice point 100° and the boiling water point 0°, but eight years later Maerten Stroemer, a colleague of Celsius, made the inversion of the scale which places the melting point at 0° and the boiling point at 100° (Fig. 48). Scientific measurements are usually made with a centigrade thermometer. Since mercury freezes at −39° centigrade, thermometers used for very low temperatures contain colored alcohol, which has a freezing point of −117° centigrade.

How to change from one temperature scale to the other. The same instrument could be given a Fahrenheit and a centi-

grade scale, as is illustrated in Fig. 49. But often one wishes to read temperature in centigrade degrees on a Fahrenheit thermometer and *vice versa*. Because of this it becomes necessary to know how to change from one scale to the other. We have already pointed out that the change in temperature from melting ice to boiling water is represented by 180 Fahrenheit degrees and 100 centigrade degrees. This means that there are $\frac{9}{5}$ as many Fahrenheit degrees as centigrade degrees in any given temperature change. Were it not for the fact that the zero on the centigrade scale corresponds to 32° on the Fahrenheit scale, the change from one scale to the other would involve simply the multiplication by $\frac{9}{5}$ or $\frac{5}{9}$. But because of this fact, the change is made by the following rules: First, to change centigrade degrees to Fahrenheit degrees, multiply the temperature reading by $\frac{9}{5}$ and add 32; second, to change Fahrenheit degrees to centigrade degrees subtract 32 from the temperature reading and then multiply by $\frac{5}{9}$.

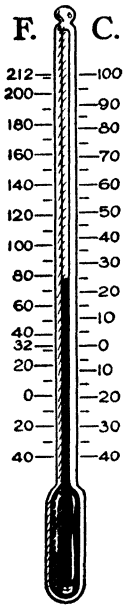


Fig. 49. Fahrenheit and Centigrade Scales Compared.

Thermostats. A commercial thermostat is often used to maintain a constant, predetermined temperature in an oven, incubator, or home. Its essential element is usually a coil made up of a thin strip of brass securely welded to a strip of invar. With an increase in temperature, brass expands more than most metals, while invar expands but little. Thus, the strip bends in one direction when heated and in the other when cooled (Fig. 50). The coil winds up or unwinds with a change of temperature; and as it does

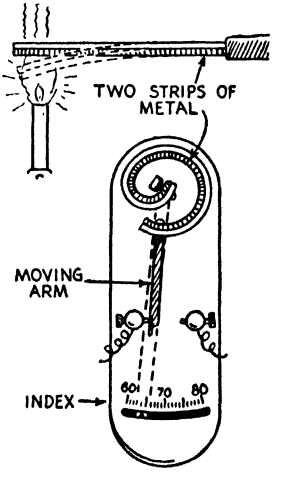


Fig. 50. Thermostat Used in Homes.

so, an electric circuit is opened or closed and the source of heat energy (electricity, gas, oil, or coal) is cut off or turned on. Such a thermostat, used to regulate the temperature of a home, is shown in Fig. 50. The index is set on the desired temperature. When the temperature of the room falls below this value, the moving arm touches the contact on the right and fuel is turned on. As the house temperature rises, the coil winds up and the arm moves to the left, finally touching the contact and turning off the fuel. In this manner the temperature is maintained approximately constant.

What is it that temperature measures? An object seems hot or cold to us, and we judge its temperature. We see a thread of mercury go up or down the stem of a thermometer to a certain mark, and we read off the temperature. What is it that temperature measures?

In 1787, Jacques A. C. Charles (1746–1823) made a careful study of how a *given amount* of gas kept at the *same volume* would change its pressure as the temperature was changed (Fig. 51). It is this increase of pressure with temperature that often causes "blow-outs" on a hot road when the automobile tires have been filled on a cool morning with air at too great a pressure.

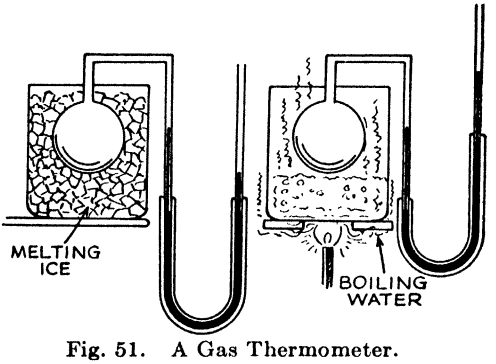


Fig. 51. A Gas Thermometer.

Charles discovered that the pressure increased a definite amount for each degree of rise in temperature. In fact, taking the gas pressure at 0° centigrade as the base of comparison, he found a pressure change of $\frac{1}{273}$ of this reference pressure for each degree centigrade change above or below 0° centigrade. Thus, decreasing the temperature 10° centigrade would mean a drop of pressure equal to $\frac{10}{273}$ of the 0° centigrade pressure; a decrease of 100° centi-

grade would mean a drop of pressure equal to $\frac{100}{373}$ of the 0° centigrade pressure; a decrease of 273° centigrade would mean a drop of pressure equal to $\frac{273}{273}$ of the 0° centigrade pressure—there would be no pressure left! Here we have a temperature, −273° centigrade, at which there is no pressure! But pressure is due to molecular bombardment. It must mean that the pounding has ceased, that the decrease in temperature has been accompanied by a decrease in molecular motion. Does it mean that temperature in some way measures molecular activity? In any event, this −273° centigrade seems a logical temperature to make the absolute zero, the real bottom of all temperatures. The absolute scale—often called the Kelvin scale in honor of Lord Kelvin, Sir William Thomson (1824–1907)—is built out of centigrade degrees but uses this lowest of temperatures as its zero. On this scale ice melts at 273° Kelvin and water boils at 373° Kelvin (Fig. 52).

Charles's law. The experimental results obtained by Charles are proof of the following law: *If temperature is measured on the Kelvin scale, a quantity of gas kept at a constant volume will exert a pressure directly proportional to its temperature.* This means that if the Kelvin temperature is halved, the pressure is halved; if the Kelvin temperature is doubled, the pressure is doubled; and so on.

Temperature the measure of the mean kinetic energy of a molecule. With the aid of mathematics let us make use of Charles's law and investigate the possible relation between temperature and molecular activity. In the last chapter, the pressure of gas was found to be

$$p = \tfrac{1}{3}nmc^2, \tag{6.1}$$

where p is the pressure, n the number of molecules per cc., m the mass of each molecule, and c an average molecular speed. The equation may be rewritten thus,

$$p = \tfrac{2}{3}n(\tfrac{1}{2}mc^2). \tag{6.2}$$

If the number of molecules per unit volume is kept constant,

which was true in Charles's experiment because he kept the quantity of gas and the volume constant, we may write,

$$p = k(\tfrac{1}{2}mc^2), \tag{6.3}$$

showing that the pressure is directly proportional to the average kinetic energy of a molecule. But Charles found that the pressure of a gas is also directly proportional to the Kelvin temperature. The pressure, then, is directly proportional to two things: the mean kinetic energy of a gas molecule and the Kelvin temperature of that gas. This must mean that *the Kelvin temperature of a gas is directly proportional to the mean kinetic energy of a representative molecule of that gas.*

The molecule we have been considering is an idealized thing. It belongs to a gas that obeys Boyle's and Charles's laws *perfectly.* It must not attract its fellows; the space it occupies must be infinitesimally small as compared with the volume it is free to move in. We introduced the ideal molecule, first, when we set up the ideal conditions necessary in the calculation of the formula for pressure, the volumes of the molecules and their mutual attractions being omitted. Second, we assumed that a gas continues to act the same at very low temperatures as it does at room temperature, and on this basis defined the absolute zero and built up the new temperature scale. The introduction of the idealized molecule may seem at first to be a kind of evasion. But science often proceeds in this manner. It sets up the ideal case, solves the problem thus greatly simplified, and then modifies the ideal solution to meet the actual situation. It is done in the case at hand. The Kelvin temperature of an *ideal* gas is found to be an accurate index of the mean kinetic energy of one of its representative molecules. When a study is made of *real* molecules—an investigation too long and difficult for the purposes of this text—we find that *the Kelvin temperature is still a true index of the kinetic energy of a representative molecule* of a gas, liquid, or solid, provided only that the kinetic energy of

translation (moving about), in the case of gases and liquids, and the energy of vibration (except at very low temperatures), in the case of solids, is the only energy taken into account.

There are no real gas molecules just like the molecules of an ideal gas, but helium, the gas used to inflate large dirigibles, is made up of molecules that meet the specifications very nearly. Its molecules consist of one atom each. They show so little attraction for each other that the gas must be cooled down to the exceedingly low temperature of 4° Kelvin before it changes to the liquid form under atmospheric pressure. Through rapid evaporation induced by the action of a huge pump—a process which strips the liquid of its fastest molecules—liquid helium has been cooled down to 0.7° Kelvin. Except for having two atoms to the molecule, hydrogen, oxygen, and nitrogen show the characteristics of an ideal gas until low temperatures or high pressures are reached. As these gases are made to approach the liquid state by a process which compresses and cools them, their molecules are crowded together and stripped of so much kinetic energy that molecular attractions are definitely revealed. The temperatures at which these gases become liquids at atmospheric pressure are shown in Fig. 52. It is easy to see why liquid air (a mixture of nitrogen and oxygen) boils as vigorously on a block of ice as water does on a red-hot stove.

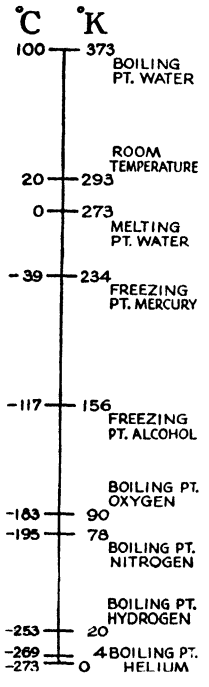


Fig. 52. Kelvin and Centigrade Scales Compared.

By chilling certain chemical salts down to the temperature of liquid helium while they are under the influence of a very strong magnetic field, then insulating them from their surroundings, and finally removing the magnetic field, temperatures as low as 0.015° Kelvin have been reached. Such man-made "cold spots" are probably not equalled

in the universe, because in interstellar space a material body does not very likely get colder than 2° or 3° Kelvin.

The standard thermometer. The gas thermometer, built after the pattern of the apparatus used by Charles (Fig. 51) and filled with an ideal gas, is the *standard* thermometer, the thermometer with which all other thermometers must be compared directly or indirectly if they are to give standard readings. Gas thermometers, using real gases such as helium, hydrogen, and nitrogen, are our very best approach to the ideal standard; and, although such instruments need slight correction at extreme temperatures because a real gas fails to act exactly as an ideal one, all official temperatures made on the many kinds of thermometers must show a pedigree extending back to measurements made by competent scientists on helium, hydrogen, or nitrogen thermometers. Temperatures ranging from a few degrees above absolute zero to 1,600° centigrade may be measured on gas thermometers, the upper limit being reached, not because the gas fails, but because at such temperatures the container (often of platinum) cannot keep the gas from diffusing through its walls. Helium and hydrogen are used for the lowest temperatures and nitrogen for the highest. Temperatures outside the range of gas thermometers are inferred from those within the range by the use of well-established scientific principles. Gas thermometers are inconvenient and, except as standards, are seldom used. Of the practical type, mercury and alcohol thermometers are well known. To these we may add the less known platinum resistance thermometers, thermocouples, optical pyrometers, radiation pyrometers, and "magnetic" thermometers—a formidable array of scientific instruments which you will find described in more advanced texts.

The Measurement of Heat

If one could take an inventory of the kinetic energy of all the molecules of an object before and after a change in temperature, he would be able to calculate the amount of

heat energy absorbed or given out during the temperature change. Even if possible, this would be a stupendous task. We must turn to a practical procedure. We know from experience that a bucket of water and a pond of water do not require the same amount of heat in order to change one degree in temperature. Yet, no doubt, equal masses of water would require the same amount of heat for such a temperature change. Thus, although a thermometer measures changes in temperature, it cannot alone measure how much heat an object takes in or gives out as it warms up or cools down. The kind of material in the object and the mass of the object, as well as the change in temperature, must be known. Thus, a standard unit which is to measure quantity of heat must involve a certain kind of substance, a definite mass of that substance, and a definite change in temperature.

The calorie. Scientists have agreed that *if one gram of water changes its temperature one degree on the centigrade scale, the heat absorbed or given out shall be designated as the* **calorie.** If the change is from 3.5° centigrade to 4.5° centigrade, the unit is called the *small calorie;* if from 14.5° centigrade to 15.5° centigrade, it is called the *normal calorie.* The *mean calorie* is $\frac{1}{100}$ the quantity of heat needed to change one gram of water from 0° centigrade to 100° centigrade. The *large calorie* is equal to 1,000 small calories. These are definitions which are needed in very accurate work. We shall use the simple definition given first as being sufficiently accurate for our purposes.

In the English system, the *British thermal unit* (B. T. U.) is the name given to the amount of heat absorbed or given out by one pound of water when its temperature changes one degree on the Fahrenheit scale (more exactly from 39° Fahrenheit to 40° Fahrenheit). One British thermal unit is equal to 252 calories.

Specific heat. By the definition of the calorie, one gram of water requires one calorie of heat to change its temperature one degree centigrade. But when the quantities of

heat required to change the temperature of one gram of other substances through one degree centigrade are measured, varying amounts of heat are found to be necessary—amounts usually much less than one calorie. This may be demonstrated by placing at the same instant equal weights of water and iron, both at the temperature of the room, in contact with gas burners of the same intensity (Fig. 53). After a few minutes, the iron will be "burning" hot, while the water will be but slightly warmed. *The quantity of heat measured in calories required to raise the temperature of one gram of a substance one degree centigrade is known as the* ***specific heat*** *of that substance.* Thus, to calculate the number of calories of heat absorbed by an iron object, we simply multiply the mass of the object measured in grams by the specific heat of iron and then multiply this product by the increase in temperature measured in centigrade degrees.

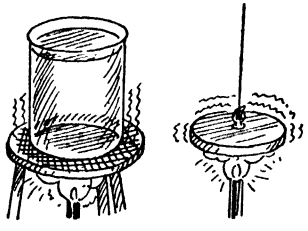


Fig. 53. Comparing the Specific Heats of Water and Iron.

TABLE OF SPECIFIC HEATS

Aluminum	0.218	Water	1.00
Brass	0.094	Ice	0.504
Iron	0.113	Water vapor	0.421 (100°C., 76 cm. Hg)
Mercury	0.033	Air	0.237 (const. pressure)
Clay, dry	0.22		

Heat and Work

The equivalence of heat and work. Let the hands be rubbed briskly. The increase in temperature is a manifestation that heat has been supplied. The work of rubbing has been converted into heat energy. When a piece of lead is pounded very vigorously with a hammer (Fig. 54), the kinetic energy of the hammer disappears, but in our "mind's eye" we see the lead molecules taking on this energy. We touch the metal and discover that this is so, because we detect an increase in temperature, and this sense datum, we have found, can be used to measure

roughly the change in the mean kinetic energy of molecules. Is there an equivalence between the mechanical energy of the hammer and the heat energy gained by the lead? Can we extend the principle of the conservation of energy so as to include heat energy along with mechanical energy? The experiments of Rumford and Davy indicated that heat is a mode of motion, but these experimenters did not make a careful check on the amount of work used and the amount of heat gained. These experiments do point the way, however, and we may guess that the conservation principle is true; but in science, guesses must be supported by facts.

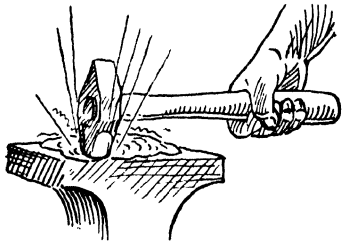


Fig. 54. Mechanical Energy Changed to Heat Energy.

Joule's experimental work. James Prescott Joule (1818–1889) experimented on the mechanical equivalent of heat for about forty years. The work done by falling weights was used to churn water by the action of paddles, the process being very much like the action in an ice-cream freezer (Fig. 55). The work done by the weights and the heat produced were very carefully measured. This permits the accurate determination of the number of ergs of work expended in the production of one calorie of heat—the so-called mechanical equivalent of heat. In other experiments and in search for conclusive proof that heat energy

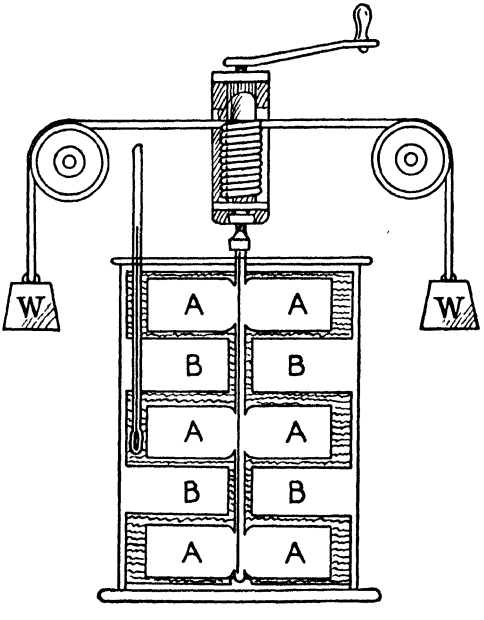


Fig. 55. Joule's Apparatus.

and mechanical energy are related in an exact manner, Joule compressed air (at constant temperature) to 22 atmospheres and measured the work done on the gas and the heat given out. Then he allowed it to expand (at constant temperature) and measured the work done by the gas and the heat absorbed. He rubbed cast iron wheels together under mercury and measured the work and heat; he forced water through small tubes and determined the work done and the heat produced; he measured the electric energy used in heating a coil of wire and determined the heat given out.

In 1879 Henry A. Rowland (1848–1901) repeated the "paddle friction in water" experiment, but used more sensitive and accurate thermometers and took into consideration that the specific heat of water is not quite constant. All these experiments as well as subsequent experiments by D'Arsonval, Miculescu, Griffiths, and others give conclusive proof that whenever mechanical energy changes to heat energy, or whenever heat energy changes to mechanical energy, there is always an exact equivalence. One calorie of heat energy always appears when 41,830,000 ergs of mechanical energy disappear, and vice versa.

Thus, the principle of the conservation of energy has been definitely broadened to include heat energy as well as mechanical energy. The waste of energy noted in the study of machines is now clearly shown not to be an annihilation of energy but rather a transfer of mechanical energy into heat energy.

How may we get work from heat? Mechanical energy is more valuable, as prices go, than heat energy. Only rarely do you find a machine constructed to convert mechanical energy into heat energy by friction. On the contrary, all designers and operators of machines are constantly struggling to keep the valuable mechanical energy from passing into the less valuable heat energy by that ever open door, friction.

Many machines are designed to convert heat energy into mechanical energy. The steam engine and the gasoline

engine are examples. A careful study of heat engines, too extensive and involved for the purpose of this text, shows definitely that, if a quantity of heat is to be converted into work, it must be taken into the machine at a high temperature, usually as hot steam or hot products of combustion, and then *part of it must be given back* at a lower temperature (the temperature of the exhaust), unless this temperature is at absolute zero. Only if the exhaust could be at the temperature of absolute zero would it be possible, therefore, for a quantity of heat to be converted completely into work. Because of the impracticability of having an exhaust at such a temperature, we may never expect an engine with 100 per cent efficiency. To give some notion of how much of the energy of coal or gasoline may be converted into useful work and where the energy actually disappears, the following table is submitted:

STEAM ENGINE		GASOLINE ENGINE	
Useful work	15 per cent	Useful work	25 per cent
Friction	5 " "	Friction	10 " "
Exhaust	45 " "	Exhaust	30 " "
Up the chimney	35 " "	Jacket	35 " "

The human machine. Food is the source of bodily energy (Fig. 56). To determine the total amount of energy that may be derived from a definite quantity of food, we need only to burn it up and measure the amount of heat liberated. This we accomplish by placing a known weight of foodstuff, along with a sufficient amount of compressed oxygen, in a very strongly built metal container called a *bomb calorimeter*. The calorimeter is immersed in a definite weight of water at a known temperature. The combustion is started with an electric spark and the rise of tem-



Fig. 56. Food, the Source of Bodily Energy.

perature of the water is noted. With due allowance for the heat used to warm the calorimeter, the heat given out by the process of combustion may be calculated from the energy known to have been taken on by the water during its rise in temperature. From such experiments it is found that,

1 gram of protein furnishes, on an average	5.6	large calories of energy
1 gram of carbohydrate furnishes	4.1	" " " "
1 gram of fat furnishes	9.3	" " " "

Will these foods if burned in the body yield the same heat energy as when burned in the bomb calorimeter? The answer may be found by placing a man in a calorimeter similar to that shown diagrammatically in Fig. 57. The living quarters are a small room just large enough to accommodate the man and the necessary furnishings, such as a bed, a chair, a table, and a bicycle by which the mechanical energy of pedaling may be converted by friction into heat energy. Another somewhat larger room completely incloses the first, and the temperature at the walls on either side of the dead air space is kept constant by the aid, if necessary, of a heating coil mounted on the outside room. Thus, no heat will leave the inside room through the dead air space, and perfect insulation, so necessary in this experiment because long times are involved, is assured. By the action of a pump, the air of the chamber is circulated, carbon dioxide and water are removed by the proper chemicals, and the supply of oxygen is replenished. The temperature of the room, which would normally rise in such a complete enclosure, is kept constant by radiator pipes through which flows a stream of cold water. Knowing the temperature of the water at intake and outlet and the rate of water flow, one may at once calculate the heat taken from the chamber. Through specially designed

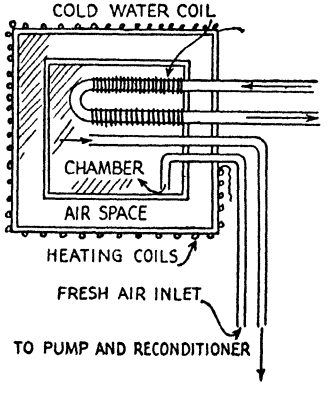


Fig. 57. Respiration Calorimeter.

doors, food is introduced and excreta removed. The latter is burned in a bomb calorimeter and its small energy content is determined. All the energy in the form of food which passes into the chamber in a given length of time is carefully listed. In a similar manner all the energy in the form of uneaten food, excreta, and warmed water which leaves the chamber in the same length of time is properly recorded. At the end of such an experiment it is found that, within the limits of experimental error, the energy which goes into the chamber is equal to the energy which comes out of it.

Here we have evidence that the principle of the conservation of energy applies to the human machine just as truly as it does to other machines. We find no hidden source of energy in the human machine itself. What energy this machine uses it must take in as the pent-up potential energy of foodstuff, and the energy thus made available is that measured by the rapidly burning process of the bomb calorimeter. The following table gives the energy requirements of the average man for a period of 24 hours.

Lowest possible to keep alive (basal metabolism)	1,700	large	calories
Lying down at rest at a comfortable temperature	2,600	"	"
At sedentary tasks	3,000	"	"
At really hard physical work	7,000	"	"

According to tests by Atwater and Benedict, mental work of the most difficult type does not increase the metabolism in any detectable amount and thus requires no extra energy over that listed above. Is it not surprising that, from the standpoint of energy, the monumental achievements of the creative mind have issued from merely the crumbs of a bounteous table?

If the intake of energy into the human machine exceeds the expenditure, storage must of necessity take place. Often this energy is tucked away as surplus fat. If a person plans to remove this surplus by vigorous exercise, how much work will be required to remove one pound of fat? Suppose that as a heat engine the human body is

10 per cent efficient, and that therefore only one-tenth of the energy of the fat is used to lift weights, manipulate hand tools, or turn machinery, the other portion being converted into bodily heat. The energy in a pound of fat is equivalent to 13,000,000 foot-pounds. One-tenth of this is 1,300,000 foot-pounds. This is equivalent to the work done in lifting 65 tons of coal through a distance of 10 feet, or, if you weigh 130 pounds, the work needed to lift yourself a distance of 10,000 feet! A hint to the wise is sufficient. Balance your intake against your expenditure. It is the safest and surest way to attain a graceful figure.

The sun the source of energy. For ages the sun has been supplying the earth's surface with a *total energy during each minute* sufficient to run all the machines in the United States for two or three years. It is estimated that each square centimeter placed at right angles to the sun's rays and located at the outer limits of the earth's atmosphere receives 1.94 calories of heat each minute. This is indeed a very small part of the total radiated energy, since the sun sends out two billion times the energy received by the whole surface of the earth. What is the source of the sun's vast supply of energy? Modern physics suggests that the sun is actually destroying its substance and radiating it as light and heat. According to this hypothesis, each gram of matter when annihilated furnishes 900,000,000,000,000,000,000 (nine hundred million million million) ergs of energy. Sir James Jeans estimates that the mass of the sun one million million years ago was 7 per cent greater than today, the energy radiated during these many years having come from the destruction of approximately 7 per cent of the sun's mass. This point of view would abolish the



Fig. 58. The Sun, Our Source of Energy.

conservation of mass and the conservation of energy as separate principles and substitute in their stead the conservation of a *single entity* which plays the role of both matter and energy. The shift from matter to energy, or from energy to matter, is always made at the exchange rate of nine hundred million million million ergs per gram. When a mass is given energy, as, for example, when a car is put into motion, its mass increases according to this exchange rate. An automobile, starting from rest and reaching a velocity of 300 miles per hour (a record speed), increases in mass by one part in ten million million, an increase *far beneath* our ability to detect. Then, when we remember that mass is not annihilated under the conditions found on the earth, it becomes clear why two conservation principles are so often used instead of just one.

By means of water turbines and windmills, man regularly makes use of energy derived indirectly from the sun, but he has done little to harness solar energy directly and apply it to his machines. True, this radiant energy has warmed him and given him light; but plants have within themselves the ability to take this energy and store it away. Man depends upon this potential energy pent-up in plants for his food. The energy of fuels, such as coal, oil, and gasoline, has come from the sun through the agency of the plant life of the remote past. Practically all known available energy has its origin in the sun (Fig. 58).

Questions and Problems

1. When a person speaks of "zero weather," just what does he mean?

2. How would you design a mercury thermometer so that it would register accurately to tenths of a degree?

3. Why does liquid air boil on a block of ice?

4. When one gram of water changes its temperature one degree centigrade, the heat absorbed or given out is designated as the .....

5. The quantity of heat measured in calories required to raise the temperature of ..... of a substance one degree centigrade is known as the ..... of that substance.

6. List the steps in an experiment in which you could test the proposition: Aluminum has approximately twice the specific heat of iron.

7. The principle of the conservation of energy may be extended to include ..... as well as mechanical energy.

8. What type of energy transfer takes place because of friction?

9. If heat is to be converted into work, it must be taken into the engine at a high temperature and part of it .... at a ....

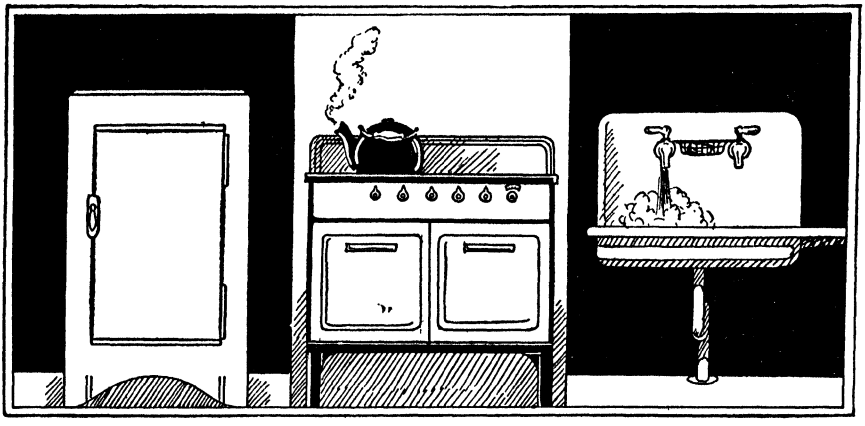
10. Why is it consistent to speak of the human body as the "human machine"?

11. Name a simple and safe plan which a person might follow in order to reduce his weight.

12. Justify the statement "Practically all known available energy has its origin in the sun."

Suggested Readings

(1) Dampier-Whetham, *A History of Science*, The Macmillan Company, New York, 1931, pp. 221–231 and 245–257.
(2) Farnham, *et al.*, *Profitable Science in Industry*, The Macmillan Company, New York, 1925, Chap. VIII.
(3) Jeans, Sir James, *The Universe Around Us*, The Macmillan Company, New York, 1934, pp. 195–212.
(4) Parsons, T. R., *The Materials of Life*, W. W. Norton and Company, Inc., New York, 1930, Chaps. V, VI, and VII.



CHAPTER VII

Physics of the Household

Water Supply

The use of water enters into many of the details of everyday living. A supply of pure water is certainly one of the most important requirements of a home. This supply should be guarded against pollution and, if possible, be distributed to various parts of the house and made conveniently available. The modern method of water distribution depends upon the fact that water is mobile and under the action of gravity seeks lower levels (Fig. 59).

A modern water system consists essentially of a storage tank located at a considerable elevation above the outlets in the homes to be served. In mountainous or hilly country, water is impounded in reservoirs at high elevations and is fed through large conduits to the distant city, where it is distributed to the homes through a complicated network of pipes. In rural communities, water may be supplied directly from a series of springs, the pipe system serving as the storing reservoir. In flat country, water is often pumped from wells or lakes into storage tanks built high above the ground.

In all these cases, the pressure which causes water to flow from the taps has its origin in the force of gravity, which urges the water to seek lower levels. If all the taps of a water system were closed, the pressure in pounds per

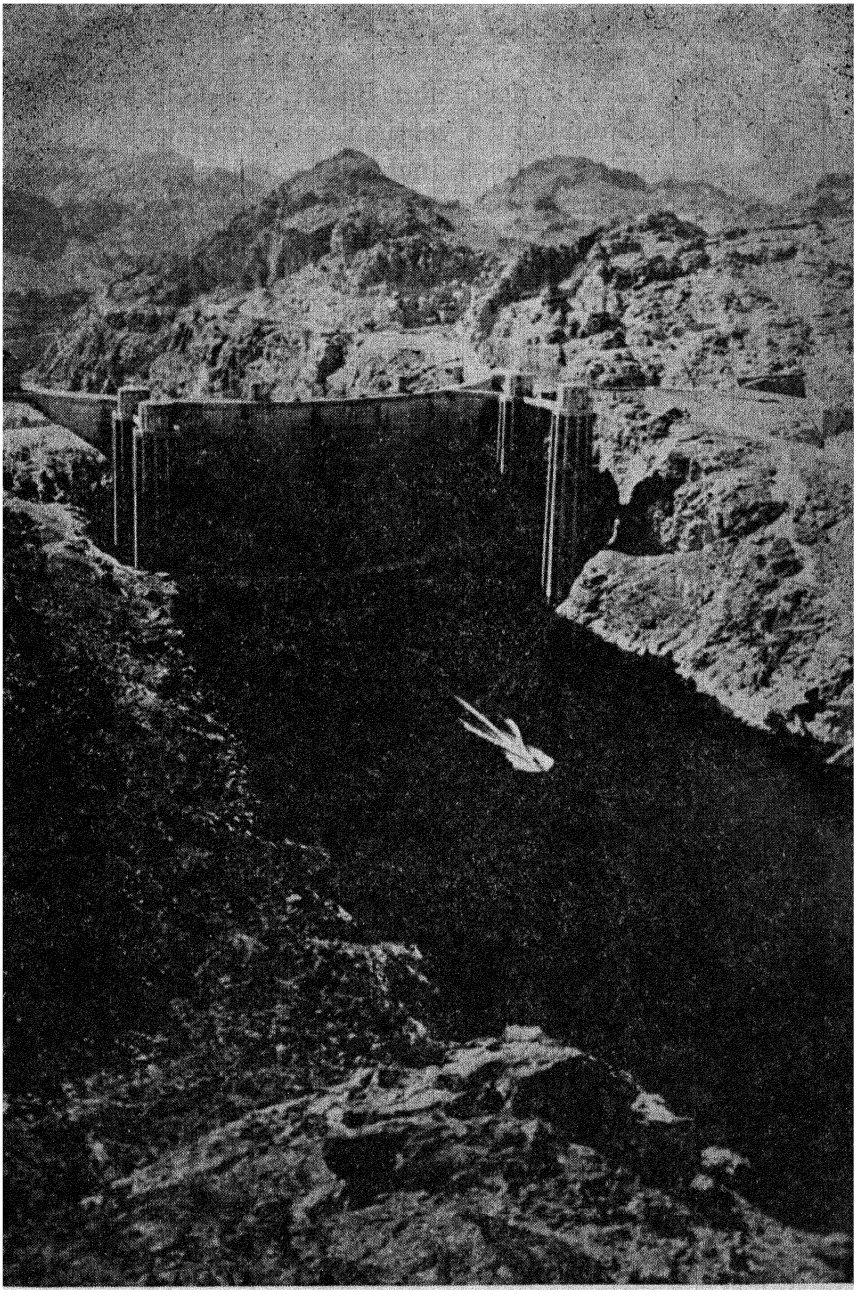


PLATE III. This is a photograph of the world's greatest dam. Five and one-half million tons of cement were poured into the canyon of the Colorado to form Boulder Dam. It is 582 feet high, and its top serves as a roadway linking Nevada and Arizona. The tiny specks you see on the top of the dam are automobiles on this highway. The lake when filled will hold 30,500,000 acre-feet of water—5,000 gallons for every man, woman, and child in the world. This great storage reservoir will be used as an aid in flood control. The conserved water will be used for irrigation and culinary purposes in the Southwest and will furnish vast quantities of electrical power as it rushes from its imprisonment.

square inch, at any particular outlet, would be equal to the weight in pounds of a vertical column of water with a cross section of one square inch and a height equal to the *vertical distance* from the outlet to the top of the water in the reservoir. In actual practice, the pressure at a given tap seldom if ever attains this maximum value. As water flows in the supply main, energy is needed to produce the motion, and the pressure along the main is less than if the water were at rest. The pressure is also decreased by the friction experienced by the water as it flows through the pipes. For these reasons a decrease in pressure is noticeable in the spray of a lawn sprinkler when someone opens a tap in the house.

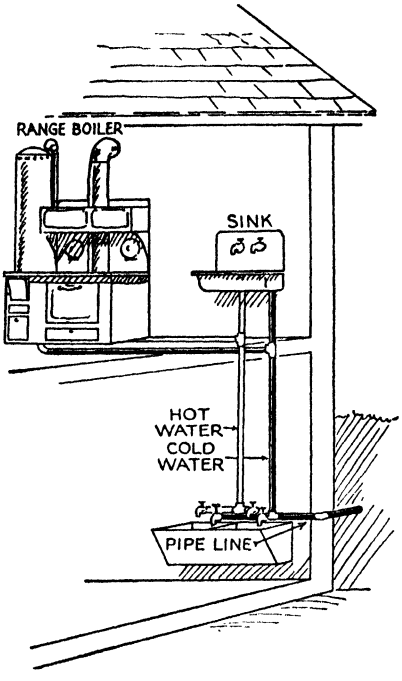


Fig. 59. Home Water Supply.

Water in Three Forms

Water, ordinarily in a liquid form, may be changed to solid ice by being placed in the freezing chamber of a mechanical refrigerator. By the process of evaporation and boiling, water may be changed into a gaseous form called *water vapor* or *steam.*

We may interpret these three states—gaseous, liquid, and solid—by assigning to molecules the property of mutual attraction, or cohesion, as well as the motion demanded by heat phenomena. In the gaseous state, the energy of motion is supreme, and a free, haphazard, to-and-fro motion is attained by the molecules. In the liquid state, the molecular forces of cohesion are slightly in excess of the forces of disruption caused by the kinetic energy of molecular motion, and a condition of restricted freedom results.

The mobility of water, which permits it to flow about, is an example of this condition. In the solid state, the molecular forces have the free hand. Form and structure, as indicated by crystallization, are in evidence. Steel bridges, great locomotives, and towering buildings are held together by such forces. Cohesive forces were present in the steel springs we used in illustrating Hooke's law (see Chapter II).

Ice. We all admire the fantastic ice patterns found on window-panes after a cold winter night. These forms portray the orderliness achieved by water molecules when marshaled in a bit of ice. By making use of the freezing chamber of a mechanical refrigerator, you may easily discover for yourself that the volume of water increases when it freezes. This explains the common occurrence of ice floating on water, and the "growth" of a bottle of milk when left outdoors, unprotected, on a cold night (Fig. 60).

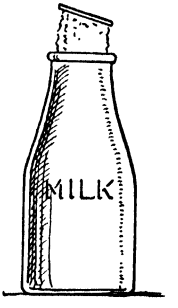


Fig. 60. Expansion Due to Freezing.

Heat of fusion of ice. We all agree that a pound of ice is more valuable as a cooling agent than a pound of ice-cold water; that is why we discard the ice-cold water that accumulates in a refrigerator as the ice melts. Careful experiments show that *80 calories of heat are required to melt one gram of ice, without a change of temperature, into one gram of ice-cold water.* This heat energy of 80 calories is not used to give kinetic energy to the water molecules. This we know is true because the water coming from the melting ice is found to have the same temperature as the ice. The heat energy, therefore, must be used in breaking down the crystal structure. It is stored, no doubt, as potential energy when the molecules are forced out of the orderly structure of the crystal against the attractive forces of that assemblage. If the crystal should take form again, this potential energy, often called *latent*, or *hidden, energy*, might be expected to appear as heat energy. This is

found to be true. One gram of ice-cold water gives out 80 calories of heat to its surroundings as it freezes to ice without a change of temperature.

In an ice box the heat required to melt the ice comes from the food and air contained in the refrigerator and from the outside through the walls and door of the structure. The most efficient boxes are those with best insulation against outside heat and with a construction which permits a positive circulation of air between the food and the ice (Fig. 61).

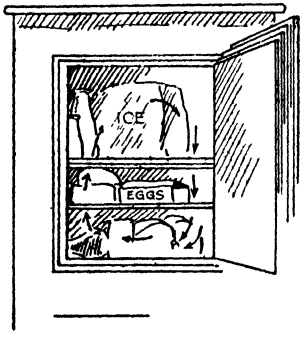


Fig. 61. A Refrigerator.

Effect of pressure and impurities upon the freezing point of water. The zero on the centigrade scale has been defined as the temperature at which *pure* water freezes or *pure* ice melts under the *pressure of one atmosphere* (76 centimeters of mercury). The freezing point of water falls below 0° centigrade under pressures greater than one atmosphere. This is beautifully illustrated by the following experiment:

A ten-pound weight is placed on each end of a small steel piano wire, and the wire is placed over a piece of ice so that the weights hang freely (Fig. 62). The wire cuts gradually and completely through the ice, yet the piece is not severed. An attempt to break the ice will show that it is just as strong at the cut as at any other part. How may this cutting and healing phenomenon be explained? Just beneath the wire, the pressure is very high. At this position the freezing point of water is lower than the temperature of the block of ice, the crystal structure is broken down, and the ice melts. But just as soon as the liquid moves out from under

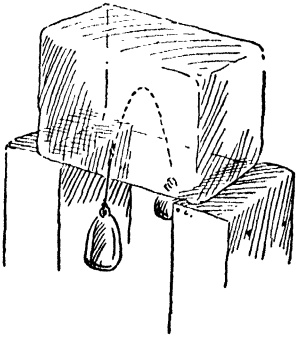


Fig. 62. Regelation.

the wire, the pressure is released, crystallization reoccurs, and the cut is mended.

When salt is added to ice, as in the process of freezing ice cream, the salt exerts a sort of "internal" pressure, which breaks down the ice structure and forces water to exist in a liquid form at a temperature below 0° centigrade (Fig. 63). The energy needed to melt the ice comes from the kinetic energy of the molecules of the ice, salt, and ice cream mixture. The temperature is reduced and the ice cream freezes. The stirring paddles are present to keep the crystals of water formed in the ice cream from assuming too large a size. The icy nature of materials frozen in a mechanical refrigerator is due mainly to the fact that large ice crystals are allowed to form. Frequent stirring during the early stages of the freezing process will greatly refine the texture of the frozen mixture.

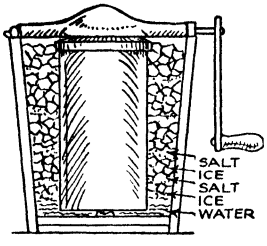


Fig. 63. Freezing Ice Cream.

Molecular forces in water. As we approach this topic, recall the puzzling childhood experience of seeing insects run over the surface of a pond. How is such a feat possible?

Let the end of a glass rod be dipped into water and then removed. The adhering water drop has the appearance of being contained in a tiny rubber bag. When water is sprinkled on a pane of glass or a wooden table top covered with eraser dust or talcum powder, the tiny particles take a spherical shape. The larger drops are somewhat flattened out. In every case, these drops act as if they were inclosed in a sort of skin that continually compresses the water into the smallest possible volume. A needle covered with a thin film of oil floats on water. The action of the water surface has every

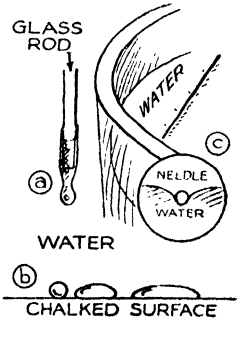


Fig. 64. Water Surface Acting Like a Stretched Membrane.

appearance of being a thin, stretched membrane (Fig. 64). Certain insects are able to skip over such a surface.

Of all forms, the sphere has the smallest surface for a given volume. The simple experiments just sketched furnish evidence, therefore, that forces are present which attempt to make the water surface just as small as possible. This so-called "surface tension" is not actually due to a membrane stretched over the surface of the liquid, but certainly at the surface there are resultant molecular forces not found in the interior of the liquid. In the interior a molecule is completely surrounded by other molecules for depths far beyond its sphere of action. But near the surface this is not the case. Here a molecule is surrounded by many more molecules on the interior side than on the surface side (Fig. 65). Referring to the figure, we see that well within the body of liquid the resultant force acting on a molecule is zero, and it moves freely about. At the surface, however, it experiences a force directed toward the interior, and its freedom to leave the liquid, or even to become a surface molecule, tends toward a minimum. The surface contracts if possible; and when we observe the phenomenon, we ascribe to the surface the properties of a thin, stretched membrane.

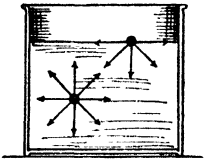


Fig. 65. Molecules at the Surface Are Attracted Toward the Interior.

Surface tension. Being a strictly surface phenomenon, *surface tension* is defined as *the force with which the liquid surface on one side of a line one centimeter long pulls against the surface on the other side of the line.* Thus, the molecular forces which compel a droplet to take a spherical shape are directly proportional to the circumference, and hence to the diameter, of the drop. The pull of gravity on this drop is directly proportional to its volume, or simply to the cube of its diameter. Halving a drop's diameter, therefore, halves the total surface force, but the pull of gravity is reduced to one-eighth of its former value. On

the other hand, doubling the drop's diameter simply doubles the total surface force, yet the pull of gravity is increased eight times. This explains why the tiny drops on the dusty table appear to be free from the pull of gravity while the larger ones are flattened out by the action of this force. In very small drops surface tension dominates, but with an increase in volume the pull of gravity increases much faster than the surface forces and, among the larger drops, gravity definitely manifests its presence. It now becomes clear why the water surfaces we ordinarily notice are flattened out by the action of gravity.

Oil on water. As a drop of oil or grease strikes a water surface, three surface tensions go into action. The surface tension of the air-water surface pulls against the surface tensions of the air-oil and oil-water surfaces, and, being stronger than their combined pull, wins out by spreading the oil over the surface of the water. You have seen drops of fat spread out as soup cools. The reason for this is that the surface tension of water increases more with a decrease in temperature than does that of oil.

Attraction between water and glass. The fact that a drop of water adheres to a glass rod when it is removed from a glass of water is evidence that the force of adhesion between the water and the glass is greater than the force of cohesion between water and water. This phenomenon may be further illustrated as follows: When water is sprinkled over the surface of a pane of glass, thoroughly washed with soap and water and dried, the droplets do not take spherical form, but flatten out to a marked degree, indicating now that the adhesive force of the glass plays an important role. On the dusty surface the water was kept from wetting the glass, and the adhesive forces between water and glass could not manifest their presence. Grease or oil would have acted as did the chalk dust. This is why glassware, and dishes in general, may be known to be

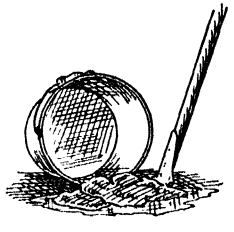


Fig. 66. Capillary Action.

free from grease when water covers the surfaces as a thin film rather than as a series of droplets.

Capillary action. If water is spilled on a linen table cloth, the wet spot rapidly increases in size, indicating that forces are present. Moisture on the hands and face is readily absorbed by a linen towel. A dampened cloth soaks up water and permits the drying of a floor being mopped (Fig. 66). Cloth is made by weaving threads, and threads are twisted fibers. A thread may contain many duct-like passages included between the fibers, and cloth will be full of small openings between the woven threads. It seems that water has a property of moving up through such small openings. This may be clearly shown as follows:

When the ends of a series of glass tubes with varying sized bores are dipped in a container of water, the liquid rises to varying heights, the height being the greatest for the smallest tube. This is known as *capillary action.* The water wets the glass and tends to spread out over it. This causes a cup-shaped surface which can easily be seen (Fig. 67). The surface of a cup is always larger than the area of the top. Similarly, the water surface in a cupped shape is larger than it would be if the surface were to become flat. Surface tension, as we have shown, tries to make the liquid surface as small as possible. Because of this, the water moves up in an attempted surface reduction. Yet all the while, adhesive forces keep the water spread out on the surface of the glass, thus preserving the cupped shape. This is possible because glass attracts water more than water attracts itself. Equilibrium is attained only when the weight of the column of water is just balanced by the surface tension.

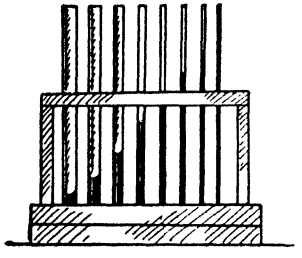


Fig. 67. Water Rising in Capillary Tubes.

Why does the liquid rise to the greatest height in the tube with the smallest diameter? When a tube of half

diameter is used, the ring of contact between the water surface and the glass is halved, the total force due to surface tension is halved, and the weight of the liquid column which can be held up is halved. But in this tube of half diameter, a liquid column is only *one-fourth* as heavy for a given height. To become *half* as heavy, the requirement for equilibrium, the liquid must rise to *twice* the height. Thus, the halving of the diameter of a tube doubles the height to which liquid rises as the result of capillary action. In general, this rise varies inversely with the diameter of the tube. Now we see why capillary action is conspicuous when small tubes are used. The height to which a liquid rises in a capillary tube depends also on the density of the liquid and its surface tension. Then, too, a rise takes place only if the liquid wets the tube. If it does not, the liquid is depressed, as in the case of mercury and glass.

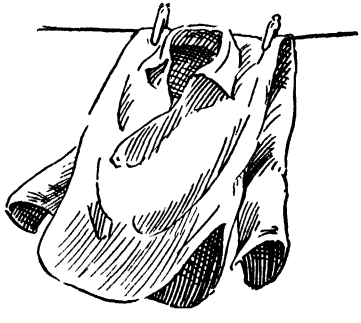


Fig. 68. Drying by Evaporation.

The soaking-up property of a towel may now be explained. Water wets cloth. This is evidence that water attracts cloth more than water attracts itself. Also, cloth contains many tube-like passages between its fibers and threads. The conditions for a capillary rise are satisfied.

Evaporation. Wet cloths become dry (Fig. 68), the earth dries up, and water boils away. How does this happen? Water molecules must leave the liquid during this evaporation process. But the surface forces, discussed above, must be overcome if molecules are to escape. Those molecules with the greatest kinetic energy might be expected to succeed, but those with the least kinetic energy will surely stay behind (Fig. 69). Since temperature is an index of the average kinetic energy of the liquid's molecules, the process of losing the most energetic ones should give rise to a cooling effect. This

agrees with everyday experience. A face covered with perspiration is cooled by fanning. As wet hands dry, a cooling is felt. A thermometer placed in a shallow dish of water registers a temperature below that of the surroundings. Because of the ease with which ether molecules escape from the liquid, the cooling effect of evaporation may be made very impressive if ether is used instead of water.

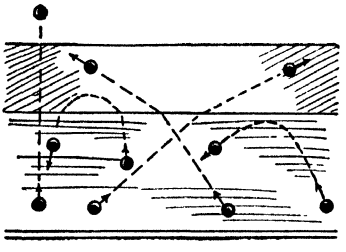


Fig. 69. Illustrating the Process of Evaporation.

Saturated vapor. Water may be kept indefinitely in a corked bottle, but it finally disappears from an open vessel. The following experiment illustrates this phenomenon in a vivid manner. Equal amounts of ether are poured into evaporating dishes and one is placed under a glass vessel. In a short time the liquid that was in the outside dish will have disappeared, but the liquid in the covered dish will have only partly evaporated. What is the explanation? In the outer vessel, the molecules which succeeded in getting through the surface and up into the air only rarely found their way back to the liquid. Thus, many more molecules left per second than returned—the condition called *evaporation.* The conditions surrounding the inclosed dish are different (Fig. 70). At first, more molecules leave than return; however, in time, owing to the increase of the number of ether molecules in the inclosed space, *as many will return as leave.* We call this *saturation,* and the molecules of ether in this space a *saturated vapor.* This vapor exerts a pressure which may be added to the air pressure in the enclosure if the total pressure is desired. Had the enclosure not been airtight, an equilibrium would have been established between the inside and outside

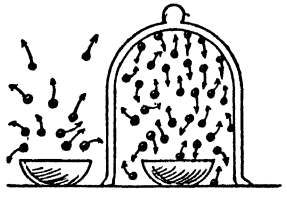


Fig. 70. Evaporation and Saturation.

pressures, and the total inside pressure would not have exceeded atmospheric pressure.

Pressure of a saturated vapor. Let two Torricellian tubes be set up as shown in Fig. 71. The heights of the mercury columns are measured. Then, by means of a curved medicine dropper, a few drops of ether are introduced into one tube and a few drops of water into the other. Immediately the mercury falls, indicating a rapid evaporation. This decrease in height is a measure of the vapor pressure, because the space was originally a vacuum and a sufficient amount of liquid was introduced into each tube so that all of it was not evaporated. Ether is found to have a greater vapor pressure than water.

Changing the volume of the vapor does ***not*** *change its pressure,* as is indicated when the tube is permitted to lean to one side and the vertical height of the mercury is measured. This means that Boyle's law does not apply to vapors as it does to gases.

An increase of temperature ***does*** *increase the pressure of a saturated vapor,* as is shown by the lowering of the column of mercury when the flame of a Bunsen burner is passed quickly over the top of either tube. Careful measurements show that a liquid has a definite saturated vapor pressure characteristic of a particular temperature. For example, at 0° centigrade water has a vapor pressure of 4.5687 millimeters of mercury, and at 100° centigrade a vapor pressure of 760.00 millimeters of mercury. At 0° centigrade mercury has a vapor pressure of 0.0004 millimeters of mercury, and at 100° centigrade a vapor pressure of 0.28 millimeters of mercury. At room temperature, water has a vapor pressure of about 2 centimeters of mercury while ether has a vapor pressure of about 40 centimeters of mercury.

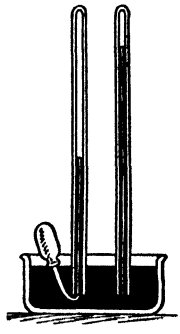


Fig. 71. Vapor Pressure.

Can the picture of molecular activity that we have painted aid in the interpretation of these facts? First, why does a change of volume not change the saturated vapor pressure? As the vapor molecules are crowded closer together, a larger number strike the liquid surface per second than before, and thus a larger number fall back into the liquid than leave it. *This condensation takes place till the density increase due to the decrease of volume is compensated for, and the equality between the outgoing and incoming molecules is re-established.* When the volume is increased, the density of the vapor is decreased and there is a lack of molecules, but evaporation brings the number up to the value demanded for equilibrium. Boyle's law *does not* apply here, because the liquid acts as a reservoir from which molecules may be obtained, and to which molecules may be returned as the circumstance demands. Air would act as does water vapor if the experiment were performed at a temperature and a pressure which would permit air to be in contact with liquid air in a closed tube or vessel.

On the other hand, the change in temperature gives the liquid molecules greater kinetic energy, a greater ascendency over the surface forces, and a greater rate of evaporation. When saturation is finally reached, therefore, more molecules are returning per second to the liquid than before the temperature was changed. For this to be possible, the space above the liquid must contain a greater supply of molecules per unit volume than formerly. This, rather than the increase of the kinetic energy of the vapor molecules, is the main source of the increased pressure due to the increased temperature. At the same temperature, ether has a greater vapor pressure than water, not because the mean kinetic energy of translation of its representative molecules is greater—this must be the same—but because the surface forces of ether are weaker and easier to get through.

Evaporation and boiling point. The evaporation of water from a bucket, pond, or lake is greatly retarded by the air

molecules. But many water molecules do penetrate very short distances, and at the surface of the liquid there is a layer of almost saturated vapor. If a current of air passes over the surface, the vapor is swept away, the vapor pressure is lowered, and evaporation increases. Thus, a wind dries the land, and we blow on soup and fan our faces to take advantage of the cooling effects of increased evaporation. If it were not for the retarding influence of the air, we would be obliged to live in an atmosphere always completely saturated with water vapor. As it is, the saturated layers next to the surfaces of oceans, lakes, and rivers furnish the moisture to the winds, which carry it into high altitudes where it is condensed to form clouds.

If water in an open kettle is placed over a flame, evaporation increases and the water soon boils. The increased temperature causes an increase of vapor pressure, an increased ability of water molecules to drive air molecules up and away from the liquid surface. Soon an upward current of water vapor, made visible by the tiny droplets of water (condensed vapor), gets under way. Then, a new phenomenon is noticed. Small bubbles appear at the bottom of the vessel; they rise a little way and disappear. These can be distinguished from the air bubbles which often appear at an earlier stage because the air bubbles increase slightly in size as they rise. Finally, the liquid is so hot that the vapor bubbles reach the surface and burst, and we have the characteristic process called *boiling* (Fig. 72a).

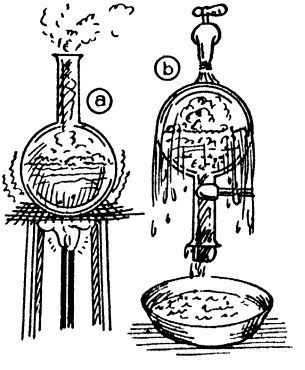


Fig. 72. Boiling at Atmospheric and Reduced Pressures.

At what temperature is boiling reached in an open kettle? If bubbles which try to form in the liquid are not to be crushed, the vapor pressure inside of them must be equal to the pressure in the liquid. Part of the liquid pressure is caused by the weight of the water, but except in the case of

geysers, this part is very small. The main part, then, must be due to atmospheric pressure, and bubbles successfully form when the vapor pressure within them is equal to this outside pressure. But the bubbles and water are in contact and at the same temperature; hence, the *boiling point of water in a kettle open to the atmosphere is reached at a temperature at which the vapor pressure is equal to the atmospheric pressure.* If a thermometer is placed just above boiling water, the temperature is found to be the same as in the liquid itself. Owing to the upward rushing of water vapor, the air is driven out and the atmospheric pressure in this region is maintained by the vapor itself. The vapor just above the liquid, therefore, establishes the same pressure as the bubbles in the liquid and hence must be at the same temperature (Fig. 73).

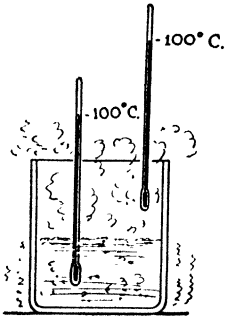


Fig. 73. During boiling, the vapor just above the surface has the same temperature as the water.

At sea level, where the atmospheric pressure is 76 centimeters of mercury, water boils at 100° centigrade; at Salt Lake City, where the average atmospheric pressure is 65 centimeters of mercury, it boils at 96° centigrade. Under a pressure of 15 pounds to the square inch (above atmospheric pressure), water boils at 122° centigrade.

From this it is clear that water may be made to boil at any temperature by subjection to the proper pressure, the pressure of its saturated vapor for that particular temperature. At room temperature, this pressure is about 2 centimeters of mercury. Thus, by the use of a pump that can maintain this pressure through pumping out the air and then the vapor, water may be made to boil at the temperature of the room. In fact, by the proper reduction of pressure, water may be made to evaporate so rapidly as to freeze itself by the cooling effect of evaporation. The change of boiling point with pressure is utilized in the pressure cooker, where boiling temperatures

as high as 122° centigrade may be obtained. The time required to cook food is thus very greatly reduced. Sugar is refined by the use of vacuum pans in which boiling temperatures as low as 70° centigrade are reached.

Water may be shown boiling under reduced pressure in a vivid manner as follows: Let the water in a half-filled flask be brought to a vigorous boil, thus making sure that all the air is forced out, and let the flask be removed from the flame, corked airtight with a rubber stopper, and inverted (Fig. 72b). Then let cold water be poured over the flask. The steam above the liquid condenses, the vapor pressure is greatly reduced, and the water in the flask springs into vigorous boiling.

Canning Fruit

In the cold pack method of canning fruit, glass jars are filled with prepared fruit and sirup to within about an inch of the top. Lids are loosely fastened in place, the jars are placed on a rack in a boiler, and warm water is added until it stands three or four inches from the tops of the jars. The water is brought to the boiling temperature and the cooking process is continued for twenty minutes or more, depending on the type of fruit being canned (Fig. 74). The jars are then sealed tightly and removed to cool. In such a process, it is possible to boil away large quantities of sirup, leaving the jars only partly full when the cooking is finished. Such a result is certainly not to be desired.

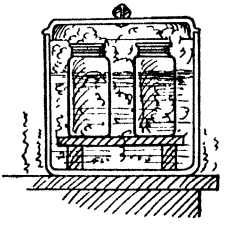


Fig. 74. Canning Fruit.

Let us investigate the problem in terms of the physics we have studied. Suppose that one of the jars described above is sealed so tightly that no liquid, vapor, or air can escape. Atmospheric pressure will exist in the air space because of the direct contact between this space and the atmosphere just before sealing. This pressure will be due to water vapor and air, the fractional part due to air depending

upon the temperature of the fruit and the sirup, that is, upon the vapor pressure of the sugar solution at the temperature of the mixture. (This vapor pressure at a given temperature will be slightly less for a sugar solution than for pure water; see below for a discussion of the boiling points of solutions.) As the bath is brought to a boil, the liquid in the jar will expand, thus compressing the confined air and increasing its pressure (Boyle's law). This volume change will not affect the vapor pressure. The temperature change will increase the air pressure (Charles's law) and also the vapor pressure. But owing to the sugar in the water, the vapor pressure will approach but not reach the atmospheric pressure. The lid may bulge, and this increase of volume will lower the air pressure (Boyle's law). It is safe to say that when all these facts are taken into accurate account, the final pressure in the jar will seldom if ever exceed two atmospheres, one atmosphere above outside pressure. Such a pressure might break some of the jars.

However, the caps may be left slightly loose so that the air might escape while the bath is brought to the boil, and then they may be tightly fastened and kept so during the cooking process. In this procedure, the air is gradually pushed out by the water molecules as the sirup is brought to the boil. Finally, with the boiling point attained, the pressure within the jar is just a little above the pressure of the atmosphere and is principally due to vapor pressure. Since from this point on there is no change in the temperature of the bath and of the fruit, vapor pressure does not increase and it is safe to fasten the lids tightly.

Most jars are supplied with lids which leak if pressure is supplied from within, and many of them, especially the metal tops which may bulge a little, may be screwed down tightly from the very start. If this is done and the boiling of the water bath is kept just above a simmer, seldom if ever will there be more than a slight loss of sirup. The lids may be bulged out, if of metal, when the jars are removed:

but after cooling, the vapor pressure decreases, no air returns to maintain the atmospheric pressure, and a partial vacuum results. The lids will then be pushed in, and when the jar is opened an inrush of air will be noticed. It is important that an inch of space be left above the liquid, for if the jar were completely filled, the liquid expansion might break the container.

The Heat of Vaporization

A thermometer placed in the open vessel mentioned above indicates a steady increase of temperature as heat is added; but as soon as the boiling point is reached, the temperature rises no further. At the boiling point, it seems that the cooling effect of evaporation is just balanced by the supply of heat. This is so because, when the flame is lowered, the boiling reduces to a simmer; when increased, the boiling becomes turbulent. In both instances, water molecules gain kinetic energy from the flame continually while the water loses its fastest molecules. The average speed of the molecules left behind remains constant. Even though the fastest molecules leave the liquid for the vapor, they arrive in that state with much less energy than that with which they started, because they have to move through the opposition of the liquid surface and aid in the expansion needed to furnish the space in which to collect. The work done by a molecule in changing from the liquid to the vapor state, under the action of its kinetic energy, is stored by the molecule as potential energy. This explains why, after the boiling point is reached, the kinetic energy (temperature) of the average vapor molecule is just equal to that (temperature) of the average liquid molecule.

The quantity of heat measured in calories required to change one gram of water into vapor at the same temperature is known as its ***latent heat of vaporization.***

By careful measurements, the heat of vaporization of water is found to be 541 calories per gram at 96° centigrade

and 539 calories per gram at 100° centigrade. This means that one gram of water, in order to be boiled away, must receive seven times as much heat as is required to heat it from room temperature to the boiling point. In condensation, this latent heat of vaporization is returned. Hence, evaporation is an energy-absorbing process—a cooling process, and condensation is an energy-liberating process—a heating process.

The Boiling Point of Solutions

When salt is added to water, the boiling point rises, the increase being directly proportional to the amount of salt added, if the solutions are dilute. Salt acts, therefore, as if it were a pressure added to the water. It probably has the effect of an "internal pressure" adding to the attractive forces which the water molecules must oppose in freeing themselves from the liquid. At the same temperature, then, a solution of water and salt cannot maintain as high a vapor pressure as water without the salt. For the same reason, a concentrated salt solution has a lower vapor pressure than a dilute solution at the same temperature. Now we see why temperatures higher than the boiling point of water must be reached before a salt solution can produce a vapor pressure equal to one atmosphere—the requirement for boiling.

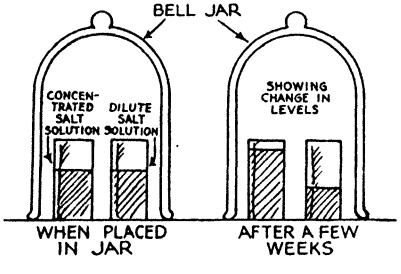


Fig. 75. Evaporation from a Dilute Solution, Condensation to a Concentrated Solution.

A simple experiment illustrating this phenomenon may be carried out as follows: Let equal amounts of two salt solutions, one concentrated and the other dilute, be poured into beakers of the same shape. Then let both beakers be covered with a single bell jar (Fig. 75). After a time, the quantity of the more concentrated solution will increase, and that of the less concentrated will diminish. At first, both liquids evaporate

molecules into the air space, but by the time the saturated vapor pressure has been reached for the concentrated solution it will not have been reached for the other. As a result, evaporation continues at the surface of the dilute solution and condensation takes place at the surface of the more concentrated solution.

Osmosis

We all have observed that dried fruit when placed in water swells up, owing to the passing of water through the skin. We may have heard that the exchange of oxygen and carbon dioxide in the lungs, the passage of digested food into the blood stream, and the passage of water up and of sugar down in a sugar beet are all due to osmosis.

Let us try a simple experiment. Let the interior of a carrot be cut out, the space filled with a thick sirup, a rubber stopper containing a long glass tube be placed firmly into the mouth of the cavity, and the carrot be immersed in a glass of water (Fig. 76). The water begins to flow through the pores of the carrot and into the sirup. The sirup tends to come out through the carrot, but the cell structure does not permit an easy passage of the sugar molecule. As a result, more water flows in than sirup flows out, and the liquid rises in the tube.

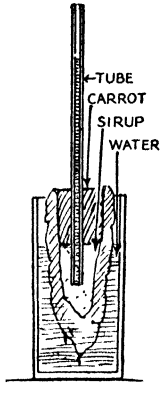


Fig. 76. Osmosis.

Now recall the experiment of the last section. A water surface is permeable to water molecules, but not to salt molecules. It is therefore equivalent to a semipermeable membrane. Under the conditions described above, water molecules pass through such a membrane as they flow from the solution where the concentration of water is greater, through the air, and into the solution where the concentration of water is less. Thus, the water moves from a region of higher water concentration to one of lower water concentration. Salt molecules cannot migrate because the water surface does not permit it.

Similarly, water moves into the carrot in an attempt to make the water concentration the same on both sides. The sugar tries to do the same, but because of the semipermeable nature of the carrot, fails. Each substance tries to move in the direction of its lesser concentration. This diffusion, when attempted through a semipermeable membrane or its equivalent, gives rise to the phenomenon called *osmotic pressure.*

Ventilation

Ventilation is the process of maintaining the air of a dwelling in a condition conducive to maximum health. Recent investigations in this field permit the conclusion that bad ventilation does not result from a diminished oxygen content or from an increase of carbon dioxide or inorganic poisons, but is caused by an unduly high air temperature, an excessive humidity, and a lack of air circulation.

In the average home, ventilation is accomplished by making use of the natural draft generated through differences in air temperature. In the summer, the windows and doors may be left open, the main problem being the control of the house temperature by their proper regulation. In some climates, it is found advisable to open all doors and windows during the night, thus permitting the house to cool, and then to close all openings in the early forenoon in order that the warm air from the outside may not enter.

In the winter, the difference in temperature between inside and outside may be great, and large drafts of air may be produced with small openings. In many residences, in mid-winter a sufficient amount of new air finds its way through cracks and crevices and through doors as the occupants pass in and out.

These natural methods of ventilation may not adequately mix the air. Cool air tends to remain near the floor area and warm air at the ceiling. Unless there is some means of stirring up the entire air volume, the ventilation arrange-

ment will be inadequate. In the process of natural ventilation, the entering air should either be admitted at the highest part of the room or be directed towards this region, in order that the cold air may have a chance to mix with the warm air of the room (Fig. 77).

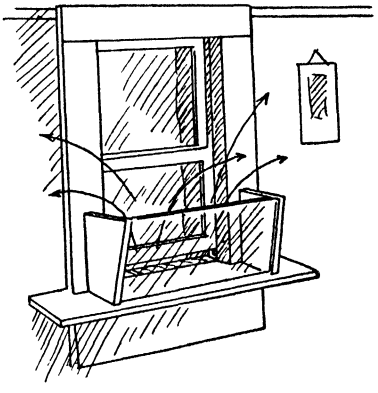


Fig. 77. Ventilation.

Mechanical devices may aid greatly in mixing and circulating the air of a dwelling, and are often made a part of its heating equipment. In such a scheme, by means of a fan or blower, warm air is forced in near the floor in the central part of the house and sucked out from registers located near the floor at the outer walls of the various rooms, or, in latest practice, vice versa. This insures an adequate circulation of the air, and the cold air layer, which often is near the floor, is removed by means of the forced ventilation.

Humidity

The relative humidity of air at an observed temperature is defined as the ratio of the water vapor actually present in a volume of air to the amount that would be contained in this volume if the vapor were saturated at this temperature. A relative humidity of 50 per cent, for example, means that if the moisture content were doubled without a change of temperature, the space would be saturated. We have explained the change of the pressure of a saturated vapor with the temperature as being due, in the main, to a change in the number of vapor molecules present in a unit volume. But since the vapor pressure decreases with a decrease in temperature, it must mean that fewer molecules per unit volume are required to saturate a cold space than a warm one. *Hence, if a partially saturated space is sufficiently cooled, it may become saturated.* The temperature

at which this condition is reached is called the *dew point.* On a summer day when the relative humidity is rather high, dew may be seen to collect on a glass of ice water. During the winter months water collects freely on the inside of cold window panes, thus decreasing the humidity of the dwelling. The dew point is reached on the surfaces of these cold objects.

Air that filters into a home through the cracks around windows and doors has a very low humidity when warmed up to the room temperature, even though its relative humidity may have been high when outside. For example, when air at a temperature of 10° Fahrenheit and a relative humidity of 50 per cent is warmed to 70° Fahrenheit, it has a relative humidity of only 5 per cent. Unless water is added in some manner, a home may in the winter time be as dry as a desert.

The nasal passages have the task of warming and humidifying the air before it enters the lungs. Thus, if a person is walking in the open on a cold winter day with the air at a temperature of 32° Fahrenheit, sufficient heat must be supplied by these passages to change the air to body temperature, 98° Fahrenheit. Moisture must also be supplied, because air at a temperature of 32° Fahrenheit and a relative humidity of 100 per cent will have a humidity of only 9 per cent when warmed to the temperature of the body. If during an hour's walk this person breathes 750,000 cubic centimeters of air, 31 cubic centimeters (1 ounce) of water must be supplied by the mucous membranes of the nose and throat. Approximately three times as much heat is needed to evaporate this water as is needed to heat the air breathed.

As this person enters a home on the same day, he will find the air warmed to at least 70° Fahrenheit, and his nasal passages will not need to supply much heat to warm the air which he breathes. In many homes, however, very little moisture is added deliberately, and since most of the air comes at one time or another by infiltration from the outside, it will have a very low humidity, probably as dry as

desert air. The nasal passages are again called upon to furnish moisture and the heat required to evaporate it. Thus, in winter in an unhumidified room, the mucous membranes must furnish approximately the same amount of moisture indoors as out-of-doors and nearly as much heat. Such a room has a drying effect on the mucous membranes not noticed when out-of-doors, since in the cold environment the mucous secretion is stimulated.

Hence, nasal passages become irritated and inflamed because of excessive drying. It is believed that colds in the head and throat are favored decidedly in this abnormal situation. The obvious solution is to add moisture to the room. It is usually considered that a relative humidity of between 40 and 60 per cent is adequate for proper living conditions. If the humidity is kept between these figures, a person in the room will feel comfortable at a temperature of 68° Fahrenheit. At such humidities, the air will seem oppressive if temperatures as high as 75° Fahrenheit are reached. On the other hand, if the air is very dry, a person may feel cold at a temperature of 72° Fahrenheit. With forced air circulation, the bodily feelings just described can be had only at somewhat higher temperatures.

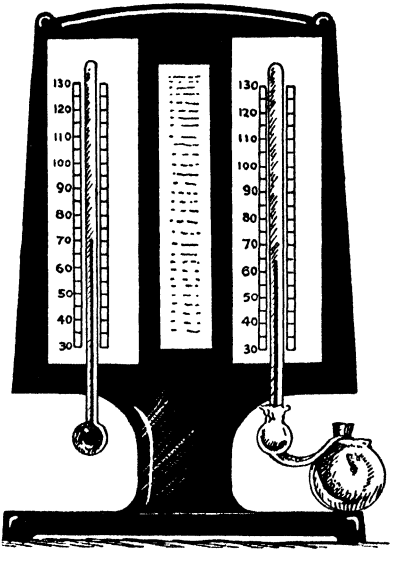


Fig. 78. Hygrometer, Wet and Dry Bulb.

The actual humidity in a home may be measured in a fairly reliable manner by reading the temperature on a thermometer with a dry bulb and then on one with a bulb surrounded by a wet wick. The wet bulb thermometer reads lower than room temperature because of the cooling effect of evaporation, the lowering being an index of the evaporation rate, which in turn depends on the humidity of the air. Knowing this lowering of tempera-

ture and the room temperature, the humidity may be obtained from specially prepared tables. Such an instrument is called a "wet-and-dry-bulb" thermometer, and is often equipped with tables which are made easy to read by some mechanical device (Fig. 78). Instruments used to measure relative humidity are called *hygrometers*. Fairly reliable ones, made upon the principle of the variation in the length of a hair under different conditions of moisture, are now on the market at moderate prices (Fig. 79).

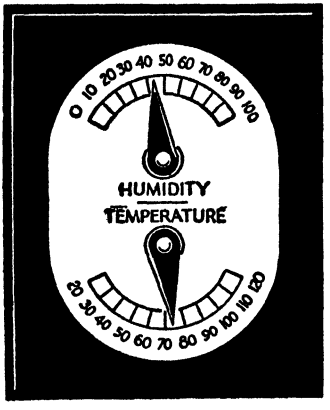


Fig. 79. Hygrometer, Hair Type.

Heating the Home

The transmission of heat. Homes are heated with the energy liberated by fuels as they are burned in stoves or furnaces. Thus, if all parts of a residence are to be warmed, heat must be transmitted by some means from this central source. It is found that heat may be transmitted by radiation, conduction, and convection.

Radiation. All hot bodies radiate energy, the rate increasing very rapidly with an increase in temperature. One is made aware of this changing phenomenon while cooking over a stove that is becoming red hot. The sun is very hot; the temperature of its visible surface is about 6,000° centigrade; its interior may reach 18,000,000° centigrade. The radiant energy from this hot body comes to the earth with the velocity of light. Directly and indirectly it is the source of man's available energy. Heat radiation is a wave motion, belonging to the great family of electromagnetic waves of which visible light, X-rays, and radio waves are members.

Conduction. The handle of a cooking utensil, though not in direct contact with the flame, becomes warm, sometimes very hot. The energy is transmitted in a sort of hand-

to-hand fashion from molecule to molecule. It is by this means that the heat of the flame goes through the metal walls of a stove or furnace, through the walls of a home, and through the panes of glass in the various windows.

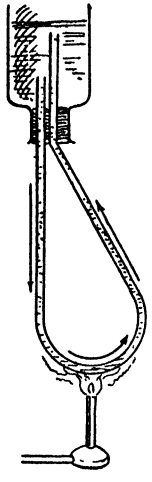


Fig. 80. Illustrating Liquid Convection.

Convection. We know from previous experience that air and water expand when heated. This means that a given volume of these substances when hot will be lighter than when cold. Thus, warm air and warm water tend to ascend and cold air and cold water to descend. When a little sawdust is added to a vessel of boiling water, convection currents become clearly visible. The apparatus illustrated in Fig. 80 also gives a vivid demonstration of convection.

Home heating devices. *The chimney.* The Indian builds a fire in the center of his teepee and allows the smoke to escape through a hole at the apex of the conical structure (Fig. 81). The chimney is the outgrowth of this idea. As the flue is warmed, cold air passes through the fireplace or stove, forces up the warmed air, and produces the so-called draft which brings the necessary oxygen to the burning fuel.



Fig. 81. The First Chimney.

The heating stove. The air near the stove becomes warm by conduction and is then forced upward by cold air. This action creates convection currents which in time bring the air from all parts of the room to the stove to be heated.

Hot air heating system. A furnace is surrounded with an air jacket. The lower part of this air space is connected by means of large pipes to registers placed near or in the floor at the outer walls of the rooms above. The upper part of this space is connected by means of similar pipes to registers placed in the central part of the house. When the fire is lighted, the air in the jacket is heated, and the cold air near the outside walls of the rooms above flows down the pipes and forces the warm air up into the central part of the house. In this way the coldest air is brought continually to the furnace to be heated. The circulation from the hot registers to the cold registers may be made more positive, and greater efficiency obtained, if an electric fan is used to force the air through the furnace. This practice avoids the need of heating the air to a high temperature before good circulation is assured. A pan of water placed over the furnace and inside the air jacket serves as a humidifier.

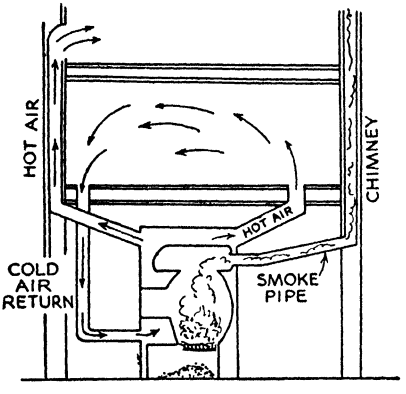


Fig. 82. Hot Air Heating System.

Hot water tank. The water in the jacket is heated by the fire and expands. The cold water, by the action of gravity, is brought into the jacket from the tank through the lower pipe and forces the hot water out into the tank through the upper pipe (Fig. 83). In time this circulation brings all the water into the heated jacket, where by conduction it receives heat from the burning fuel.

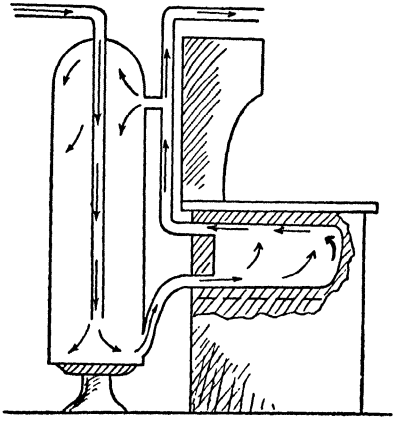


Fig. 83. Hot Water Tank.

Hot water furnace. The principle used here is the same as that used in the hot water tank, except that a number of radiators replace the tank, and the water jacket is much larger and specially designed. The water is not connected directly with the water mains, and so this system must have an expansion tank to take care of the increase in volume due to heating. A radiator heats a room in the same way that a stove does, that is, by air convection. The radiators are equivalent to a series of small stoves placed in different parts of the home, but the hot water system has many advantages over this series. First, one fire will heat the house; second, owing to the high specific heat of water, the radiators will remain hot for a time after the furnace fire dies down; third, the radiators do not overheat the air as stoves often do unless very carefully managed.

Steam heating plant. In this system, steam is generated under pressure in a boiler located in the basement and is allowed to flow through insulated pipes to cast-iron radiators located in the rooms above, where it condenses to water, giving out its heat of vaporization. In the low-pressure gravity system, the steam operates at a pressure of from three to four pounds per square inch (above atmospheric pressure), and the condensed water returns by the action of gravity. In the so-called vapor system, the steam operates at a few ounces of pressure. This system, if well designed, is efficient and quiet, and such equipment is usually installed if steam is selected as a home heating agent.

All home heating devices should be equipped with some type of humidifier.

Heat insulation. By careful experimentation it is found that the quantity of heat which passes through a wall depends upon the time (being twice as great for twice the time), upon the difference of temperature between the hot side and the cold side (being twice as great if this tempera-

ture difference is twice as great), upon the surface exposed (a wall with twice the area will pass twice the heat), upon the thickness of the wall (if the thickness is doubled, the amount is one-half), and finally, upon the nature of the substance composing the wall. We may now put these ideas into the language of algebra, as follows,

$$H = \frac{k(t' - t)AT}{d}, \tag{7.1}$$

where H is the quantity of heat measured in calories, k is the coefficient of conductivity of the substance (values listed below), t' is the temperature of the hot side, t is the temperature of the cold side, A is the area of the surface, d is the thickness of the wall, and T is the time of flow.

In cold weather, fires are built in homes to take care of the heat loss through the walls, ceilings, and windows. Heat insulation is an important subject for our study (Fig. 84). Let us first list data on the heat conductivity of different substances. The coefficients change very slightly with the temperature but are sufficiently constant for our purposes.

COEFFICIENTS OF HEAT CONDUCTIVITY

Silver	1.006	Asbestos	0.0004
Copper	0.918	Dry soil	0.00033
Aluminum	0.480	Wood (with grain)	0.00030
Iron	0.161	Flannel	0.00023
Ice	0.005	Sawdust	0.00012
Glass	0.0025	Mineral-wool	0.000094
Concrete	0.0022	Wood (across grain)	0 00009
Brick	0.0015	Attic-wool	0.000088
Water (no convection)	0.0013	Felt	0.000087
Old snow	0.0012	Air (no convection)	0.000057
Plaster	0.0005	Eiderdown	0.000011

How does a unit area of window compare with a unit area of brick wall in the heat loss due to conduction? Let us find the answer by making calculations using the following data:

Inside temperature	20°C.	Outside temperature	−10°C.
Thickness of glass	0.4 cm.	Thickness of wall	20 cm.
Coefficient of heat conductivity for glass			0.0025
Coefficient of heat conductivity for brick			0.0015
Time for comparison purposes			1 sec.
Area of both glass and brick, for comparison purposes			1 sq. cm.

For glass.....$H = (0.0025 \times 30) \div 0.4 = 0.19$ calories.
For brick.....$H = (0.0015 \times 30) \div 20 = 0.0022$ calories.

This means that a square foot of window area will conduct nearly as much heat as 100 square feet of wall space. Because of this, in very cold countries storm windows are added. The air space between the two sets of window sash serves as a good insulator. Were it not for convection currents, this air space would have exceptional insulating properties, the coefficient for air being 0.000057.

Assuming the ceiling (plaster) to be one-tenth as thick as the wall (brick) and making use of the data in the table, a simple calculation shows that about three times as much heat is conducted out through a square foot of ceiling as through a square foot of wall. It is true that the temperature of the attic will not be as cold as the outside, but the inside ceiling temperature is always warmer than the average inside wall temperature, so in the calculations we assumed the difference of temperature to be the same for ceiling as for wall. By using an insulating material on the ceiling (often this is placed in the attic directly over the lath), this heat loss may be reduced greatly in winter and the house may be kept cool in summer.

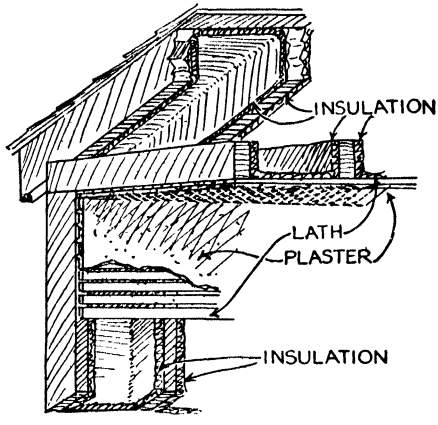


Fig. 84. Insulating Against Heat and Cold.

A tile or concrete floor seems colder to the touch than a wooden floor, even though a thermometer indicates the same temperature. The explanation is that heat is conducted away more rapidly by the concrete than by the wood. A hammer that has been in the hot sun will not feel very warm if picked up by the handle, but if picked up by the head it will seem very hot indeed. Try it.

Farmers realize that a concrete floor conducts heat more rapidly than a wooden floor and hence they supply their animals with additional straw, which serves as a good insulator. Apples and vegetables are covered with straw and earth and so are kept from freezing. Heat gradually leaves the vegetables, and if the winter is very cold and protracted, their temperature may even go below the freezing point unless there is a sufficient depth to the covering. As spring comes, heat gradually finds its way into the pit, but not until the weather is rather warm does the temperature of the vegetables become too great. This lag of ground temperature with the seasons and the insulating properties of soil make it possible to draw cold water from the mains in the summer time.

Ice boxes, fireless cookers, and insulated ovens furnish other important examples of how the flow of heat may be checked by proper insulation. The insulating materials used, such as sawdust, celotex, celoair, asbestos, and cork, are materials which contain "dead" air spaces. The very small pockets of still air—still because convection currents are lacking in such small air pockets—add to the insulating ability of the substance because still air has very small heat conductivity.

Questions and Problems

1. Why does horizontal distance not contribute to water pressure in a water system as vertical distance does?

2. Describe a practical experience which is evidence that ice may change to water vapor without first becoming a liquid.

3. Explain why during a very cold snap a large tub of water is often placed in a vegetable cellar.

4. Explain why the banks of a stream remain moist even though they are above the level of the water.

5. Explain why the readings on a thermometer may be used to give accurate information concerning the concentration of sirup in the process of candy-making.

6. If the hand is placed in the foggy part of the vapor stream issuing from the mouth of a boiling tea kettle, little damage will be done; but

if it is placed in the invisible region near the mouth of the spout, a severe burn may result. Explain.

7. Why does boiling sometimes cease for a moment when water is salted?

8. Describe how you would humidify a home equipped with (*a*) a hot blast stove; (*b*) a hot air furnace; (*c*) a hot water furnace.

9. The boiling point in a vessel open to the atmosphere is reached at a temperature at which the ..... is equal to the .....

10. ..... calories of heat are required to melt one gram of ice without a .....

11. Vapor pressure increases with ..... but remains constant with a change of .....

12. Boyle's law may be applied to a ..... but not to a .....

Suggested Readings

(1) Crew, Henry, *General Physics*, The Macmillan Company, New York, 1927, pp. 172–194.
(2) Harris and Butt, *Scientific Research and Human Welfare*, The Macmillan Company, New York, 1924, Chaps. 45–47.
(3) Keene, E. S., *Mechanics of the Household*, McGraw-Hill Book Company, Inc., New York, 1918, Chaps. I–XII.
(4) Kirkpatrick and Huettner, *Fundamentals of Health*, Ginn and Company, Boston, 1931, Chap. VIII.



CHAPTER VIII

Mechanical Vibrations

Most people have an appreciation for rhythm. Primitive man delights in the periodic thud of the tom-tom, and his sacred dances vividly portray the rhythm of bodily movements. The periodic swing attendant upon walking, the heart beat, the sweep of a pendulum, and the tick of a watch are examples of recurrent phenomena. We observe the swaying of trees and the nodding of grain under the action of a breeze. Periodically the sun appears overhead and the seasons come and go. All these are examples of periodic motion. We shall turn our attention to the most elementary of all oscillations—*simple harmonic motion*—and describe the vibration patterns of oscillating springs, rods, plates, strings, and air columns.

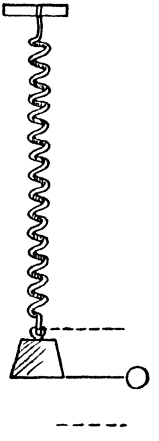


Fig. 85. Vibrating Spring.

The Vibration of a Mass Attached to a Spring

Recall the experiment described in Chapter II, illustrated in Fig. 4, in which we discovered that the distance that a spring stretches is proportional to the force applied. With every mass applied, equilibrium is achieved at a definite elongation, the pull of gravity downward being just equal to the pull of the spring upward. Such a system is illustrated in Fig. 85. If the mass is displaced either upward or downward from this zero position and then

allowed to go free, it will take on a rhythmic motion, a vibration which seems to be precisely periodic but which slowly disappears as the body finally comes to rest at its zero position.

Cause of the mechanical vibration. As the mass is pulled below its zero position and then released, the upward pull of the spring is greater than the downward pull of gravity and an upward *accelerated motion results.* By the time the mass has reached the zero position, the upward pull of the spring is just equaled by the downward pull of gravity, but because of the inertia of the mass it continues to move upward until a force—which is now the excess of the pull of gravity over the diminishing force of the contracting spring—brings it to rest. At this highest position, the excess of the force of gravity over the force of the spring is at a maximum, and the mass is accelerated downward. Again because of its inertia the mass does not stop at the zero position, but carries through until the increasing unbalanced force, due now to the excess of the pull of the spring over the pull of gravity, brings it to rest. Again the unbalanced force is at a maximum, the mass is accelerated upward, and the process is repeated over and over again.

The conditions for simple harmonic motion. Thus we see clearly that the vibration just described arises out of the combined action of force and inertia. It takes a force to accelerate the mass, and inertia to carry it through the zero position. Inertia *alone* keeps a body in motion at a constant velocity. A constant unbalanced force imparts uniformly accelerated motion to the body on which it acts. But in the mass-spring-gravity system, or simply the oscillating spring, just described, the force acting is *not constant,* but increases directly with the distance moved on either side of the zero position and acts upward when the displacement is downward and downward when the displacement is upward. The motion, which always arises under the combined effects of this type of force and

the inertia of the moving mass, is called *simple harmonic motion.* Since all bodies have mass, we have only to discover the presence of this type of force to predict that the resulting motion will be simple harmonic.

Stiffness of a spring. In the oscillating spring, we acquire a quantitative measure of the inertia effect of the system by determining its mass in grams. But if the magnitude of the force acting is to be obtained, we must know the stiffness of the spring—the force of gravity being constant. *The stiffness of a spring is defined as the force measured in dynes required to stretch it a distance of one centimeter.* Thus, for example, to obtain the upward pull of the spring in dynes, we have only to multiply the downward elongation in centimeters by the stiffness of the spring.

Energy of oscillations. The oscillations of a spring may be interpreted in terms of the principle of the conservation of energy as follows: Let the mass (Fig. 85) be pulled down a distance D below the zero position and brought to rest. Work is done and a definite amount of potential energy is stored. During this process of spring elongation, the force of opposition *increases* with the displacement, and the problem of calculating the work done is not the simple one encountered in lifting weights against the constant force of gravity. However, when the problem is solved by a method too advanced to be presented here, the potential energy is found to be

$$P.\ E. = \frac{1}{2} f D^2 \tag{8.1}$$

where f is the stiffness of the spring and D is the initial displacement.

With the release of the mass, a *conversion* of potential into kinetic energy begins, and by the time the mass reaches the zero position all of the energy is kinetic. As the mass moves upward toward its highest position, kinetic energy is converted to potential energy, and at the highest point all of the energy is potential. Thus, we have a cyclic

transfer of energy from potential to kinetic, back to potential, then to kinetic, and so on. Owing to air friction and also the loss of energy in the stretching and contracting of the spring, the original supply of potential energy is finally dissipated, and the mass comes to rest at its zero position. *With no dissipation of energy, the sum of the potential and kinetic energy is constant at all positions in the path of a vibration.* In the above case this is very approximately true.

Amplitude of vibration. Let the mass of Fig. 86 be displaced from its zero position. When allowed to go free, it will oscillate with definite up-and-down sweeps. *The* ***amplitude of a vibration*** *is defined as the maximum distance moved either up or down from the zero position.* The complete sweep is twice the amplitude of vibration. As the body oscillates, the amplitude decreases slowly, finally reaching zero when the body comes to rest. But for a given stiffness of spring, the square of the initial amplitude is an accurate index of the potential energy given to the system (Equation 8.1); hence, the decreasing amplitude is evidence that the mechanical energy of the system is either being passed over to some other system (in this case, mechanical vibrations in air) or being transferred directly to the heat energy of moving molecules. We say that the oscillations are being *damped*, and the motion is known as *damped simple harmonic motion.* If the mass is made to vibrate under water, the damping will be greatly increased.

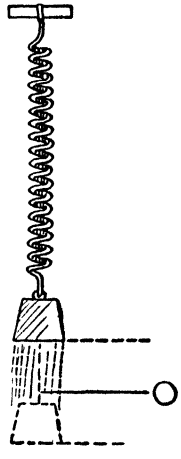


Fig. 86. Amplitude of Vibration.

Damping is often a decided disadvantage; for example, the "cat dies" too soon for a child in a swing. But were it not for its presence, all elastic bodies, which we so freely set into vibration as we awkwardly move about, would vibrate indefinitely. Our physical environment would be a shaky one indeed.

A damped oscillation is pictured in Fig. 87. With each vibration the mass actually retraces a certain portion of its path. An attempt to draw this fact would yield a straight line with no evidence of the nature of the motion clearly portrayed. For this reason the oscillations are spread out, *horizontal distance being used to represent time.* As the eyes sweep through the picture from left to right, one obtains a notion of what is happening to the amplitude of the vibration with increasing time.

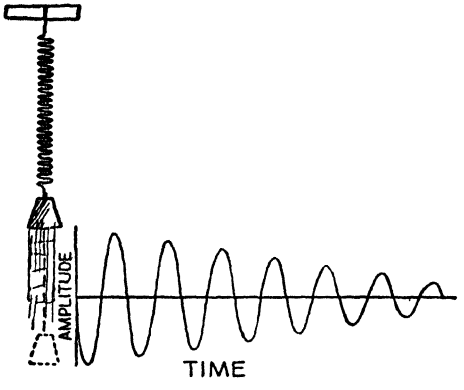


Fig. 87. Damped Oscillation.

Period and frequency of vibration. The reader has already anticipated the need of measuring the time required for a single to-and-fro motion of the vibrating mass. Let us add a 100-gram mass to a spring, set the system into oscillation, and by means of a stop watch measure the time of ten complete vibrations, a *complete vibration (or cycle) being composed of a sweep up and a sweep down.* The time observed divided by ten is the *time required for one complete vibration and is called the* ***period*** *of the vibrating system. The number of complete vibrations which take place during the time of one second is called the* ***frequency*** *of the vibrating system.* If n represents the frequency in cycles per second and t the period in seconds per cycle, then

$$n = \frac{1}{t} \quad \text{and} \quad t = \frac{1}{n} \tag{8.2}$$

Factors affecting the period of an oscillating spring. Keeping all else the same, let us replace the 100-gram mass with a 400-gram mass, and then measure the period as before. We find that the period for the 400-gram mass is exactly twice that for the 100-gram mass. Continuing the experiment, let us replace the 400-gram mass with a 900-gram

mass. We find the period to be three times that found for the 100-gram mass. From this it is clear that the period of the system depends upon the mass of the body in a very exact and definite manner. The period is doubled when the mass is made four times as great, and tripled when the mass is made nine times as great. This type of relationship between period and mass is stated thus: *The period varies directly with the square root of the mass.*

The period may also be modified by the kind of spring used, because if equal masses are attached to several springs of varying stiffness, the one with the greatest stiffness will have the shortest period and the one with the least stiffness the longest period. Careful measurements on such vibrating systems lead to the conclusion that *the period varies inversely with the square root of the stiffness of the spring.* This means that if two springs have the same mass attached, the one with four times the stiffness will have one-half the period.

From all this it is clear that the period of the oscillating spring varies directly with the square root of the mass and inversely as the square root of the stiffness of the spring. The exact relationship expressed in the language of algebra is

$$t = 2\pi\sqrt{\frac{m}{f}} \tag{8.3}$$

where t is the period measured in seconds, m is the mass measured in grams, and f is the stiffness of the spring measured as the force in dynes required to stretch the spring one centimeter.

Phase. Suppose that we wish to contrast the nature of the vibration of two oscillating bodies. The vibrations of the two bodies may differ in amplitude and also in period. But even with the same period, they may differ in the time at which the masses move up through the zero position. To illustrate this, we may perform the following experiment:

Let the weights on two similar springs be adjusted until they have exactly the same period, and have each pulled down and released at the same instant. The two vibrating systems are said to be *in phase* because *the weights pass through the zero positions in the same direction, at the same time.* When one weight is pulled down and the other pushed up, both held at rest, and then at exactly the same instant allowed to go free, the two systems are not in phase because they pass through the zero position in opposite directions at the same time. One of these weights is said to lead or lag the other, depending upon which is taken as the reference system (Fig. 88).

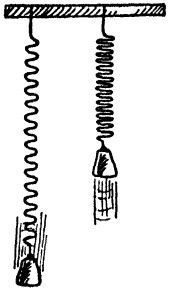


Fig. 88. Out of Phase 180°.

It is customary to measure phase in terms of degrees. Thus, the systems in the first case had a phase difference of 0°, and in the second case a phase difference of 180°. By proper technique and careful manipulation, the two oscillating springs may be made to vibrate with a phase difference of 90°, 270°, 360°, or any number of degrees desired.

Sympathetic Vibrations—Resonance

When a child in a swing attempts to set himself in motion by his own efforts, he soon discovers that he may do so only if he establishes a bodily motion of definite rhythm. When a diver prepares to jump from a springboard, he feels the natural vibration of the system, adjusts his bodily movement in rhythm with it, and establishes the required amplitude of motion necessary for a successful dive. As one travels in an automobile over a washboard highway, the period of the up-and-down motion of the wheels depends on the speed of the car. At certain speeds the body of the car takes on exaggerated and uncomfortable amplitudes of

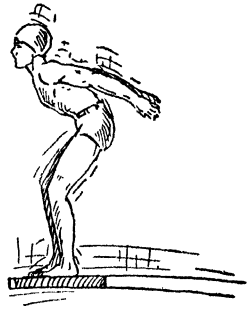


Fig. 89. Ready for a Dive.

vibration. All these are examples of sympathetic vibrations or resonance.

What are the required conditions for resonance? A simple experiment will yield the answer. Let us arrange apparatus as illustrated in Fig. 90. A meter stick is supported at each end with its flat side up. Two vibrating springs, *A* and *B*, which have been brought to the same period by the proper selection of masses, are suspended near each other at the middle of the meter stick. Let us now set system *A* into vibration with a large amplitude. Its motion is more rapidly damped than usual. Small amounts of energy are probably being sent along the meter stick. This is true because system *B* starts to vibrate, its amplitude slowly increasing. Finally *A* comes to rest, and *B* reaches a maximum amplitude. But this condition does not continue for long. *B's* motion begins to show damping and *A* starts vibrating again. Soon *B* comes to rest and *A* reaches a maximum amplitude. Thus energy radiated by one is absorbed by the other, and the receiver in turn passes it back to the giver. The process continues in this cyclic manner until finally both systems come to rest. The friction of the air, the twisting of the springs, and the bending of the meter stick have finally effected a complete transfer of mechanical energy into heat.

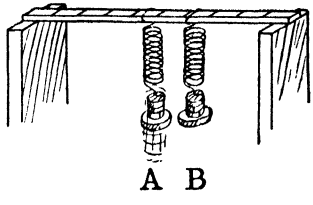


Fig. 90. Illustrating Resonance.

Next let us remove roughly half of the mass from spring *B*, and then set *A* into vibration with a large amplitude. Its oscillations are not damped as much as formerly, and even after considerable waiting we find that *B* does not vibrate. This means that, with its new period of vibration, *B* is not able to absorb energy as before.

Finally, we conclude that if the period of agitation given to the meter stick by one vibrating spring corresponds exactly to the period of vibration of the other spring, energy

is transferred in a most successful manner and resonance is achieved. In general, resonance will be noted: first, when the agitations of one vibrating system are imposed directly, or indirectly through a coupling agent (the meter stick in the case described), upon another vibrating system; second, when the period of vibration of the agitating system corresponds exactly to the period of vibration of the system to be agitated. Some response may be obtained if the systems are "off tune," but maximum response is possible only if the periods of vibration are exactly equal.

Automobile on a washboard road. The main mass of an automobile is supported by springs, and thus the car has a natural period of vibration characterized by the stiffness of the springs and the mass they support. As the automobile is driven over a washboard road at different speeds, the wheels are forced up and down with rather definite periods of vibration. At certain speeds these periods will approximate the natural period of the car, and resonance or near resonance will be established. Shock absorbers help to damp out the large amplitudes thus established, but many drivers realize that a decided shift of speed will successfully solve the probem. In this manner vibrations are set up which are not in resonance with the natural vibrations of the car.



Fig. 91. Resonance.

The action just described may be demonstrated in a simple and vivid manner as follows: Select a spring with a period of approximately one second per cycle when supporting a ten-pound weight. If the upper end of the spring is grasped and forced to vibrate up and down, the large mass will respond vigorously when the period of agitation corresponds to that of the system (Fig. 91). A state of resonance is attained. Bring the system to rest. Now when the agitation of the hand is made very rapid, the large mass responds practically not at all.

Forced Vibrations

Forced vibrations may be illustrated by the use of the vibrating system just described. Let the system be made to oscillate and its motion observed long enough for the period of vibration to be definitely sensed. The mass may now be grasped and the system deliberately forced into oscillation with a period *different from its natural period.* The vibrations thus produced are not natural to the system and are said to be *forced* vibrations.

Transverse Vibrations in Rods and Bars

Recall your childhood experience of crossing a small stream on an improvised footbridge. Usually in fear, you hurried across the long, straight sapling, but surely at some time or other you remained long enough to experience the thrill of setting up oscillations with amplitudes great enough to bring you periodically into the water—an impressive demonstration of resonance. This kind of oscillation is called a *transverse* vibration because the movement is a to-and-fro passage *across* the equilibrium axis of the pole. We shall be interested in discovering whether this periodic swing is an example of simple harmonic motion, as was the motion of an oscillating spring.

Let us turn back a few pages and read again the requirements of such a motion. First, we must have inertia; this the footbridge has without question. Second, the bending must give rise to an elastic force which *opposes* the change in shape; this we know is true because it required just such a force to balance your weight. Third, the elastic force must be *directly proportional* to the amount of bending; this we would not know without direct experiment, or without consulting the experiments of others. It is found that most bodies can yield a little to a change of form and then regain their shape. Within this limit of elasticity, Hooke discovered experimentally that any deformation produced is proportional to the force acting. This being the case,

Hooke's law holds for the bending of a pole as truly as it does for the stretching of a spring. This means that for amplitudes within the elastic limit of the pole, the motion was truly simple harmonic. Probably during your enthusiastic moments, the elastic limit was exceeded and the motion ceased to be simple harmonic.

The vibration pattern assumed by the pole may be duplicated and made more vivid by the use of a two-meter stick, supported at each end with the flat side up (Fig. 92a). As before, the equilibrium conformation, with the weight balanced by elastic forces, will be called the *zero* position. If the stick is pulled down at the middle and then allowed to go free, for reasons outlined above its many parts will execute simple harmonic motion, all with the same

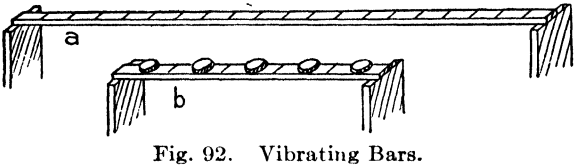


Fig. 92. Vibrating Bars.

period but with amplitudes varying from zero to a maximum. At the points of support the amplitudes are zero, and these positions are called *nodes*. At the middle the amplitude is a maximum, and this region is designated as an *antinode*. At the highest and lowest positions the stick actually comes to rest and the segment between two successive nodes takes on the appearance of a *loop* and is so designated (Fig. 93a).

Factors controlling the period of an oscillating bar. Let a regular meter stick be supported at each end with the flat side up and placed alongside the two-meter stick described above (Fig. 92b). On comparison it will be found that *the shorter stick has a much shorter period.* Now let five or more 100-gram weights be distributed equally along the short stick, thus effectively *increasing its mass per unit length* without changing its stiffness. The period will *show a marked increase.* Finally, let one of the sticks, preferably the longer one, be placed on edge and clamped at each end.

Its stiffness to a downward thrust is *greatly increased* and its period of vibration is *definitely shortened.* These results may be summarized in the following qualitative statement: For the same mode of vibration, a bar will have its period shortened if its length is shortened, if its stiffness is increased, and if its loading is decreased. In general, short, light, stiff bars have shorter periods than long, heavy, weak bars. The quantitative statement may be found in reference (4), Suggested Readings.

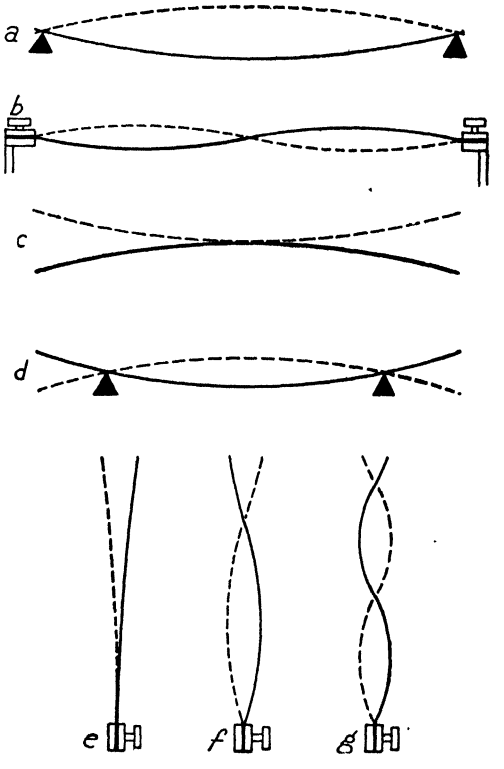


Fig. 93. Vibration Patterns on Bars.

A bar has many modes of vibration. The vibration patterns of a few of the simpler and more common modes of transverse vibration in bars are illustrated in Fig. 93. Mode (a) has already been described and discussed. It is possible to obtain mode (b) by clamping the two-meter stick, or better, a thinner slat, loosely at both ends, holding it at the middle with one hand, pulling down on it one-fourth the distance from an end, and then allowing the system to oscillate. To illustrate mode (c), grasp the thin slat at the middle and then impart rhythmic impulses to it. One soon gets the "feel" of the stick, resonance is achieved, and the vibration pattern becomes clear and definite. Mode (d) may be demonstrated with the hand in the same position as for (c). It is necessary only to move the hand through rather large amplitudes and to shorten the period of vibration. Reso-

nance is soon obtained and the vibration pattern clearly appears. One can illustrate the same mode of vibration even better by supporting the stick at points 0.22 the length of the bar from each end. A slight depression of the middle will cause the system to oscillate with its vibration pattern vividly portrayed. To illustrate mode (e), clamp the stick in a vise, pull the free end aside, and allow it to go free. One can also demonstrate it by grasping the end of the stick in the hand and then imparting to it rhythmic impulses which by the "feel" are known to be in resonance with this mode of vibration. The vibration patterns of (f) and (g) may be obtained in a similar manner, except that it is necessary to decrease the period of the rhythmic pulses as the pattern becomes more complex.

From the study of all of these vibration patterns, we observe: first, a node always exists at a fixed end; second, an antinode always exists at a free end; third, the period of vibration in a general way decreases as the distance between two adjacent nodes is shortened. In the case of vibrating bars, the exact relationship between the period of vibration and the distance between adjacent nodes is very complex.

The xylophone. The bars of the xylophone vibrate with the pattern shown in Fig. 93d, the points of support being located at the nodes. The set of bars of different lengths makes possible a series of definite vibrations. The period of each bar is controlled in the main by its length. If the elements are struck in order, a musical scale is heard, and we have evidence that sounds may have their origin in the mechanical vibrations of bars. A homemade xylophone may be constructed out of pieces of straight-grained wood, flat rocks, or plate glass. The simple rule given above may be used in adjusting the period of vibration (the pitch, as we shall see later) of each bar. For example, in the wood instrument, to shorten the period of a bar, shorten the length of the bar; to lengthen the period, cut out material from the middle portion of the bar.

The tuning fork. Ernst F. F. Chladni (1756–1827) suggested that the tuning fork could be considered as having developed from the bar of the xylophone. The evolution of such a bar, illustrating the shift of the nodes as the fork takes form, is shown in Fig. 94a. Notice that during vibration the prongs move toward and then away from each other (Fig. 94b).

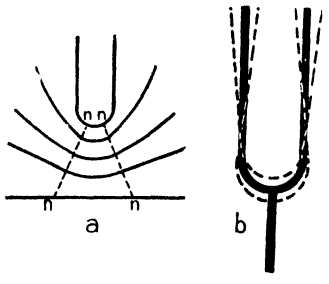


Fig. 94. Tuning Fork.

The factors which affect the period of a vibrating bar also control the period of a tuning fork. Specifically, if the prongs are made lighter by filing on the ends or sides, the period is decreased; if the yoke is made weaker by filing near it, the period is increased. As a temporary measure, wax may be added to the ends of the prongs to increase the period of vibration. The tone produced by the fork is evidence that the mechanical vibrations of a tuning fork may be the source of sound.

Reeds. The reeds of musical instruments are essentially bars clamped at one end, and the vibration pattern, in its fundamental form, is that shown in Fig. 93e. The period of vibration is shortened by a shortening of the reed or a reduction of its cross section at the free end. The period is lengthened by a loading of the free end or by a reduction of the cross section near the clamped end.

Transverse Vibrations in Plates

We can make a plate vibrate by striking it with a mallet or by stroking its edge with a violin bow. From our experience with oscillating bars, we might anticipate that plates also have definite vibration patterns with nodes and loops. Ordinarily, in small plates the oscillation is not easily visible; but if salt or sand is scattered freely over the plate's surface, the small particles move from the regions of motion and pile up along the lines of no motion. Thus, if a square plate is firmly clamped at its middle, covered with salt,

and stroked on an edge with a violin bow, the particles of salt are seen to pile up in a series of lines, indicating the presence of nodes (Fig. 95). The shrill note emitted is striking evidence that sound may have its origin in vibrating plates. By touching the plate at definite points, thus predetermining the location of nodes, and bowing an edge in a certain region, thus predetermining antinodes, many figures of artistic design, known as *Chladni's figures*, may be produced. In all the vibration patterns, we note that the shortest average distance between nodes (the shortest period of vibration) appears when the highest-pitched notes are emitted.

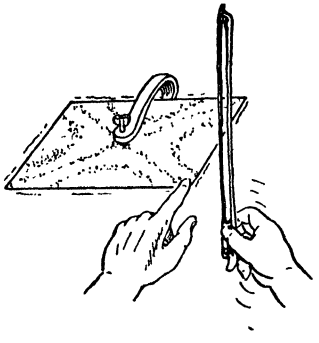


Fig. 95. Chladni's Figures.

Transverse Vibrations in Strings

When the middle of a long cord or elastic spring, fixed at both ends, is pulled to one side and allowed to go free, a vibration pattern consisting of one loop with a node at each fixed end takes form. The motion is simple harmonic, the tension in the cord being the required type of force. If now one end of the cord is held in the hand and rhythmic up-and-down movements of the proper period are imparted, resonance will be established and the same vibration pattern will be displayed (Fig. 96a). If the rhythmic movements of the hand are impressed at two, three, and then four times the rate, a series of patterns illustrated in Fig. 96b, c, and d are obtained.

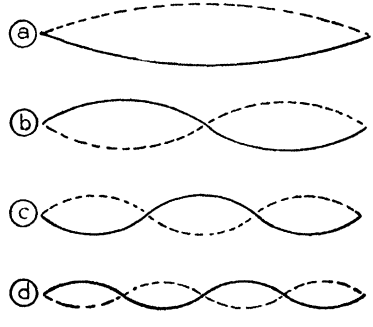


Fig. 96. Vibrating String.

Fundamental and overtones. In every case, the cord is divided up into equal segments with vibration rates or fre-

quencies in the ratio of the whole numbers, 1:2:3:4, and periods in the ratio $1:\frac{1}{2}:\frac{1}{3}:\frac{1}{4}$. Thus, the fundamental vibration, or simply the *fundamental*, is related in a very simple manner to the more rapid vibrations, the so-called *overtones*. When overtones stand in this simple relationship to the fundamental, they are called *harmonics*. (The fundamental is the first harmonic of the series.) In the case of bars clamped at both ends (Fig. 93a and b) or free at both ends (the xylophone type), overtones are in the ratio of 1:2.756:5.404:8.933; for bars clamped at one end (the reed type, Fig. 93e, f, and g), the overtones are in the ratio of 1:6.267:17.55:34.39. At once it is clear that the overtones in bars are not in a *harmonic series*, as are the overtones in strings. The overtones in bells likewise are usually not harmonic.

Complex vibrations. With a little care a long elastic spring may be made to vibrate as a whole (in a fundamental) and in halves (in a first overtone) at the same time. Thus, the oscillations of a certain particle of the spring will be, not a simple harmonic motion, but the result of compounding two such motions. On the stretched strings of musical instruments, it is the usual occurrence to have the fundamental and many of its overtones present at the same time—a harmonic series of simple harmonic motions compounded into a complex motion. The tones produced by vibrations in a harmonic series sound well together, as we shall see in a later chapter. The overtones in bells are usually not in a harmonic series, a good bell being one in which the tones produced by the most prominent vibrations give a pleasing effect when sounded together.

A traveling deformation. To interpret our experiences with bars, plates, and cords in which a slight up-and-down motion at one end of a body sets the whole body into oscillation requires an explanation of how energy is transferred from one end of an elastic body to the other. That a quick up-and-down motion actually travels along an elastic body

may be demonstrated by a simple experiment. Let one end of a coil spring be fixed to a rigid support and then pulled tight. With the spring at rest, let a little hump be quickly produced near the end held by the hand. This deformation will be seen to travel to the far end and return, probably making a number of round trips before being completely damped out. Let the spring as a whole now be set into its fundamental mode of vibration, and *during the oscillation* let a little hump again be made near the end. *The deformation will be seen to travel down and back in exactly the time required for one complete vibration* (Fig. 97).

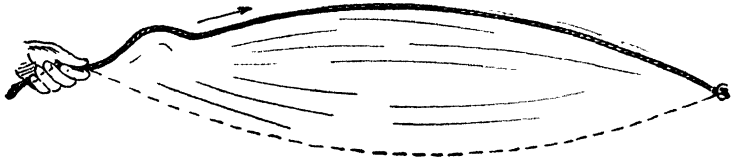


Fig. 97. Comparing Time of Wave Travel with Time of Vibration.

This coincidence of action and the fact that it is possible to establish vibrations by feeding small amounts of energy into one end of a spring suggest that the vibration patterns produced are the results of wave motion. This problem, the problem of so-called *standing* waves, will be investigated in the next chapter.

Longitudinal Vibrations in Rods

When a rod is stretched in the direction of its length, opposing forces which are directly proportional to the change in length arise. Since the conditions for simple harmonic motion are achieved, we have reason to believe that the various thin slices of the bar, when displaced and allowed to go free, will move forward and back along the rod in simple harmonic motion. *This kind of oscillation is called a* ***longitudinal vibration.*** That such vibrations actually exist may be shown as follows: When a six-foot brass rod is grasped firmly at its middle and gently tapped transversely near its end, a distinct tone, which is due to a

transverse vibration pattern of the type illustrated in Fig. 93c, is heard. Now if the rod is gently struck *at the end in a direction along the bar*, a very much higher-pitched note will be heard (Fig. 98). This is obviously the *longitudinal vibration* which we anticipated. Since the ends are free and the middle of the rod is held by the hand, an antinode (maximum longitudinal motion) exists at each end and a node is located at the middle. Other longitudinal vibration patterns may be obtained, as may be shown if the rod is grasped at two positions, each located one-fourth the length

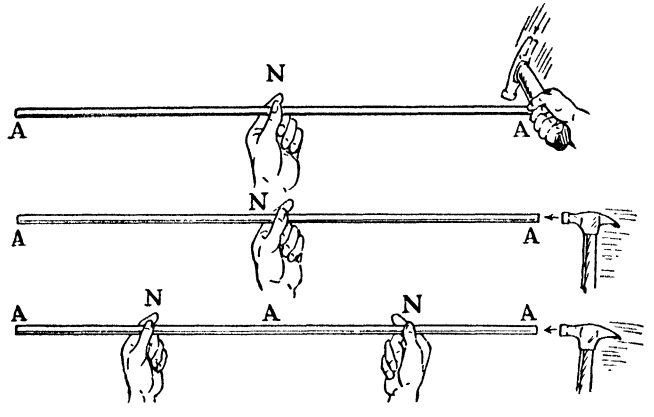


Fig. 98. Vibrating Rods.

of the rod from an end. A still higher-pitched note is heard when the rod is gently struck as before. In this new vibration pattern an antinode is located at each end, and two nodes are stationed approximately at the two points of support.

Again we seek wave motion as a recourse to explain how an impact at one end produces a general vibration of a rod. Finally, we have evidence that sound may have its source in the longitudinal vibration of rods.

Vibrating Air Column

It has been shown conclusively that *sound has its origin in mechanical vibrations.* But since air is the elastic medium usually connecting the source of sound with the

ear, we shall be interested to discover if air may be set into oscillation. Evidence that longitudinal vibrations may be produced in air is easily obtained by the use of Kundt's apparatus (Fig. 99). This apparatus is essentially a vibrating steel rod coupled to an air column. The free end of the rod is stroked with a resined cloth, longitudinal vibrations are thus established, and a disk on the other end of the rod gives a series of very rapid, rhythmic impulses to the end of the air column with which it is in contact. If lycopodium powder or fine cork dust is spread along the

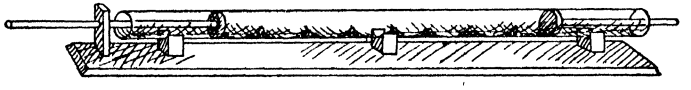


Fig. 99. Kundt's Apparatus.

tube, the presence of nodes may be detected in much the same manner as they were located by the use of a covering of salt on a vibrating plate. Actually, when resonance is established through the adjustment of the length of the air column, the powder is brought into a series of little piles located at equal intervals along the tube. It is possible to make the response of the cork dust to the vibrating air column exceedingly sensitive by turning the tube till the cork dust is just about ready to fall from its position on the side of the tube. This experiment furnishes evidence that under proper stimulus, achieved most effectively at resonance, the air column takes on a vibration pattern with nodes and loops evenly spaced, a node being located at the rubber stopper which forms the fixed end of the tube and an antinode being located at the vibrating disk. Again we need wave motion to explain how a series of rhythmic impulses on one end of an air column may set it into a general vibration. Air waves would also explain how the energy of vibrating bodies may be brought to the ear, a phenomenon so essential to hearing. Wave motion will be discussed in the next chapter.

Questions and Problems

1. List five examples of simple harmonic motion.

2. When the amplitude of vibration of an oscillating spring is halved, what change has taken place in the energy of the system?

3. If the period of an oscillating body is one-half of a second, what is its frequency of vibration?

4. List five examples of (*a*) sympathetic vibrations; (*b*) forced vibrations.

5. Describe how a xylophone can be made from plate glass.

6. Explain how you would decrease the period of a tuning fork.

7. Explain how a rod may be made to oscillate (*a*) with transverse vibrations; (*b*) with longitudinal vibrations.

8. Draw two (fundamental and first-overtone) vibration patterns of an oscillating string, and label nodes, antinodes, and loops.

Suggested Readings

(1) Miller, D. C., *Science of Musical Sounds*, The Macmillan Company, New York, 1916, Lecture I.
(2) Saunders, F. A., *A Survey of Physics*, Henry Holt and Company, Inc., New York, 1930, Chap. XVI.
(3) Tyndall, *Sound*, D. Appleton-Century Company, Inc., New York, 1875, Chaps. III–V.
(4) Wood, A. B., *A Textbook of Sound*, The Macmillan Company, New York, 1930, Section II.



CHAPTER IX

Waves

Waves are so fascinating that when we are in the presence of a placid lake and a pebble-strewn beach, we have an urge to throw a stone through the unruffled surface. The stone goes into the water with a splash. Beginning at the point of entrance, a short wave-train moves out in a series of ever expanding circles. When a group of such wave centers are produced by the dropping of a number of stones through the surface, the circular waves move freely through each other and over the surface without a distortion of form.

Such water waves may be demonstrated in the laboratory by the use of a ripple tank—a glass-bottomed, shallow tank with an area of at least four square feet covered with a thin layer of water about one-fourth of an inch thick. A naked arc light, or a clear incandescent lamp of high power, is placed beneath the tank. Ripples produced on the water surface are shown as light and dark lines on the ceiling above.

What Travels in a Wave?

John Tyndall (1820–1893) had extraordinary ability for simplifying the presentation of difficult subjects. As a demonstration of wave motion, he suggested an experiment of the following kind. Let 20 or more persons stand in a single file, each person placing his hands on the shoulders of the one in front of him. A forward thrust exerted by the end man will then be felt progressively through the whole group. A person in the column feels the compression go by, and an observer sees a state of disturbance move from one end of the column to the other.

Obviously, the disturbance felt or seen traveling along the column is not the moving of a person through the full distance, but the passing of a forward thrust along the column at a definite speed. The slight forward displacement of the first person is passed on to the next person, and by him in turn to his neighbor. Thus the state of pushing goes from one end of the column to the other. That which travels in a wave, therefore, is not the material composing the medium, but rather a disturbance in it. This simple demonstration is an example of a *longitudinal* wave, because the slight displacements of the persons in the column are in the same line as the direction of the travel of the wave.

A *transverse* wave, one in which the particle displacement is back and forth across the line of wave travel, may be demonstrated by the use of a long, heavy rope laid along the full length of a table. While one end of the rope is hanging freely over the far edge of the table, the other end is grasped and moved transversely back and forth at a rapid rate. A wave moves down the rope in an impressive manner, but because of friction between the rope and the table top, the amplitude diminishes and the wave will probably be damped out before it reaches the far edge (Fig. 100). Thus, the effect of reflected waves is eliminated.

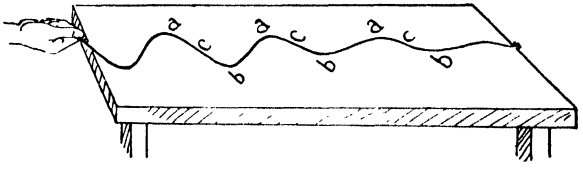


Fig. 100. A Transverse Wave.

Wave Length

Water waves furnish the most common wave pattern of our experience, and no doubt you know already what is meant by the *crest* of a wave. The depressed region between the crests is called a *trough.* An important observation to be made of any wave motion (this is clearly shown in the waves just produced on a rope) is that the

particles of the medium vibrate with the same frequency as the oscillating system that sets the wave going. The particles of the medium (the rope) located at any instant in the various crests (points *a*, *a*, *a*, Fig. 100) are at that moment in those parts of the wave which were sent out periodically by the oscillating system (the hand) as it reached its most extreme position to the left. All these particles are at that instant ready, along with the oscillating system, to start the trip to the right, and thus they are all *in phase.* For the same reason, the particles in the troughs (points *b*, *b*, *b*), or any set of symmetrically located particles, such as, for example, *c*, *c*, *c*, are also in phase; but the particles between are not in phase with those at the end position, the phase difference varying from 0° to 360°.

The *distance* between two adjacent crests or two adjacent troughs or, generally, for longitudinal as well as transverse waves, *between any one particle and the next one vibrating in the same phase is called a* ***wave length.***

Frequency of vibration, wave velocity, and wave length. Each wave length λ of a train of waves is sent out during one complete vibration of the oscillating body. Thus, during *one second* a body making n oscillations per second sends out n waves, and the length of this wave train is $n\lambda$. But this length is also the distance the wave travels in one second, or simply the magnitude of the velocity, V. Hence,

$$V = n\lambda \tag{9.1}$$

and we have a simple relation between the velocity V of a wave, the number of vibrations (cycles) per second n of the particles of the medium, and the distance λ between any one particle and the next one vibrating in the same phase.

Speed, Reflection, and Interference of Waves on Ropes

Speed of waves on ropes. Let two 30-foot ropes, of the same material but one twice the diameter of the other, be stretched side by side between two supports, the tension on each being measured by a spring balance inserted at one end

(Fig. 101). If the shape of one rope is changed near the end, a hump-like deformation will travel down the rope and back a number of times. A wave pulse is used instead of a train of waves because its exact location is much easier to follow with the eye, and thus the speed of the wave is determined with greater ease. A few simple tests soon reveal clearly that *the speed of the wave increases with an increase of tension and decreases with an increase in linear density (mass per unit length).*

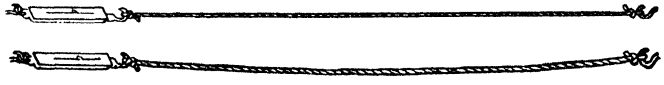


Fig. 101. Speed of Waves on Ropes.

Proceeding in a more definite manner, let the tension on the two ropes be made the same. Then, if at the same instant a hump is started down each rope, one will travel *down* to the far end of the larger rope while the other goes *down and back* on the smaller rope, the speed of the wave on the former being one-half that on the latter. But the larger rope has twice the diameter and hence four times the cross section and linear density. Thus we conclude that, *under the same tension, the speed of a wave on a rope is halved if its linear density is made four times as great.* Next, let the tension on the *larger* rope be made four times that on the smaller one. The speeds are now found to be equal, and we conclude that, *with a four-times increase in tension, the speed of a wave on a given rope is doubled.* These results agree with the experimentally tested relation,

$$V = \sqrt{\frac{T}{d}}, \tag{9.2}$$

in which V is the speed of the wave in centimeters per second, T is the tension in dynes, and d is the linear density in grams per centimeter.

Reflection of waves. *At a fixed end.* If a long coiled spring with one end fastened to a rigid support is grasped and stretched moderately tight, a hump-like distortion

produced near the hand will quickly sweep down the coil and be reflected back. Let special attention be given to the form of the reflected deformation. The hump is reflected as a hollow (Fig. 102). Actually, then, the reflection is more than an inversion, a sort of right-about-face action; it consists also of a complete reversal of phase, because the hump causes the particles to take on a displacement *above* the zero position as it passes, and the hollow causes a displacement *below* this position as it goes by.

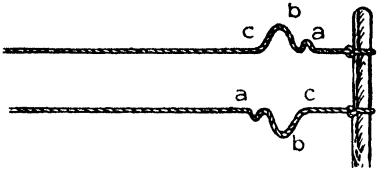


Fig. 102. Reflection at Fixed End.

At a free end. Let the coiled spring now be connected to a rigid support by means of a few feet of string. The end of the coil is thus virtually free. A study of the form of the reflected deformation reveals that a hump is reflected as a hump, there being an inversion but no change of phase (Fig. 103).

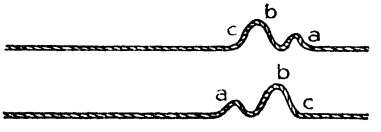


Fig. 103. Reflection at Free End.

Interference of waves. Let a series of humps and hollows—a short wave train produced by three or four rapid to-and-fro movements of the hand—be sent along a large rope under slight tension. The forward-moving waves in no way obstruct the returning ones. This ability of a rope to carry at the same time waves traveling in different directions is possessed by all media unless the disturbance is too violent.

When the forward and returning waves pass a given rope particle simultaneously, the particle experiences the equivalent of two displacements, an actual displacement which is the algebraic sum of the two. If the displacements are equal and opposite, the particle remains at rest; if they are equal and concurrent, the amplitude is doubled. In the first case, the interference is completely destructive; in the second case, its re-enforcing action is at a maximum. In general, the *destructive* or the *re-enforcing*

action of two or more waves acting simultaneously on a particle is called *interference.*

Interference may be vividly shown by the following experiment: Let two ten-foot ropes and one two-foot rope be tied together at a point A, and have the short rope fastened to one end of a table. Then, standing at the opposite edge, grasp the end of a long rope in each hand and gently tighten the whole system so that each rope is under the same tension but is still resting on the surface of the table (Fig. 104). Now, when the right hand is moved right and left in rapid rhythmic motion, a wave moves down the table. With the right hand still oscillating, let the left hand begin a right-and-left motion with exactly the same frequency and no difference in phase. The hands move right and left together and the point A becomes more agitated. At the point A, the interference has resulted in re-enforcement. Next, let the system be brought to rest and the right-and-left motion of the right hand started. The point A again oscillates. Now let the left hand move with the same frequency but with a phase difference of 180°. The hands move toward each other and then away from each other. Point A is brought to comparative rest; destructive interference has destroyed its motion.

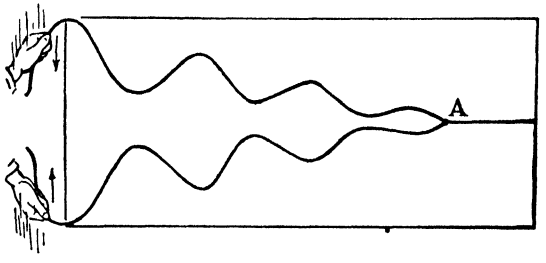


Fig. 104. Destructive Interference.

It is possible to demonstrate destructive interference very clearly by substituting the prongs of an electrically driven tuning fork for the moving hands and a silk floss for the ropes.

Standing Waves

Recall the vibration patterns described in the last chapter. At that time we expressed the need of wave

motion to explain how rhythmic impulses supplied at one end of a coiled spring or rope could finally set the whole system into oscillation. With our added knowledge of waves, let us reconsider the problem.

As before, let one end of a rope be fastened to a rigid support and the other grasped by the hand. As the hand is moved from the zero position up and back again, a hump is sent along the rope, and all the particles through which it passes have displacements above the zero line (Fig. 105a, b, c, d, and e). As the hand is moved from the zero position down and back, a hollow is sent out and all the particles through which it passes have displacements below the zero line (Fig. 105e, f, g, h, and i). Thus, during each complete vibration a hump and a hollow are sent along the rope.

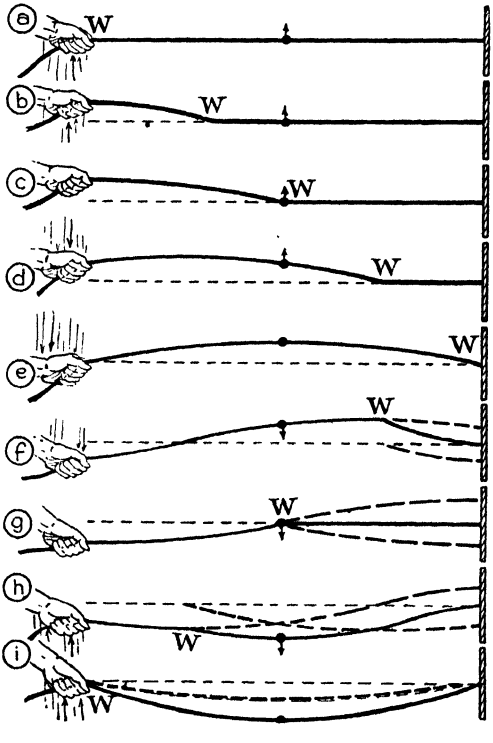


Fig. 105. Interference Producing a Standing Wave.

When a hump is reflected from the fixed end, a complete reversal of phase results, and the hump returns as a hollow. If the timing of the hand is just right, when the outgoing hollow meets the reflected hollow, their maximum amplitudes will be at the middle of the rope at exactly the same instant (Fig. 105i). In order to do this, the hand must make one complete vibration while the wave moves down the rope and back. In the figure, *W* marks the front of the wave. You will notice that *W* moves down the rope and back during one complete vibration of the hand. With this

adjustment, the forward wave and the returning wave will continue to be in phase at the middle of the rope, and the resulting amplitude at this position will continue to be the greatest along the rope. From this position of maximum amplitude toward either end of the rope, the phase difference changes from 0° to 180°; and the interference changes from a complete re-enforcement at the middle of the rope to a complete destructive action at its ends. The net result is the loop of the vibration pattern. It follows, therefore, that if energy is supplied with just the proper rhythm, an antinode will definitely appear at the middle of the rope and nodes will be present at the ends. Thus, the rope's oscillation results from the interference of oppositely directed wave motions. The vibration pattern produced, although not really moving like a wave, is often called a *standing* wave in reference to its origin.

To establish the fundamental mode of vibration, it was only necessary to make a complete vibration with the hand while the wave moved twice the distance between nodes, down the rope and back (Fig. 105). But by definition this distance of wave travel during one complete vibration (or cycle) is a wave length. Thus, *the distance between two adjacent nodes is one-half a wave length.* If L is the length of the rope (in this case, the distance between two adjacent nodes) and λ is the wave length, it follows that $\lambda = 2L$. From Equation (9.1) $V = n\lambda$, and solving for n, we have $n = V/\lambda = V/2L$. Making use of Equation (9.2) we obtain

$$n = \frac{1}{2L}\sqrt{\frac{T}{d}}, \tag{9.3}$$

as the frequency of the fundamental mode of vibration of a wire or rope with length L, linear density d, and tension T. We shall need this relation in the study of stringed instruments.

When the hand is moved up and down with twice the frequency, the wave length will be half as long. In traveling the full length of the rope, the outgoing wave will now consist of a hump followed by a hollow, and the reflected wave will consist of a hollow followed by a hump. The hollows will meet in phase at one-fourth the length of the rope from the hand, and the humps will meet at the same distance from the other end. Thus, the timing will be just right for the outgoing and reflected waves always to be in phase at these points; therefore, antinodes will form at these positions. At the middle of the rope, displacements will be entirely eliminated because of destructive interference, and a node will result. In the same general manner, the standing wave patterns of the other overtones may be easily interpreted and described (Fig. 96).

Had the end of the rope been free instead of being fixed, the reflection would have taken place without a reversal of phase, the interference would have been an act of re-enforcement, and an antinode would have formed at the free end.

The interpretation of the vibration of an oscillating rope in terms of wave action may be extended to the various modes of vibration found in bars, rods, and air columns. The vibration patterns of bars and rods have been discussed adequately (in Chapter VIII), but the standing waves found in air columns should receive further consideration.

Standing waves in air columns. *Air column, one end closed.* The standing waves formed in a column of air, closed at one end, must be of such a pattern as always to give a node at the closed end, where reflection takes place with a reversal of phase, and an antinode at the open or free end, where reflection takes place *without* a change in phase. The simplest, and therefore the fundamental, mode of vibration is attained with a node at the closed end and an antinode at the open end. The more complex modes, the overtones, are formed by an alternate sequence of

nodes and antinodes, with a node always at the closed end and an antinode at the open end (Fig. 106).

Just as in the case of oscillating ropes, the distance between two adjacent nodes in an air column is one-half a wave length, and the distance between a *node* and an adjacent *antinode* is one-fourth a wave length. Thus, in the above series of fundamental and overtones, the *length of the column* is successively $\frac{1}{4}$, $\frac{3}{4}$, $\frac{5}{4}$, $\frac{7}{4}$, and so forth, of the wave length of the vibration which is characteristic of that particular mode. The wave lengths then bear the ratio $1:\frac{1}{3}:\frac{1}{5}:\frac{1}{7}$, . . . The velocity of wave travel in air is constant for these modes of vibration, and hence the frequency of a vibration varies inversely with its wave length ($V = n\lambda$ and $n = V/\lambda$). It follows, therefore, that frequencies of the series bear the ratio 1:3:5:7:9 . . .

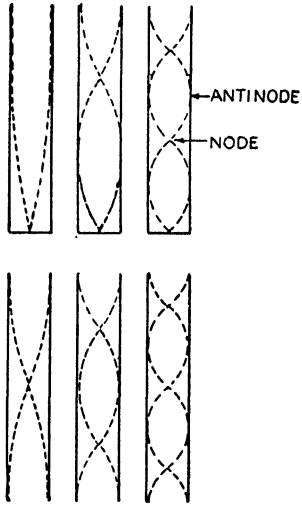


Fig. 106. Standing Waves in Air Columns.

Air column, both ends open. In this type of air column, the wave is reflected from open ends and, therefore, without a reversal of phase. Hence, an antinode is formed at each end. The fundamental mode of vibration is attained with an antinode at each end of the column and a node at its middle. The more complex modes, the overtones, are formed with an alternate sequence of antinodes and nodes. An antinode always appears at each open end (see Fig. 106).

In the series of fundamental and overtones produced in this air column, the length of the column is successively $\frac{1}{2}$, $\frac{2}{2}$, $\frac{3}{2}$, $\frac{4}{2}$, $\frac{5}{2}$, etc., of the wave length of the vibration having that particular mode of oscillation. The frequencies of this series stand in the ratio 1:2:3:4:5:6:7 . . . We shall need this knowledge of standing waves in air columns as we study organ pipes and wind instruments.

SUMMARY OF STANDING WAVE PATTERNS

TRANSVERSE VIBRATIONS

Kind of Vibrating Body	*Ratio of Fundamental and Overtones*	*Nature of Series*
Bars:		
Free or clamped at both ends.	1:2.756:5.404:8.933	Inharmonic.
Clamped at one end.	1:6.267:17.55:34.39	Inharmonic.
Wires, ropes, etc.:		
Flexible and fixed at ends.	1:2:3:4:5:6:7:8....	Harmonic.

LONGITUDINAL VIBRATIONS

Kind of Vibrating Body	*Ratio of Fundamental and Overtones*	*Nature of Series*
Rods:		
Free at both ends.	1:2:3:4:5:6:7:8....	Harmonic.
Clamped at middle.	1:3:5:7:9:11:13....	Even harmonics missing.
Air Column:		
One end closed.	1:3:5:7:9:11:13....	Even harmonics missing.
Both ends open.	1:2:3:4:5:6:7:8....	Harmonic.

Waves in Free Air

We have shown that the standing wave patterns, known to exist in an air column (Kundt's tube experiment), may be interpreted if we assume that longitudinal air waves travel forward and back in the tube. Let us extend the picture to free air.

If a small spherical ball expands and contracts in a rhythmic manner, the air immediately surrounding it will first be compressed and then be allowed to expand. We know that longitudinal waves may exist in air. Therefore, if the velocity of the wave is the same for all directions and no objects obstruct the path, the compressional disturbance will pass from air layer to air layer in an ever-expanding

spherical wave front. The compressional disturbance will be followed immediately by a state of expansion. This will be followed by another compression and this in turn by another expansion. An instantaneous picture of the process is illustrated diagrammatically in Fig. 107.

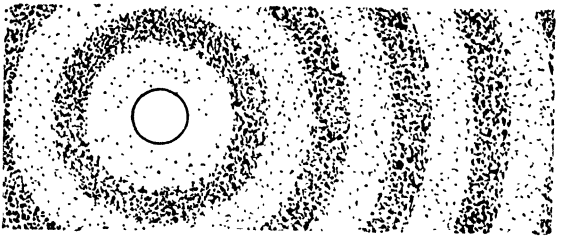


Fig. 107. Illustrating Waves in Air.

The region of each compression in the wave train is called a ***condensation*** *and the region of expansion is called a* ***rarefaction.*** These regions travel outward with the speed of the wave, and an air molecule in the path experiences in turn the effect of a condensation, a rarefaction, a condensation, another rarefaction, and so on. During one complete vibration, the pulsating ball originates a condensation and a rarefaction. Thus the distance from one condensation through a rarefaction to a similar position in another condensation is also the distance that the wave travels during one cycle, and therefore a wave length.

The idealized pulsating ball would be difficult, if not impossible, to construct; but the open end of a long tube, in which standing waves are maintained by some driving mechanism such as a telephone receiver coupled to the other end, makes an excellent substitute for the ideal point-source.

The first attempt to measure accurately the speed of waves in air was made in 1738 by a commission of the French Academy of Sciences. Two parties stationed about three miles apart measured the time between the flash of a cannon and the sound of the report. To eliminate the effect of the wind, observations were taken first in one direction and then in the other. But the most reliable results are those which have been obtained in large underground pipes, such as the sewers of Paris, and in the labora-

tory, where variations of temperature and wind may be avoided. The speed of air waves at 0° centigrade is found to be 1,088 feet per second and at 20° centigrade, 1,129 feet per second.

A theoretical study of longitudinal waves in air leads to this formula for the speed of wave travel:

$$V = \sqrt{\frac{1.40P}{d}}, \tag{9.4}$$

in which V is the speed of the wave in centimeters per second, P is the air pressure in dynes per square centimeter, and d is the density of the air in grams per cubic centimeter. Notice the similarity between this equation and the one we found for the speed of waves on ropes and wires (Equation 9.2). Air pressure is directly proportional to air density if the temperature is kept constant; hence, it follows from Equation (9.4) that *the speed of air waves is independent of the barometer reading. However, the speed is affected by temperature changes, increasing approximately two feet per second for each degree centigrade rise in temperature.*

This large variation of the speed of air waves with temperature may readily be interpreted in terms of air molecules. We have already shown that the molecules of a gas are in rapid random motion. As a wave passes, a directed motion is superimposed upon the irregular motions already existing among the molecules. The wave, therefore, may be passed along from molecule to molecule no faster than the molecules themselves move in this direction. But, as we know, the kinetic energy of the molecules is directly proportional to the absolute temperature, and their average speed directly proportional to the square root of the absolute temperature. When these facts are applied to waves in air, the change of speed with temperature is quantitatively accounted for.

The following table permits a comparison of the speeds, expressed in feet per second, of longitudinal waves in various common substances:

Air 0°C	1,088
Carbon dioxide 0°C	846
Hydrogen 0°C	4,165
Water 13°C	4,728
Pine Wood	10,900
Brass	11,480
Brick	11,980
Steel	16,360

Sound Waves

Sound, a wave motion in an elastic material medium. That which we interpret as sound we have definitely shown to have its origin in vibrating elastic bodies. We have anticipated that these vibrations, under most circumstances, are transmitted to the ear by the atmosphere. Finally, to prove this point, we shall perform the following experiment: Let a small bell be suspended freely on the inside of a glass jar by means of rubber bands (Fig. 108). Waves transmitted by rubber are rapidly damped. To be heard, therefore, the bell must transmit its vibrations, for the most part, through the air contained in the vessel. The jar is completely airtight except for the opening produced when a stopcock is opened. When the stopcock is closed and the jar is shaken, the sound of the bell is definitely heard. When the vessel is evacuated and again shaken, a very enfeebled sound is heard. This is final proof that mechanical vibrations transmitted through the air to the ear are interpreted as *sounds*.

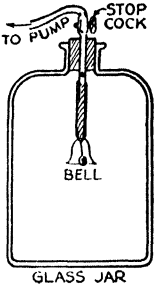


Fig. 108. Bell in Vacuum.

Because of this, the word *sound* has been extended in meaning to include *any wave motion in an elastic material medium.* Hence, all the waves we have considered in this chapter might with propriety have been called "sound waves." The other aspect of sound, the *sense datum* produced through the ear by the *stimulus* of a wave motion, will be considered in the next chapter.

The intensity of sound. As a sound wave travels through a medium such as air, it carries energy with it.

When a person talks into the open from a tall tower, the sound spreads out in spherical waves. The area of a sphere increases four times if the radius is doubled; it increases nine times if the radius is tripled. This means that the surface over which the energy of a spherical sound wave front spreads increases with the square of its distance from the source. The energy, therefore, which flows during one second through a square centimeter varies inversely with the square of the distance from the source, becoming one-fourth as great if the distance is doubled, one-ninth as great if the distance is tripled, and so on. Rarely do we have this ideal condition, but it is often quite closely approximated. In such a free progressive wave, *the intensity of sound is defined as the energy in ergs which flows perpendicularly through one square centimeter during one second.* Thus, for spherical waves, the sound intensity varies inversely as the square of the distance from the source.

Direction of sound travel. In the spherical waves described above, the obvious direction of sound travel is straight out from the sound source on lines perpendicular to the wave fronts. In general, we shall take the line perpendicular to an element of wave front as the direction in which that portion of the wave is traveling. This means that if the wave front becomes more curved or is flattened out because of a change in the velocity of the wave in different parts of its path, the line of travel will be bent and the sound waves will be *refracted.*

Refraction of sound in the open air. The speed with which sound travels in air will often be modified by variations in temperature, water vapor, and wind. The speed of sound increases with the temperature; water vapor is lighter and therefore transmits sound faster than dry air; the speed of sound is increased if a wind is blowing in its direction of travel. Often a change of altitude is accompanied by a change in temperature, air humidity, and wind velocity, and these changes, taken together or singly, may

cause the speed of sound to increase with an increase of altitude. Such a condition is often achieved on a still, foggy night. Under such circumstances, a sound wave emitted from a source S, instead of going forward on straight lines with wave fronts of true spherical form, travels on a curved path, as indicated in Fig. 109. This is so because the increase of speed with altitude causes the spherical wave fronts to be flattened out and finally to become concave toward the listener L, thus roughly focusing the sound upon him.

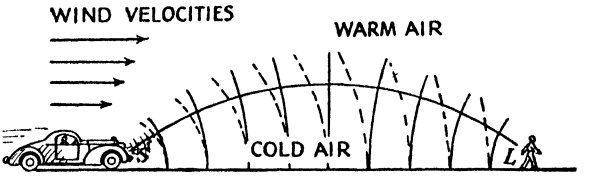


Fig. 109. Refraction of Sound.

If, on the other hand, an opposing wind velocity increases with altitude, or if the stratification of the atmosphere is such that cold layers are above and warm layers are beneath, or if the lower layers of air contain more moisture than those above them, the speed of travel will be greater near the surface of the earth than at higher altitudes. This condition is often reached on clear, warm days. In this case, the sound is bent up away from the earth, making hearing difficult.

Often these special air conditions are contiguous and the sound is first bent up and then down, so that it may be heard at a distance with greater intensity than at a near-by position. If the atmosphere is very inhomogeneous, the sound becomes scattered and loud sounds may be heard only a short distance.

Reflection, transmission, and absorption of sound. Whenever a sound wave in one medium, such as air, meets another medium with different density and elasticity, such as the walls of a room, part of the energy is thrown back as a reflected wave, part is absorbed in the new medium, and part is transmitted. If the transmitted wave does not meet another medium, it will continue till its

energy is completely absorbed. Porous material, such as hair felt, reflects only a small amount of energy but absorbs a rather large amount in its porous channels. On the other hand, solid plaster reflects over 97 per cent of the energy falling on it. This is caused by the marked difference between the elasticity and density of the air and the elasticity and density of the solid material. These phenomena will be investigated more extensively in Chapter XI.

Questions and Problems

1. The speed of a wave on a rope is ten feet per second. If the tension is increased four times, what will be the new speed?

2. A rope is making two vibrations per second and oscillating in its fundamental mode. If its tension is increased four times and its length is doubled, what will be the new frequency of vibration?

3. The fundamental of an air column with one end closed is 100 cycles per second. List the frequencies of vibration of six possible overtones.

4. The fundamental of an air column open at both ends is 200 cycles per second. List the frequencies of vibration of six possible overtones.

5. The distance between two adjacent nodes is ..... wave length.

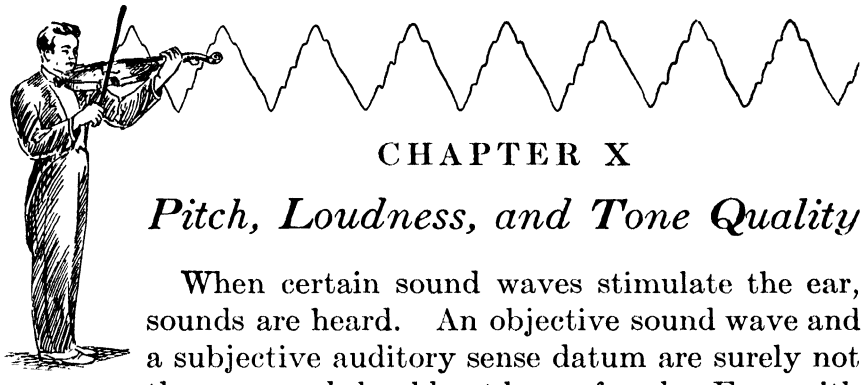
6. The speed of sound in air is independent of the .....

7. A wave length of a sound wave in air is composed of a ..... and a .....

8. Explain how and why the frequency of an air column changes with the temperature.

Suggested Readings

(1) Bragg, W. H., *The World of Sound*, E. P. Dutton and Company, Inc., New York, 1922, Lectures I–VI.
(2) Cajori, Florian, *History of Physics*, The Macmillan Company, New York, 1914, pp. 174–177.
(3) Tyndall, *Sound*, D. Appleton-Century Company, Inc., New York, 1875, Chaps. I and VII.
(4) Watson, F. R., *Sound*, John Wiley and Sons, Inc., New York, 1935, Chaps. I–IX.



CHAPTER X

Pitch, Loudness, and Tone Quality

When certain sound waves stimulate the ear, sounds are heard. An objective sound wave and a subjective auditory sense datum are surely not the same and should not be confused. Even with no ear to hear, the former no doubt exists in a forest when a tree falls, but the latter certainly does not. Yet the *subjective* characteristics of sounds and the *physical* characteristics of waves are unquestionably interdependent—the subject to be discussed in this chapter.

Musical Tone and Noise

When a sound is heard, it is interpreted as a tone or a noise. The difference between noise and tone is one of degree. In extreme cases, the two phenomena are clearly distinct, but lack of musical training, fatigue, or the absence of so-called proper standards may cause a person to classify tones as noises and noises as tones. After Wagner's "Tannhauser Overture" had been known to the musical public for ten years, the London *Times* characterized it as "at best but a commonplace display of noise and extravagance."

Probably as good a definition for noise as we can find is the following given by Miller (3)[1]: "Noise is a sound that is of too short duration or too complex in structure to be analyzed or understood by the ear."

In terms of mechanical vibrations, a *pure tone* consists of a pure, simple harmonic vibration. Such a pure tone is emitted by a properly mounted tuning fork. A *tone* is

[1] References given in this manner will be found in the Suggested Readings at the end of the chapter.

made up of a *series of pure tones*, a fundamental and a series of overtones sounded simultaneously. A violin, a piano, and a flute emit tones. Tones, when *heard*, have three characteristics: pitch, loudness, and tone quality. We shall seek the ordinal relationships between the *subjective* and *physical* characteristics of a tone.

Pitch

Pitch and frequency. The pitch of a tone, as usually understood, is the place assigned to it on the musical scale. An acute sound is of high pitch, a grave sound is of low pitch. Let us seek to discover just what physical characteristics of a mechanical vibration give rise to the pitch characteristic of a sound sense datum.

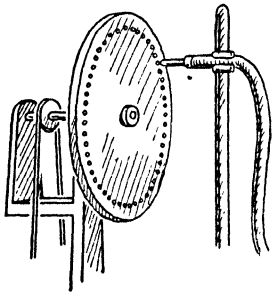


Fig. 110. The Siren.

A siren produces sound by the periodic interruption of a jet of compressed air (Fig. 110). As each evenly spaced hole of the revolving disc passes by, the air goes through, and the number of puffs per second is equal to the number of holes in the circle times the number of turns made by the disc in one second. As rotation is started, distinct puffs are heard; but as the speed increases, a low note, then a higher note, and finally a very shrill sound are heard. This is conclusive evidence that a musical tone is produced by a truly periodic vibration, and further, that *the frequency of this vibration is a physical characteristic which controls pitch.* Finally, let the jet be turned upon a ring of holes which are not evenly spaced. A noise is heard.

By turning the jet upon a set of evenly spaced holes and by properly adjusting the speed of the disc, it is possible to tune the tone emitted to the sound from a tuning fork. The number of holes in the series may be counted; the revolutions per second of the disc may be measured; and hence, the number of puffs produced per second may be calculated and the frequency of the fork determined.

The pitch of the fork, then, may be designated by the frequency of its vibration.

When we study the human ear, we shall describe how the *pitch* of a pure tone, principally controlled, as we have shown, by the *frequency of vibration*, depends slightly upon the *loudness* of the tone (Fig. 122b). The pitch of a complex tone with components in a harmonic series corresponds, as we shall see, very closely to that of the *fundamental* of the series, the change of pitch with loudness being controlled, however, by the joint action of all the harmonics and not by the fundamental alone. For example, as the loudness increases, a pure tone of 200 cycles per second is lowered more in pitch than a complex tone with this frequency as its fundamental. This difference in pitch caused by great loudness does not give rise to discord when these two tones are sounded together. In fact, a tone with a new pitch results. Thus, frequencies rather than pitch should be considered when a harmonious blending of tones is desired.

THE RELATIONSHIP BETWEEN FREQUENCY AND PITCH

Frequency in Terms of a Reference Tone S	*Pitch in Octaves Below or Above the Reference Tone S*
$\frac{1}{32}S$	−5
$\frac{1}{16}S$	−4
$\frac{1}{8}S$	−3
$\frac{1}{4}S$	−2
$\frac{1}{2}S$	−1
S	0
$2S$	1
$4S$	2
$8S$	3
$16S$	4
$32S$	5

Careful experiments with the siren, or with tuning forks of known frequency, lead to the conclusion that if the number of cycles per second of a sound is doubled, an increase of pitch called the *octave* is sensed; another doubling of frequency gives rise to another octave increase; still another doubling gives another octave rise. On the other

hand, halving the number of cycles per second lowers the pitch one octave; another halving lowers the pitch another octave; and so on. This definite relationship between frequency and pitch is illustrated in the preceding tabulation.

Students acquainted with logarithms will at once recognize the exact logarithmic relationship between the frequency of vibration and the pitch. All may get the meaning of the relationship, for it may be stated as follows: *Octave increases in pitch are attained, not by adding up equal frequencies, but by doubling the frequency each time one additional higher octave is required; and octave decreases in pitch are attained, not by subtracting equal frequencies, but by halving the frequencies each time one additional lower octave is required.*

Standards of pitch. Before musicians had standardized pitch in terms of frequency, Rudolf Koenig selected the frequency of 256 cycles per second for middle *C*. The musical scale with this as a standard is called the *Scientific Scale. The International Tempered Scale* gives *A* above middle *C* the frequency of 435 cycles per second, and middle *C* takes on the frequency of 258.65 cycles per second.

Loudness and Intensity

The sound intensity which the ear interprets is that found at the ear. But the sound intensity at the ear increases when the quantity of sound energy radiated per second increases and when the listener comes nearer to the sound source. Just how will such a variation be interpreted by the ear? When a tuning fork is struck lightly, a faint sound is heard. When it is struck vigorously, more energy is radiated per second. The sound is louder. When a vibrating fork is held at arm's length and then is brought nearer the ear, the sound intensity at the ear increases because of the shortening distance. The loudness is found to increase also. *Thus, we have definite evidence that an ordinal relationship exists between the intensity of sound and the loudness perceived.*

Limits of audition. Although we have shown a relationship between pitch and frequency of vibration, and between sound intensity and loudness, we may not conclude that sounds of all frequencies or all intensities may be heard. If a pure tone is kept at a *constant intensity* and at the same time is raised or lowered in frequency, it ceases to be heard as the pitch becomes either too low or too high. These

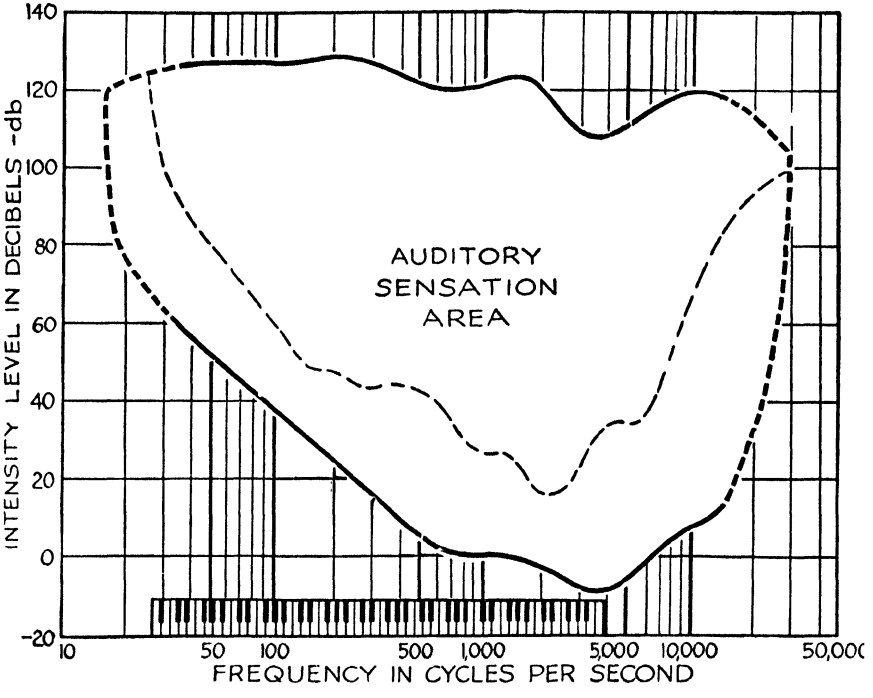


Fig. 111. Illustrating the Auditory Sensation Area. (*After Fletcher. Courtesy of Bell Telephone Laboratories.*)

pitch limits change with a shift in sound intensity. Again, if a pure tone is kept at a *constant frequency* and at the same time is raised or lowered in intensity, lower and upper limits are reached. These intensity limits are not constant with pitch, but have the form given by Fletcher (2) and illustrated in Fig. 111. The lower line represents the sound intensity just sufficient to be heard and is called the *threshold of audibility.* The upper line corresponds to the intensity at which hearing changes to feeling and pain

and is called the *threshold of feeling.* A sound wave with a frequency and an intensity within the area bounded by these curves—the so-called *auditory sensation area*—may be sensed as a sound. The sensation area depicted is an average of many normal ears. Threshold curves for a single individual are never so smooth as that shown. Dr. Harvey Fletcher and his associates at the Bell Telephone Laboratories have accurately determined the boundaries of the auditory sensation area, one of many contributions by them to the knowledge of hearing and the general field of sound.

Consulting the figure, we see that the maximum range of frequency is from approximately 20 to 20,000 cycles per second. But, as is clearly shown, this range shortens if the sound intensity is very low. At a frequency of 50 cycles per second, the sound intensity necessary to produce a just audible sound is 100,000 times greater than that needed to produce such a sound at 1,000 cycles per second. At 1,000 cycles per second, the intensity at the *threshold of feeling* is 1,000,000,000,000 times the intensity at the *threshold of audibility.*

If loudness were to vary directly with the sound intensity, this would mean that between the threshold of audibility and the threshold of feeling the loudness would increase a trillionfold. This certainly does not agree with common experience. Thus, obviously no *simple* relationship exists between intensity and loudness. The customary musical notations *ff*, *f*, *mf*, *p*, and *pp*, are used to designate loudness, but these terms are not precise. They depend upon the acuteness of hearing and, in general, upon the customs and experiences of the persons using them. In the interest of scientific measurements, a more precise loudness scale has been constructed. The student interested in the more technical aspects of the problem will find in the next topic material which the acoustical engineer must master, but which the beginning student may not care to investigate.

Intensity level and loudness level. According to the Weber-Fechner Law, the intensity of a sense datum is

generally proportional to the logarithm of the intensity of the external stimulus. Although this law is known to have many exceptions, it does suggest a new unit of sound intensity measurement based on logarithms. For the benefit of those not acquainted with the logarithms of numbers, the following table is submitted:

Number N	*Logarithm*$_{10}$ *N*
0.000000001	−9
.....	...
0.0001	−4
0.001	−3
0.01	−2
0.1	−1
1	0
10	1
100	2
1,000	3
10,000	4
.....	...
1,000,000,000	9

We shall define *sound intensity level* such that if the numbers in the left-hand column represent *sound intensity*, the numbers in the right-hand column represent *sound intensity level*. The intensity level is measured in terms of the unit called the *bel*, named in honor of Alexander Graham Bell (1847–1922), the inventor of the telephone. Usually, however, a unit one-tenth the size, the *decibel* (db), is used as the practical unit of sound intensity level. By consulting the table, you will observe that the logarithm of *unity* is zero, and thus the zero on the intensity level scale will depend on what intensity is used as the reference unit standard. The zero intensity level now accepted by engineers corresponds to a flow of sound energy equal to 10^{-9} or 0.000000001 ergs per square centimeter per second, a sound intensity so small as to be only slightly above the threshold of audibility in the region where the ear is most sensitive (Fig. 111).

Engineers have found it advantageous also to use a so-called *loudness level scale*. First a pure tone having a frequency of 1,000 cycles per second is selected as the

standard tone, and its *intensity levels* are arbitrarily taken as its *loudness levels.* This being done, the loudness level of another tone is said to be the same as the standard when the two tones are judged by the ear to have the same loudness. Thus, two tones with the same *loudness level* sound equally loud.

As may be seen from Fig. 111, the acuteness of the ear decreases with the pitch. Thus, for example, at an intensity level of 50 decibels, a pure tone with a frequency

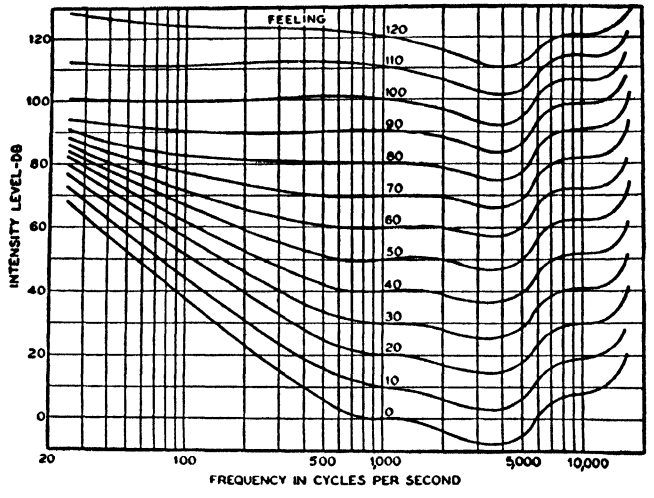


Fig. 112a. Equal Loudness Level Curves. (*Fletcher and Munson. Courtesy of Bell Telephone Laboratories.*)

of 64 cycles per second is just being heard, while the standard tone (1,000 cycles per second) has a loudness level of 50 decibels. However, if the low tone is increased in intensity level from 50 decibels to 70 decibels, its loudness level is found to change from 0 to 40 decibels (Fig. 112a). Thus, with a rise in *intensity* level, this lower-pitched tone increases in *loudness* level at a faster rate than does the standard tone. This is illustrated by the fact that the lines representing equal loudness levels are closer together in this region. Finally, at the *intensity* level of 90 decibels, the *loudness* level of the low tone is approximately the same as that for the standard tone at the same *intensity* level.

To determine the loudness level of a pure tone, therefore, one must know, not only the *intensity level* of the sound, but also the *frequency of the vibration.* Through a large part of the practical frequency range (300–4,000 cycles per second), the loudness level is approximately equal to the intensity level.

Thus, we see that the *intensity level scale* is entirely independent of the sense of hearing, and that the *loudness level scale,* insofar as the standard tone is concerned, is also free from this sense, because the loudness level of the standard is taken arbitrarily as equal to its intensity level. However, the extension of the loudness level scale to include all tones has required the use of the ear in rendering equal-loudness judgments; therefore this scale is very dependent upon the so-called normal ear. Even so, loudness level may not be taken as the *direct* measurement of the subjective sound characteristic, loudness.

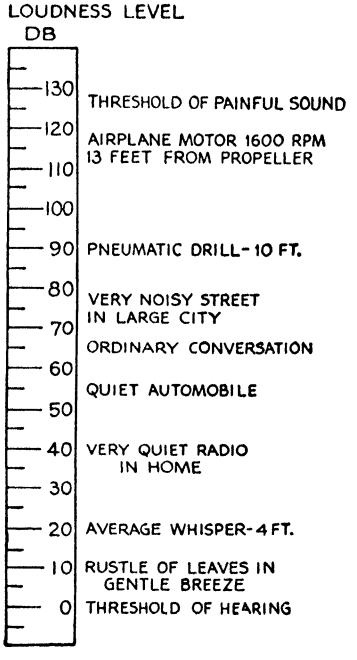


Fig. 112b. Loudness Level of Common Sounds.

To obtain a concrete notion of what all this means in terms of everyday experience, one may set a list of familiar sounds opposite a loudness level scale and see just how the sounds of his experience are interpreted in terms of loudness level (Fig. 112b).

A loudness scale. We must still seek a *true* loudness scale—one based throughout on loudness judgments. The units of such a scale would be doubled with a doubling of the magnitude of the sensation and tripled with a tripling of the loudness sensation, and so on. On this scale one could find the answer to the question: Just what intensity level change is required in order that a sound may be judged

twice as loud, four times as loud, or ten times as loud? Recently a scale of this type was created. It is based on data arising from various types of experiments of which the following is a simplified example: Let a listener, using both ears, fix his attention on the loudness of a sound such as, for example, a pure tone making 1,000 cycles per second; and let the sound intensity level at his ears be measured. Now ask the listener to close one ear. The loudness of the sound will decrease; in fact, the magnitude of the sensa-

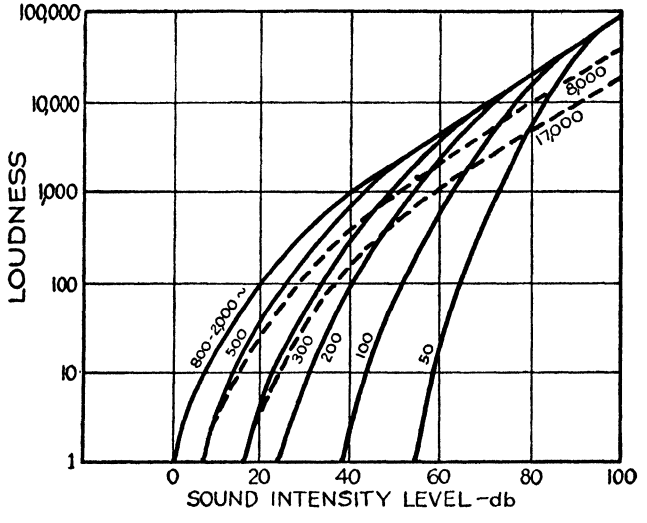


Fig. 113. A Loudness Scale for Pure Tones. (*After Fletcher. Courtesy of Bell Telephone Laboratories.*)

tion will be halved. Next, let the sound intensity level at the ear be increased till the loudness perceived with the one ear is the same as that originally heard with both ears. In this manner the change in intensity level for a doubling of the magnitude of the sensation is obtained. This, of course, assumes that both ears are identical in their response to sound waves. Fletcher summarizes the results[2] in this field by curves illustrated in Fig. 113. Clearly, the subjective characteristic, loudness, is controlled by all three of the

[2] Of studies by Fletcher and Munson, Ham and Parkinson, Firestone and Geiger, and Churcher, King, and Davis.

physical characteristics—sound intensity, frequency of vibration, and overtone structure—depending principally upon the *sound intensity* and secondarily upon the frequency and overtone structure.

Further, we find that for a pure tone of 1,000 cycles per second and for a loudness between 1,000 and 10,000 units, the doubling, the quadrupling, and the increasing ten times of the apparent loudness require an intensity level increase of 10 decibels, 20 decibels, and 33 decibels, respectively. In fact, a 10-decibel change of sound intensity level corresponds approximately to a doubling in loudness of sounds with a frequency range from 800 to 2,000 cycles per second and an intensity range from 40 to 100 decibels.

Limits of Hearing

How many pure tones may be heard? If the ear could detect the slightest change in frequency of vibration or the minutest variation in sound intensity, there would be no limit to the number of tones, each having a difference in pitch or loudness or both, which the ear could detect. But the ear does not have such sensitivity. At frequencies between 500 and 4,000 cycles per second, the normal ear can detect *changes* in *pitch* amounting to 0.3 per cent. This means that the pitch of pure tones characterized by 500 and 501.5 cycles per second or by 1,000 and 1,003 cycles per second may just be distinguished in pitch. As frequencies go below 500 and above 4,000 cycles per second, the ability of the ear to distinguish changes in pitch decreases, until finally at 64 cycles per second a variation of 0.9 per cent is required in order for the ear to be able to detect a frequency change.

The ability of the ear to distinguish *changes in sound intensity* varies greatly with the pitch and the loudness of the tone. Within the region ordinarily used in conversation, the fractional change in sound intensity which is just perceptible lies between approximately 5 and 20 per cent (0.2 and 0.8 decibels).

Making use of such facts, we may divide up the auditory sensation area into tiny regions, each representing a single pure tone which the normal ear may perceive. It is estimated that approximately 540,000 such regions exist. Therefore, this number of pure tones, differing in pitch or loudness or both, may be heard by the normal ear. Undoubtedly a very much larger number of complex tones may be heard, for reasons to be explained later.

Impaired hearing. Persons with impaired hearing have a smaller auditory sensation area than those with normal hearing, the decrease being due to *the rise of the threshold of audibility*. This higher threshold is illustrated by the dotted line in Fig. 111. In the so-called *nerve deafness*, the greatest loss of *acuteness* is for tones of high frequency, and often such persons may have normal hearing for low frequencies. In such cases the high-frequency components of speech are not heard, and great difficulty is experienced because many of the consonant sounds may not be distinguished or even heard. In the so-called *conductive deafness*, the acuteness of hearing for low frequencies is lost, the ear often being normal for high frequencies. Such deafness is not such a handicap as nerve deafness because low-frequency tones are not so necessary in the interpretation of speech as are those of high frequency. Impaired hearing is somewhat prevalent among persons advanced in years, and this deafness attendant with age may represent a loss of hearing through the whole frequency range, the loss usually appearing first at the very high frequencies.

The dotted line in Fig. 111 is the curve of the threshold of audibility of a person rated by Fletcher as having a hearing loss of 26 per cent. Such a person might understand conversational speech in a small room, but undoubtedly would experience much difficulty in understanding a speaker in an auditorium unless he were to choose a position in the first few rows of seats. This person would be at an advantage in a noisy room, because for others with normal hearing to carry on an intelligent conversation in such a

room, the sound intensity of the speech would need to be increased, and the intensity of speech would be brought well within the auditory sensation area of the deafened person.

Audiometers, which measure hearing in terms of the so-called normal ear, are used in clinics, hospitals, and schools. In schools the audiometer often consists of a phonograph turntable and a set of forty ear phones, one



Fig. 114. Testing the Hearing of School Children. (*From Fletcher's "Speech and Hearing." Courtesy of D. Van Nostrand Company, Inc.*)

for each child to be tested (Fig. 114). As the recording on a record is reproduced in the telephone receivers, the children recognize a voice telling them to write down what they hear. But because of the nature of the record, the loudness level of the voice decreases step by step. Soon a certain child begins to guess, then he gives up writing altogether. His record reveals his hearing ability, and the records of the other children reveal their hearing. Approximately ten per cent of the children in schools have defective hearing.

Deafening due to masking. In a noisy factory or subway or at a busy street corner, one finds himself talking at the

top of his voice in order to be heard. This masking effect of noise is the equivalent of a temporary deafening. We hardly realize that this type of deafening is nearly always present until we find ourselves in a very quiet place, and then we hear the surge of blood through the arteries of the head, the feeble noises of insects, the creaking of timbers, and many unfamiliar sounds. Musical tones as well as noises may produce a masking effect. Low-pitched tones with considerable intensity produce a rather marked masking effect upon high-pitched tones, but high-pitched tones produce only a feeble masking effect upon low-pitched tones. In general, *the greatest masking takes place near or above the frequency of the masking tone.* Thus, the low tones of revolving machinery produce a deafening which affects the whole of the frequency range required for an adequate interpretation of speech. Instruments based on the deafening due to masking are now used to determine the noise found in city streets, public buildings, offices, and factories (Fig. 112b).

The Quality of a Musical Tone

With very little practice, one may acquire the ability to distinguish the tones produced by a series of musical instruments even though they are all produced at the same pitch and with equal loudness. This distinguishing characteristic is called *tone quality.* The "velvet"-noted flute, the "melting" clarinet, the "humming" violin, the "voice with a smile" are all expressions in which an attempt is made to select a word adequate to describe the tone quality of the sound. What physical characteristics are interpreted as tone quality?

Let us study the tone produced by a steel wire stretched tightly over a long, light, wooden box. The box serves as a soundboard, its surfaces being set into forced oscillations by the vibrating wire (Fig. 115). In this way more energy is radiated per second and thus the loudness of the sound

produced is increased. If the string of this so-called *sonometer* is plucked near one end, a tone having the characteristics of pitch, loudness, and tone quality is heard. Let us call this note the fundamental and designate it *do*. When the string is plucked near one end and then touched very lightly at its middle, the fundamental tone ceases, but the octave *do'* is distinctly heard. When the string is again plucked and then very lightly touched at a point one-third of its length from an end, the fundamental tone again ceases and a new tone *sol'* appears. When the string is plucked again and then touched one-fourth of its length from an end, the new tone is the double octave, *do''*, of the fundamental. Still higher tones may be evoked if the process is continued, the string being touched lightly at a point located one-fifth, one-sixth, one-seventh, or one-eighth its length from an end after being plucked at the middle of the segment.

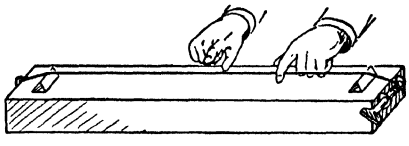


Fig. 115. Demonstrating the Origin of Tone Quality.

Recall the experience of producing standing wave patterns on a vibrating rope (Fig. 96). You will remember that it was possible to cause the rope to vibrate as a whole, in halves, thirds, fourths, or any whole number of equal parts. The simplest form we called the *fundamental;* the others we called *overtones.* The frequency of vibration for each vibration pattern increased in the same proportion that the number of segments increased. With this as a background, it is easy to understand the experiment we have just performed. The steel wire when plucked near one end vibrates as a whole, producing the fundamental *do;* in halves, producing the first overtone *do';* in thirds, producing the second overtone *sol';* in fourths, producing the double octave *do''*. The pitch of the full tone corresponds to the frequency of vibration of the fundamental, but the tone quality results from a blending of tones, a harmonic series of tones made up of the fundamental and overtones.

Each overtone is heard at its own pitch when isolated by the method outlined in the experiment.

From such results as these, after an elaborate investigation, Hermann Von Helmholtz (1821–1894) concluded that *the quality of a tone depends upon the number and relative intensity of the pure tones which combine to produce the complex tone.* The fact that the ear is able to distinguish between tones of different quality is evidence that the hearing mechanism is able to analyze a musical tone into its component pure tones. The theory of hearing which explains how this is done will be presented later.

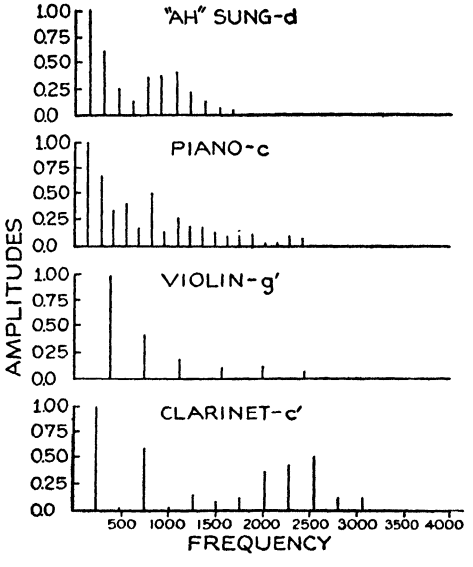


Fig. 116. Sound Spectra of Tones. (*After Fletcher. Courtesy of Bell Telephone Laboratories.*)

By means of various electrical and mechanical devices, the tones produced by the human voice and various musical instruments have been analyzed and the intensity of the overtones present have been determined. A few such sound spectra are shown in Fig. 116.

Although playing the principal role, the sound spectrum or overtone structure of a tone does not uniquely determine its quality. A general shift in the frequency of all of the components or a general increase in the sound intensity of these elements without a change in the overtone structure modifies the tone quality. For example, a violin tone from a high-quality loud-speaker may lose its violin quality if reproduced at a high-intensity level, even though in the process the overtone structure is faithfully maintained.

Finally, we conclude that the *subjective* characteristic, *pitch,* depends principally upon the physical characteristic,

frequency of vibration; loudness, principally upon *sound intensity;* and *tone quality*, principally upon *overtone structure*. Yet *each* of these three subjective characteristics depends to a greater or less degree upon *all three* of the physical characteristics.

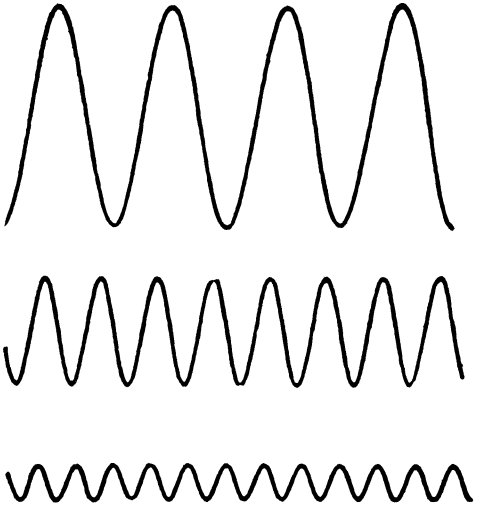


Fig. 117. Wave Forms of Pure Tones.

Wave Form

Sound waves in air are longitudinal in form and thus are rather difficult to depict. But if their true nature is kept clearly in mind, it will be instructive to illustrate them by using a transverse wave pattern. The simplest sound wave is that originated by a body vibrating with simple harmonic motion. The tone produced is a *pure* tone, and the sound wave takes the form shown in Fig. 117, in which three waves differing in frequency (pitch) and intensity (loudness) are depicted. Wave forms such as these are produced by properly mounted tuning forks (Fig. 119).

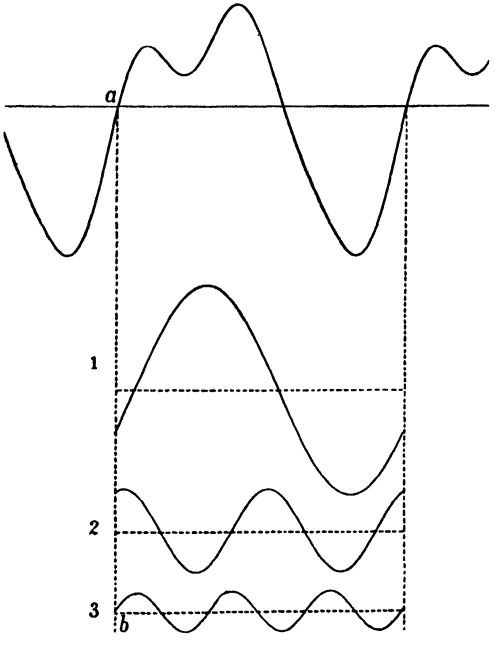


Fig. 118. Compounding to Form a Complex Wave. (*Miller.*)

In the case of a vibrating string, for example, the sound wave is initiated by a body which vibrates simultaneously in various

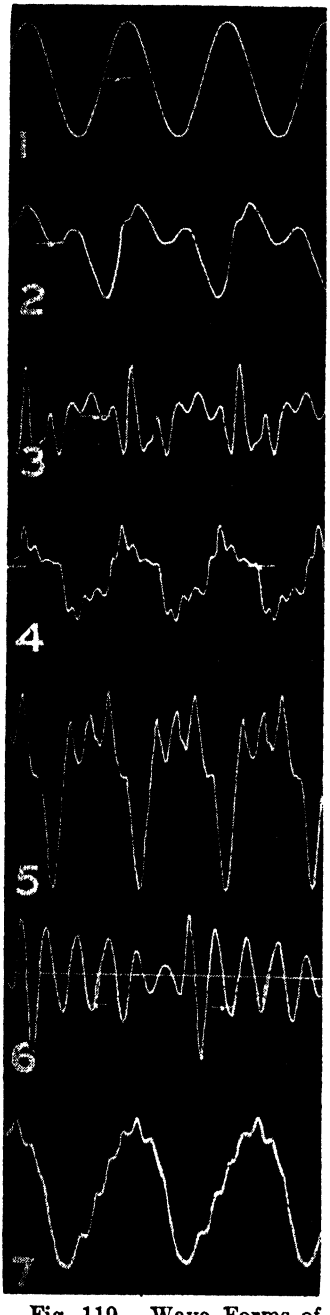


Fig. 119. Wave Forms of (1) Tuning Fork, (2) Bass Flute, (3) Oboe, (4) Clarinet, (5) Saxophone, (6) French Horn, and (7) Violin. (*Miller.*)

modes, in *a series of harmonic vibrations.* The air wave must carry the net result of all these components, and the result is a complex wave form. In Fig. 118 such a wave, resulting from compounding three simple vibrations, is shown. Clearly, a tone composed of a long series of harmonics may be expected to have a very complex wave form, a form depending upon the intensities of the components and the phase relations between them. Since changing the phase of the components produces large changes in the wave form but very little modification of the tone quality, it is incorrect to judge the quality of a tone wholly by the form of its wave. It is best to use the overtone structure (sound spectrum), as illustrated in Fig. 116. Yet wave forms (see Fig. 119) do illustrate in a vivid manner why each musical instrument has its characteristic tone quality. Dr. Dayton C. Miller has made significant contributions in the field of wave form analysis. With his permission, we reproduce a number of excellent curves. The reader is urged to consult his attractive book, *Science of Musical Sounds.*

The Human Ear

Structure. The human ear performs two distinct functions. Through the aid of the semicircular

canals of the ears, we are able to maintain equilibrium. Through the operation of the hearing mechanism, we obtain auditory sense data. The ear is divided into three chambers, as is shown in Fig. 120. The *outer ear* consists of the ear canal, closed at its inner end by a diaphragm, the ear drum. The *middle ear* contains an arrangement of three small bones known as the *hammer*, *anvil*, and *stirrup*. These so-called "ossicles" form a lever system which connects the ear drum

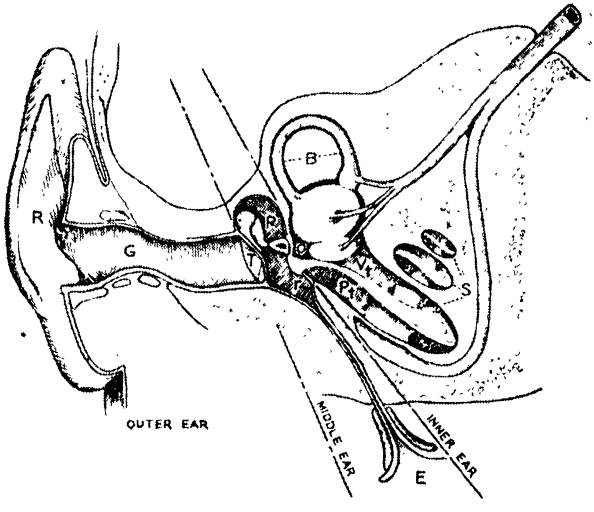


Fig. 120. The Ear. (*From Fletcher's "Speech and Hearing." Courtesy of D. Van Nostrand Company, Inc.*)

with the oval window, a membrane-covered opening which, along with another similar opening, the round window, closes in the inner ear. The lever system changes large amplitudes of motion at the ear drum into shorter but more forceful motions at the oval window. Owing to this lever action and the fact that the ear drum is much larger than the oval window, the force is increased 60 times. The *inner ear* contains the cochlea, with its bony walls shaped like a snail shell, and its coiled spiral enclosure separated into two portions, except at the very tip, by the basilar membrane. The endings of the auditory nerves are scattered along the basilar membrane, which is $\frac{1}{100}$ of an inch wide and, if uncoiled, 1.2 inches long. Thus, the

important organ of hearing is very small indeed. In Fig. 121 the three chambers of the hearing apparatus are shown diagrammatically, the whorls of the cochlea being unwound to simplify the picture.

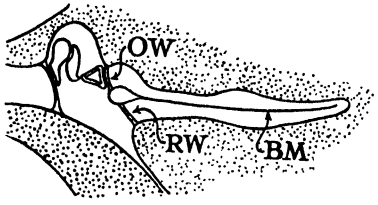


Fig. 121. Diagrammatic View of the Ear (Semicircular Canals Omitted).

Operation. When a sound wave strikes the ear drum, the vibrations are transmitted through the middle ear to the oval window *OW* by the bony lever system—the hammer, anvil, and stirrup. At this window, the vibrations are communicated to the fluid contained in the cochlea. Since the walls of the inner ear are rigid and the fluid is practically incompressible, a movement at the oval window also gives rise to a movement at the round window *RW*.

According to the theory of hearing first proposed by Helmholtz and later developed and extended by Fletcher, a mechanical vibration is transmitted a certain distance through the liquid, depending on the frequency of its vibration, then through the basilar membrane, and on to the oval window (Fig. 121). At very low frequencies, the surfaces on either side of the membrane receive the same pressure. The membrane remains stationary, the liquid virtually moving bodily back and forth through the small opening at the apex of the cochlea, thus causing the round window *RW* to oscillate. When this situation exists, no nerves in the membrane are stimulated and no sound is heard. As the frequency is decreased to approximately 20 cycles per second, this condition, which marks the lower limit of pitch audibility, is reached. As the frequency of vibration increases, the end of the membrane farthest from the oval window is first stimulated, and the little patches marking the regions of maximum membrane vibration move toward the oval window. But finally, at about 20,000 cycles per second, the inertia of the ossicles and liquid is so great that very little if any

energy reaches the cochlea, and the upper limit of pitch audibility is reached. Fig. 122a is a much simplified diagrammatic representation of the inner ear. The location along the basilar membrane of the patch of maximum response, corresponding to a certain pitch, may be approximately located by reference to the figure.

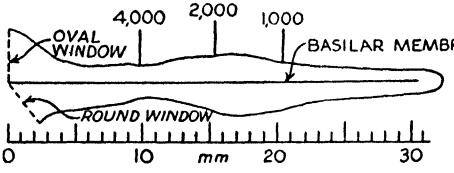


Fig. 122a. Pitch Variation Along Basilar Membrane. (*After Steinberg.*)

Loudness. It is thought that the loudness of a sound is directly related to the total nerve energy being sent to the brain from the hearing mechanism, and to the number of nerve fibers active in the process. Thus, the nature and the sensitivity of the nerve endings and the area and amplitude of vibration of the agitated patch are the aspects of the hearing mechanism which give rise to the sense impression of loudness.

Pitch. The pitch of a tone is thought to be determined by the *position* along the basilar membrane of the maximum

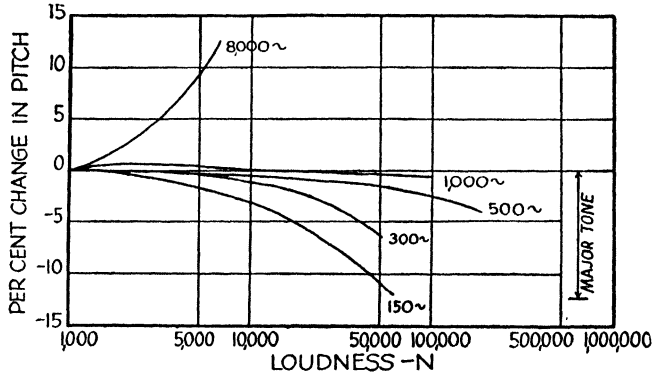


Fig. 122b. Pitch Change Due to Loudness. (*After S. S. Stevens.*)

stimulation, and also by the *frequency of the impulses* sent along the nerve fibers. The position of the vibrating patch seems especially important for high-frequency tones, but the frequency of the impulses is probably more important for tones of low frequency. The position of the vibrating

patch seems to shift slightly when great amplitudes are reached. At least, this would explain the change in pitch with loudness (Fig. 122b).

The pitch of a complex tone with components in a harmonic series corresponds, except for slight modifications due to great loudness, to that of the fundamental of the series. And this is true even though this fundamental component is actually missing in the physical wave that reaches the ear. The explanation is as follows: At low sound intensities, a simple air wave—a pure tone (Fig. 117)—which appears at the ear drum is transmitted into the cochlea without change of form. But at high intensities, the elastic members of the middle ear do not obey Hooke's law (equal displacements produced by equal forces) for such large amplitudes. Therefore, on the way into the inner ear, the simple vibration is modified so that, if the frequency of the pure tone is designated by S, the new tones, called *subjective* tones, with frequencies $2S$, $3S$, $4S$, and so forth, arise. Thus, at a high sound intensity, a *pure* tone on reaching the inner ear is composed of the *original* fundamental vibration and, in addition, a series of *subjective* harmonic overtones. A loud pure tone, therefore, produces a *complex* vibration pattern on the basilar membrane. This vibration pattern is a sort of joining together of the isolated patches corresponding to the fundamental and the subjective overtones.

The pitch of a complex tone having four components of equal intensity but with frequencies of 400, 600, 800, and 1,000 cycles per second is perceived as a tone with a pitch of 200 cycles per second—the pitch of the *fundamental* of the harmonic series, which was not even present in the original complex tone. If three additional components, 500, 700, and 900 cycles per second, are added to the complex tone, the pitch is sensed as an octave lower, or 100 cycles per second. This is the fundamental of the new harmonic series. These facts may be interpreted in terms of subjective tones as follows:

When two pure tones of high intensities fall upon the ear, the elastic members of the middle ear are given amplitudes so large that Hooke's law is no longer obeyed. Then, not only will subjective harmonic overtones be produced for each tone, but so-called subjective *difference* and *summation* tones will be produced. Thus, if S_1 and S_2 represent the frequencies of the two pure tones, then the subjective harmonic overtones, $2S_1$, $3S_1$, $4S_1$, and so forth, and $2S_2$, $3S_2$, $4S_2$, and so forth; the subjective difference tones, $S_1 - S_2$, $2S_1 - S_2$, $2S_1 - 2S_2$, and so forth; and the subjective summation tones, $S_1 + S_2$, $2S_1 + S_2$, $2S_1 + 2S_2$, and so forth, will arise as the energy passes to the inner ear. Thus, in the cases just described, the needed fundamentals are manufactured as subjective difference tones, and the effect on the basilar membrane is the same as if a complete harmonic series were present in the original tone. For this reason, the pitch in each case was perceived as that of the fundamental of the harmonic series.

Musical tones are made up of components in a harmonic series. Hence, even at low intensities, such a tone produces a complex vibration pattern on the basilar membrane. What would happen to the pitch of a tone if the fundamental and the first few overtones were suppressed? (You may hear the effect if you listen to the phonograph record *BTL 4-B* produced by Bell Telephone Laboratories.) The elimination of the *fundamental* and the *first four overtones* of a tone from a piano does not change its pitch, although the quality is greatly impaired. Probably at high sound intensities, the ear manufactures, as subjective tones, the missing components of the harmonic series; and thus the fundamental, though suppressed in the original tone, exists as a subjective tone in the final stimulus producing the sound sense datum. At very low sound intensities, this would not be true; but since the pitch of the original sound persists, it appears that the pitch of a tone depends in some measure upon where the individual patches of the composite pattern are located along the basilar membrane.

Tone quality. Musical tones which reach the ear are in most cases already complex in form, being composed of a group of component vibrations in a harmonic series. A low-intensity tone of this kind gives rise to a pattern of vibration on the basilar membrane produced by the joining together of the isolated patches corresponding to the components of the series. Thus, each tone has a space pattern of its own, and the ear interprets these patterns as tone quality. Owing to the appearance of subjective components at high intensities, the quality of a tone changes with an increase in intensity. For this reason, it is best to reproduce music or speech from a radio at the same sound intensity as existed in the original rendition.

A striking example of a resonant effect, which roughly illustrates the action of the basilar membrane in the analysis of a complex tone, may be produced by singing into the undamped strings of a piano. A tone very similar to the voice will be returned. The component frequencies of the voice found strings of the same frequency in the piano, resonance was established, and a complex tone, the net result of all the strings set into sympathetic vibration, was evoked. A different tone uses a different set of strings, or at least causes certain ones to vibrate with different amplitudes. Thus, each tone has its unique pattern of strings and amplitudes. Supply the nerves and brain, and the analogy is complete.

Volume and brightness. Low-pitched or complex tones have a large "volume" and high-pitched tones have small "volume." This characteristic is not identified as pitch, loudness, or tone quality, but is sensed probably when a rather long patch of basilar membrane is stimulated. Another psychological experience called "brightness," characteristic of high-pitched tones as contrasted to the "dullness" of low-pitched tones, is probably sensed when the amplitude of vibration of the basilar membrane falls off rapidly as the edge of an oscillating patch is approached. These characteristics, volume and brightness, are sub-

jective; they have no counterpart in sound waves, as have loudness, pitch, and tone quality.

Beats, Consonance, and Dissonance

So far we have studied the nature of single tones; let us now investigate the result of combining such tones. If two tuning forks having a very slight difference in frequency—any two common forks marked as having the same frequency will usually serve well—are sounded together, a throbbing or beating effect will be heard. The number of beats per second may be controlled by the addition of wax to the prongs of one of the forks. In this manner one fork may be tuned to the other—a condition that will be achieved when no beats are heard. This is perfect unison, the finest example of consonance. In a similar manner, musical instruments being tuned to the same pitch may first be brought to a point where beats are heard and then gradually to a point where the beats are eliminated.

But why are beats heard when the forks fail to vibrate in unison? Suppose that one fork makes exactly two more cycles per second than the other. This means that if they are just in phase at the beginning of a second, they will be just in phase at the beginning of the next second. But since the one must make two more complete vibrations during this time than the other, it must slowly move ahead in phase, first arriving at complete opposition, then at complete agreement, then at opposition, and finally again at agreement to begin the next second. Each fork originates a train of waves, and these trains of waves travel out at the same speed. The phase relation between the waves will change at exactly the rate found for the forks. Hence, twice a second the emitted waves combine in exact phase and twice in opposite phase. As these combined waves pass the ear, twice during a second a state of re-enforcement passes by, and in between, a state of destructive interference (Fig. 123). Thus, twice during a second, the sound will

wax and wane, and two beats per second will be heard. The number of beats heard per second turns out to be exactly equal to the difference in frequency.

If the difference in pitch of two tones is increased, the number of beats per second increases, and finally the throbbing becomes *so rapid as to give a roughness known as* ***dissonance.*** With further difference in pitch, a new tone having a frequency equal to the difference in the two frequencies may be heard, and finally this tone blends with the tones which produce it, and *consonance* is again achieved.

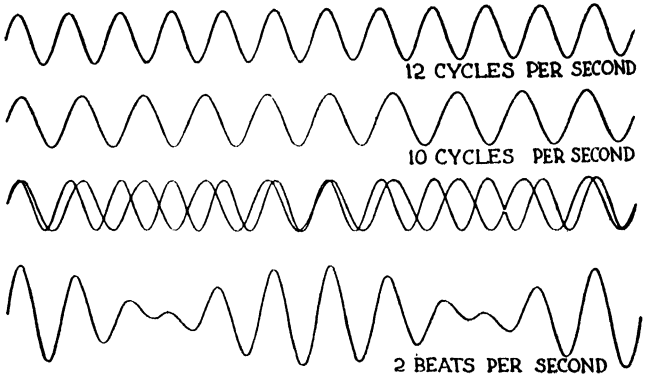


Fig. 123. Illustrating the Cause of Beats.

There is always consonance among the components of a complex tone produced on a string or in an air column. This is true, as we shall see, because the tone is made up of a fundamental and overtones in a harmonic series. For example, in the harmonic series having 100, 200, 300, 400, 500, and 600 cycles per second, the beat tones, obtained by taking the differences of the frequencies of the various components, are found to be 100, 200, 300, 400, and 500 cycles per second—frequencies already found in the original complex tone and therefore without possibilities of adding roughness to the tone. On the other hand, if one complex tone is mistuned to another, not only will the fundamentals add roughness due to beating, but the overtones will do likewise.

Musical Intervals

In the study of pitch it was pointed out that octave changes are attained, not by the process of addition or subtraction, but by a process of multiplying and dividing by 2. For example, 64 is the octave of 32; 4,096 is the octave of 2,048. The octave is called a *musical interval with ratio 2:1.* In general, the ratio of the frequencies of two notes is called the *interval* between them. Not all intervals are found to be *consonant;* as a matter of fact, only intervals which are the ratio of *small whole numbers* are found to give consonance. Even the ratio 6:5 was not considered to be a consonant interval during the early history of music, and such intervals as 7:6, 8:7, 9:8, 16:15, and 25:24 are still not pleasant intervals to the ear unless handled in a masterly manner.

The most common musical intervals are as follows:

Unison	1:1
Octave	2:1
Fifth (fifth place in major or minor scale)	3:2
Fourth (fourth place in major or minor scale)	4:3
Major third (third place in major scale only)	5:4
Minor third (third place in minor scale only)	6:5
Major sixth (sixth place in major scale only)	5:3
Minor sixth (sixth place in minor scale only)	8:5
Major second (second place in major scale only)	9:8
Semitone	25:24

The Musical Scale

Major diatonic scale. If three notes with frequency ratios 4:5:6, that is, three notes with *major third* and *minor third* intervals, are sounded together, the effect is very pleasing. Three triads of this sort are the basis for the major diatonic scale. Call the notes of this scale

Do Re Mi Fa Sol La Ti Do'

If *Do* has a frequency proportional to 24, the other members of the scale are built up out of the three triads as follows:

	Do	*Re*	*Mi*	*Fa*	*Sol*	*La*	*Ti*	*Do'*	*Re'*
Tonic	24		30		36				
	(4)		(5)		(6)				
Dominant		27			36		45		54
		(3)			(4)		(5)		(6)
Subdominant				32		40		48	
				(4)		(5)		(6)	

The following table shows the interval between each note and *Do* and also the interval between adjacent notes.

	Do	*Re*	*Mi*	*Fa*	*Sol*	*La*	*Ti*	*Do'*
Frequency proportional to	24	27	30	32	36	40	45	48
Interval between each note and *Do*	$\frac{1}{1}$	$\frac{9}{8}$	$\frac{5}{4}$	$\frac{4}{3}$		$\frac{5}{3}$	$\frac{15}{8}$	$\frac{2}{1}$
Interval between adjacent notes		$\frac{9}{8}$	$\frac{10}{9}$	$\frac{16}{15}$	$\frac{9}{8}$	$\frac{10}{9}$	$\frac{9}{8}$	$\frac{16}{15}$

Notice that the intervals between adjacent notes are of three kinds: A major tone, $\frac{9}{8}$ or 1.125; a minor tone, $\frac{10}{9}$ or 1.111; a half tone, $\frac{16}{15}$ or 1.067. On the basis that intervals expressible as the ratio of small whole numbers give best consonance, we may list *Do* with each note of the scale in the order of decreasing consonance as follows:

Do-Do' *Do-Sol* *Do-Fa* *Do-La* *Do-Mi* *Do-Re* *Do-Ti*

This is in agreement with the experimental work of Helmholtz presented graphically in Fig. 124. The magnitudes of the dissonance between *Do* and the other notes of the scale are represented as vertical distances. As the pitch moves up from *Do*, the dissonance increases rapidly, then decreases, and finally reaches another high maximum near *Do'*. In between, distinct drops in dissonance are seen at *Mi*, *Fa*, *Sol*, and *La*, bearing out the conclusion that intervals which are the ratio of simple whole numbers give consonance. Also, mistuned notes may give greatest dissonance.

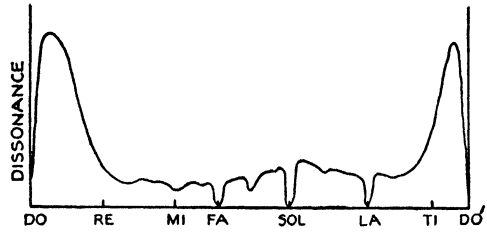


Fig. 124. Dissonance Between *Do* and the Other Notes of the Scale.

Minor diatonic scale. If notes in the ratio 10:12:15, that is, three notes with intervals of a *minor third* and a *major third,* are sounded together, the effect is not only pleasing but is of unique quality. If we begin with *Do* and give it a frequency proportional to 24, we may work out the following table in which the notes have the same names, but of course not the same frequencies, as those of the major scale.

	Do	*Re*	*Mi*	*Fa*	*Sol*	*La*	*Ti*	*Do'*
Frequency proportional to	24	27	28.8	32	36	38.4	43.2	48
Interval between each note and *Do*	$\frac{1}{1}$	$\frac{9}{8}$	$\frac{6}{5}$	$\frac{4}{3}$	$\frac{3}{2}$	$\frac{8}{5}$	$\frac{9}{5}$	$\frac{2}{1}$
Interval between adjacent notes	$\frac{9}{8}$	$\frac{16}{15}$	$\frac{10}{9}$	$\frac{9}{8}$	$\frac{16}{15}$	$\frac{9}{8}$	$\frac{10}{9}$	

You will observe that the intervals between adjacent notes are of the three kinds found in the major scale, but that they are located at different positions. For example, in the major scale the half tones are between the third and fourth positions and the seventh and eighth positions; on the minor scale the half tones are located between the second and third positions and the fifth and sixth positions.

The scale of equal temperament. Although the major and minor diatonic scales use notes which form exact and simple ratios, and therefore may be called *consonant* scales, they are not composed of equal interval steps. If one were to play the major diatonic scale in every musical key, he would need at least four notes for each one now in the scale. On a keyboard instrument, such as a piano or organ, the complexity required would make playing very difficult indeed. For this reason, a scale of equal temperament having twelve equal intervals per octave has been constructed. The interval needed is a number which when multiplied by itself twelve times gives 2. Such a number is the twelfth root of 2, or 1.05946. The frequency of vibration of each note beginning with middle *C* and making an octave on the International Tempered Scale (the stand-

ard pitch being 435 cycles per second for *A* above middle *C*) is as follows:

C	*C#*	*D*	*D#*	*E*	*F*	*F#*	*G*	*G#*	*A*	*A#*	*B*	*C*
258.7		290.3		325.8	345.3		387.5		435.0		488.3	517.3
	274.0		307.6			365.8		410.6		460.9		

Many of the notes of this scale are displaced slightly from the positions of best harmony, but the simplicity which the scale brings to fixed-keyboard instruments justifies its creation.

Questions and Problems

1. What is the double octave of a note with a frequency of 256 cycles per second?

2. List the primary and secondary physical characteristics corresponding to each of the subjective characteristics, pitch, loudness, and tone quality.

3. By means of numbers illustrate a harmonic series.

4. Contrast deafening due to impaired hearing and deafening due to masking.

5. List the important parts of the ear and describe the function of each.

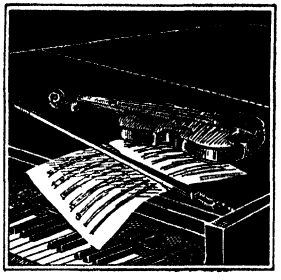
6. Quality of a tone depends upon the ..... and .... of the pure tones which combine to produce the complex tone.

7. Dissonance is the result of a roughness caused by .....

8. The ratio of the ..... of two notes is called the ..... between them.

Suggested Readings

(1) Fletcher, Harvey, "Newer Concepts of the Pitch, Loudness, and Timbre of Musical Tones," *Journal of the Franklin Institute*, Vol. 220, 1935, pp. 405–429.
(2) ———, *Speech and Hearing*, D. Van Nostrand Company, Inc., New York, 1929, Parts II and III.
(3) Knudsen, V. O., *Architectural Acoustics*, John Wiley and Sons, Inc., New York, 1932, Chap. III.
(4) Miller, D. C., *The Science of Musical Sounds*, The Macmillan Company, New York, 1916, Lectures II, III, IV, and V.
(5) Robinson and Robinson, *Readings in General Psychology*, University of Chicago Press, Chicago, 1923, Chap. VIII.
(6) Watson, Floyd R., *Sound*, John Wiley and Sons, Inc., New York, 1935.



CHAPTER XI

Musical Instruments, Speech, and Auditoriums

Musical instruments in general have two parts: the *generator* of the vibrations—vibrating strings, for example—which the performer controls and to which he supplies energy; and the *amplifier*—a soundboard, for example—which aids in rapidly and effectively unloading the energy of the generator in the form of sound waves in air. The amplifier cannot give out tones not received from the generator, and tones, though received, may not be heard unless the amplifier reproduces them. Therefore, we hear nothing from a musical instrument except that which is produced by the generator, and practically nothing which is not in turn reproduced by the amplifier. The amplifier, therefore, may be the most important part of a musical instrument.

Amplification

Amplification by resonance. In Chapter VIII we demonstrated how resonance may be established between two coupled vibrating bodies of the same frequency. In a similar manner, when a vibrating tuning fork is coupled to an air column, resonance may be expected if the frequency of one of the column's natural modes of vibration corresponds to the frequency of the fork. As a test of this inference, let the following experiment be tried: A tall glass jar is filled with water and a large glass tube which is just small enough to fit inside the jar is selected. By inserting the tube and moving it up and down, an excellent means of varying the length of an air column is established. A tuning fork is now struck and held over the air column

while its length is being changed (Fig. 125). At a definite position, the sound is greatly amplified. At this particular length, one of the natural frequencies of the air column corresponds to the frequency of the fork, and resonance is established. In this case the fork is the generator and the air column the amplifier.

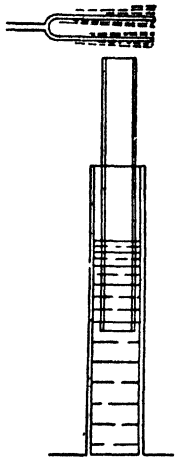


Fig. 125. Resonant Air Column.

When a mounted tuning fork is removed from its box, struck, and then brought near the open end, resonance is detected. In a mounted tuning fork, therefore, the fork is the generator, the air column the amplifier, and the top of the box the coupling agent.

Amplification by forced vibration. If an unmounted tuning fork is struck and then held firmly against a table top, increased loudness is heard. The table top is set into forced vibration by the up-and-down motion of the fork handle (Fig. 94b). Because of the large air contact furnished by the table, energy is radiated rapidly and the intensity of the sound is increased.

In stringed instruments, the vibrating strings transmit rhythmic impulses to the body or soundboard of the instrument. When properly designed and constructed, the thin sheets of wood respond to forced vibrations, and because of their large contact with air, they radiate the energy of the strings much more rapidly and efficiently than the strings could do it acting alone. One may verify this statement by listening to tones from a piano. Most of the sound comes from the soundboard (Fig. 126).

Resonance in a piano or violin is very objectionable when it occurs within the range of frequencies played on the instrument, because whenever a resonant frequency is struck, the loudness of the tone is abnormally great. Good instruments have their main resonant frequency, if they have any at all, below the range of frequencies used on the instrument. They respond evenly to all the tones played.

Poor violins may have a resonant frequency within the pitch range of the instrument, and a "howling wolf" note is attained at resonance.

Stringed Instruments

Vibrating strings. In Chapter VIII we made a study of standing waves on wires and discovered that a string may vibrate naturally in any number of equal parts. In Chapter X we explained that the quality of a tone depends upon what pure tones combine to form the complex sound, and just how these components are related in intensity. A vibrating wire we found capable of emitting a complex tone made up of components in a harmonic series, because it vibrates naturally as a whole, in halves, in thirds, in fourths, and so forth, all at the same time. We discovered that the pitch of a complex sound is determined by the frequency of the fundamental of the harmonic series, whether present in the tone or not. Hence, the pitch of a string which is vibrating as a whole and in halves, thirds, and so forth, must be determined by the frequency of the vibration as a whole, the fundamental mode of vibration. We deduced the equation for the frequency of this vibration and found it (Equation 9.3) to be,

$$n = \frac{1}{2L}\sqrt{\frac{T}{d}} \tag{11.1}$$

when L is the length of the wire, T its tension, d its linear density, and n its frequency, which we now know is a true index of its pitch.

This means that it is possible to change the pitch of tones played on a stringed instrument by varying the length, tension, and linear density of the strings. On a piano, the lengths and linear densities are fixed, and, except during the process of tuning, the tension is presumably constant. A variation of pitch is obtained when different keys are struck. As one looks at the many strings of this instru-

ment, he is impressed with the fact that a control of pitch has been accomplished by the use of the above-mentioned variables (Fig. 126). The pitch is made acute by a shortening of the wire and a decrease of the linear density; the pitch is made grave by a lengthening of the wire and an increase of the linear density to the point of the wrapping



Fig. 126. Strings and Soundboard of a Piano.

of extra wire on the main wire. Tension is changed in the process of tuning.

The law of vibrating strings. We may now express in words the facts contained in the formula: (*a*) If the tension and linear density are kept constant, *the pitch varies inversely as the length;* that is, the pitch is doubled if the length is halved, and so forth. (*b*) If the length and linear density are kept constant, *the pitch varies directly as the square root of the tension;* that is, the pitch is doubled if the tension is quadrupled, and so forth. (*c*) If the length and tension are kept constant, *the pitch varies inversely as the square root*

of the linear density; that is, the pitch is doubled if the linear density is made one-fourth as great, and so forth.

The piano. The strings of a piano are *struck* with felt-covered hammers (Fig. 126). The vibrations are forced upon the soundboard, which radiates to the air most of the sound heard. The lower notes are very weak in fundamental, but have as many as 42 harmonics. The higher notes have few harmonics, and often the loudest component is the first overtone. The middle tones are composed of ten or more partials with intensities well distributed.

The piano, no doubt, may be classed as the most musical of instruments, since the pianist can obtain great varieties of "tone color" by grouping the notes into many chords and combinations. Miller has shown that the tones of a string played at equal loudness have the same quality quite independently of how the keyboard is struck. This conclusion does not agree with the popular belief that a unique "touch" may bring to pass superior quality. Yet, the control of the variation in intensity and of the relative time of beginning the several tones of a chord or combination adds greatly to the general "tone color" and characterizes the artist. If these characteristics are added to player-piano records, an automatic piano may reproduce to a surprising degree the excellence of hand playing.

The violin. The strings of a violin are bowed. The vibrations produced are transmitted through the bridge to the body and then to the air. The fundamental is weak in the *G*-string, but the upper three strings (*D*, *A*, and *E*) have strong fundamentals. The tones from the three lower strings (*G*, *D*, and *A*) have important harmonics as high as the fourth overtone (Fig. 116), while the *E*-string, according to Miller, has a very prominent second overtone. Fletcher has shown that a violin playing *G* above middle *C* loses its violin quality if all the harmonics *above* the first overtone are eliminated, and Miller concludes from his study that the tone of a violin is characterized by *the prominence of the second, third, and fourth overtones.*

The bowing of a string (Fig. 127) is made up of the following operations: (*a*) Let the bow and string be at rest. Then, because of the friction of the rosin, the hair grips the



Fig. 127. Bowing the Strings of a Cello.

string; and as the bow moves forward, the string is moved to one side until the backward pull causes a slipping to begin. (*b*) Friction is less during slipping than when there is no relative motion. Thus, the string moves in a direction opposite to the motion of the bow toward the zero position,

passes it, comes to rest and begins to move in the direction of the bow. (*c*) Again the hairs grip the string and it is urged to move in the direction of the bow. Thus, the string vibrates while the bow moves in one direction.

Through such a complex process, a great variety of tone qualities may be produced by changes in bowing. This gives the violin the possibility of being one of the most expressive of the musical instruments. Hector Berlioz (1803–1869) said, "Violins particularly are capable of a host of apparently inconsistent shades of expression. They possess, as a whole, lightness, grace, accents both gloomy and gay, thought, and passion. The only point is to know how to make them speak."

Wind Instruments

Air columns. In Chapter IX, we found that an air column closed at one end has natural modes of vibration with frequencies in the ratio 1:3:5:7:9 . . . , the fundamental having a wave length equal to *four* times the length of the column. An air column open at both ends, we found, has natural modes of vibration with frequencies in the ratio 1:2:3:4:5:6:7 . . . , a complete harmonic series, and a wave length of the fundamental mode equal to *twice* the length of the air column. Thus, with a proper generator, an air column may be expected to vibrate in various modes, producing a fundamental and a series of overtones. The frequency of the fundamental (and therefore the pitch of the tone emitted) is controlled by the length of the air column, the column *closed at one end being an octave lower for the same length than the one open at both ends.* The quality of the tone produced depends on which overtones are present and on how their intensities are related. The closed column may never produce the even-numbered harmonics, and hence it lacks somewhat in tone quality possibilities.

Generator and amplifier of wind instruments. Wind instruments are of two types. Some use a thin stream or

jet of air as the generator; others use a mechanical vibrator, such as a reed or the player's lips. Thus, by the aid of a fluttering jet or vibrating lips or reeds, the air column of the instrument is set into resonance and a complex sound made up of the fundamental and certain other harmonics is emitted. The amplifier (the air column) in such instruments controls the pitch, the jet, reed, or lips (the generator) taking on the vibration frequency of the air column (Fig. 128). The quality of the tone is controlled by the shape of the air column and the blowing.

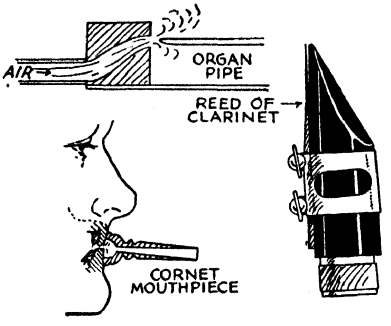


Fig. 128. Illustrating Vibrating Jets, Lips, and Reeds.

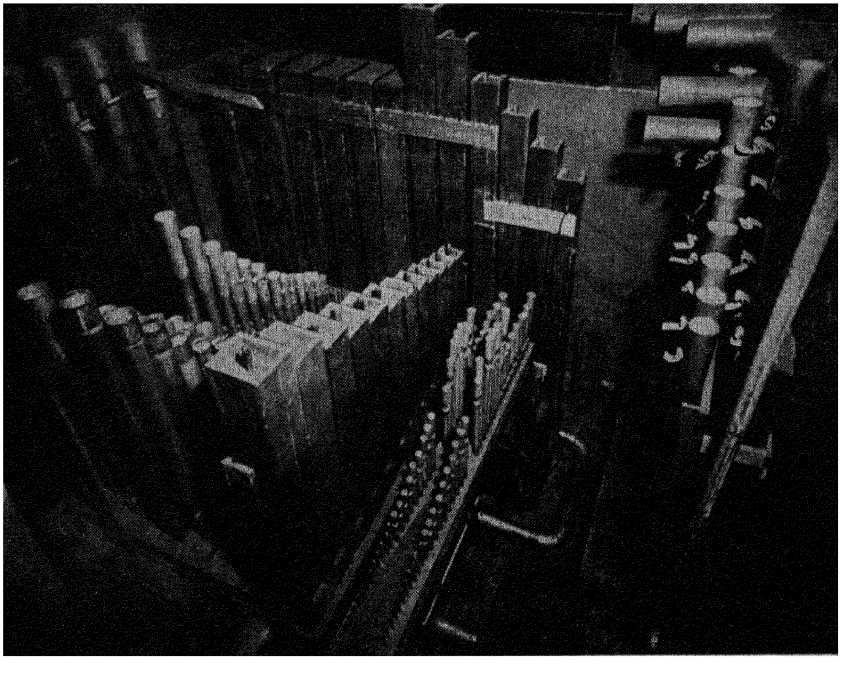


Fig. 129. Pipes of an Organ.

The pipe organ. The pipe organ is an assemblage of organ pipes of fixed pitch (Fig. 129). The pipes may be of

the open or the closed variety, usually the open type. By the act of blowing, air is caused to flutter back and forth across the edge of an opening in the pipe (the lip) and a resonant tone is produced (Fig. 128). If the opposite end of the air column is closed, it acts as a column closed at one end, and the pipe is of the closed variety. If open at the far end, the column acts as one open at both ends, and the pipe is of the open type. The quality of the tone produced depends upon the pressure of blowing, the usual tone being made up of the fundamental and only a few overtones. By "overblowing" of an organ pipe, not one tone but a series may be produced. The fundamentals of these new tones will have frequencies corresponding to the harmonics of the "normal" tone produced by regular blowing. In an "overblown" closed pipe, the tones with fundamental frequencies corresponding to the even harmonics will be missing.

The wave length of a tone produced in an organ pipe is fixed by the length of the air column. However, the speed of sound changes with the temperature, and the pitch of the pipe will likewise change, an increase of temperature increasing the pitch, and vice versa. (Remember that $V = n\lambda$. Since λ is fixed by the pipe, n is directly proportional to V.) Therefore, if an organ is to remain in tune, its temperature must be kept constant. It is possible to tune the pipes of an organ, as well as other wind instruments, by changing the length of the air column.

The flute. The flute is essentially an open pipe. As one blows across the hole near the end of the tube which incloses the air column, an irregular fluttering is produced. Among the many vibrations will be those which are in resonance with the air column. These will be sustained and all the others will be damped out, and a tone characteristic of the air column will be emitted. The holes in the body of the flute serve to control the effective length of the air column and thus determine the frequency of the fundamental mode of vibration and the pitch of the note. The upper registers

are obtained by "overblowing." The fundamental components of the notes of the middle register have the same frequencies as the first overtones of those of the lower register. The upper register makes use of the second and higher overtones of the lower-register tones as the fundamentals of its notes.

Miller has found that the tones of the lower and middle registers are characterized by a few overtones, the first one predominating. The tones of the higher register are practically pure tones. Speaking of the flute, Berlioz said, "If it were required to give a sad air an accent of desolation and of humility and resignation at the same time, the feeble sounds of the flute's medium register would certainly produce the desired effect."

The clarinet and oboe. In the clarinet the vibrations are initiated by a reed of bamboo which vibrates against the opening of the mouthpiece (Fig. 128). This reed is flexible and takes on the frequency of the standing air waves produced in the body of the tube. The tube is cylindrical in shape with a short bell-shaped enlargement at the end. Keys control the frequency of the standing waves by changing the effective length. It should be pointed out, however, that because of the action of the reed this instrument performs somewhat like a closed pipe. It practically fails to emit the first, third, and fifth overtones, as its sound spectrum (Fig. 116) shows. For this reason, when it is "overblown," it does not emit the octave, as does an open pipe such as a flute, but the *twelfth*, as does a closed pipe. Because of this fact, the key mechanism is usually complicated. The characteristic clarinet quality is probably due to the seventh, eighth, and ninth overtones.

The oboe has an air chamber which is conical throughout. Its mouthpiece consists of two thin slips of bamboo placed nearly in contact. Being flexible, these reeds take on the vibration of the air column, which is controlled in length by keys. This instrument acts as an open pipe and has three registers, the middle register being the most agreeable. Its

characteristic quality is probably due to the third, fourth, and fifth overtones.



Fig. 130. The Horn.

Brass instruments. By proper blowing, an open pipe may be made to emit a *series* of notes with fundamental frequencies in a harmonic series. If we call the lowest note

do, then the instrument may be made to emit also tones of the pitch, *do'*, *sol'*, *do''*, *mi''*, and so forth, without changing the effective length of its air column. The ease with which these different harmonic tones may be blown depends upon the shape of the air column. If the width of the tube is relatively great as compared with its length, the first harmonic tone will be easy to sound, as in the case of the tuba. When the bore is narrow, as in the horn (Fig. 130) and the trumpet, the lowest harmonic tone is difficult or impossible to blow and the higher ones are easy. The best range for the horn is from the fourth to the twelfth harmonic tone. A wide, spreading bell, as in the horn, makes the tone smooth. The horn under average playing produces tones with a strong fundamental followed by a complete series of more than twenty overtones which gradually decrease in intensity. Of this tone quality, Lavignac remarks, "It is perhaps in the expression of tenderness and emotion that it best develops its mysterious qualities."

A small bell, as in the trombone, produces brilliant quality. The type of mouthpiece also affects the quality. A shallow, cup-shaped mouthpiece gives brilliance; a mouthpiece gradually narrowing from the rim gives a smoother, more mellow tone.

The Vibrating Diaphragm

The telephone. Sound waves strike the diaphragm of the transmitter and force upon it the characteristics of the sound. These vibrations are turned into a pulsating electric current by a process to be described in Chapter XVI. The current is conducted by a wire to a distant telephone receiver, where the electrical pulsations are changed into vibrations of the diaphragm (Chapter XVI). The passing of the sound characteristics through this electrical circuit without distortion is one of the great modern accomplishments. The Bell Telephone System, named after the inventor of the telephone, has done most in this development.

The phonograph. The wave form of the original sound is cut into a wax disc, from which records are made. The needle runs along the wave (or hill-and-dale) track and sets a diaphragm into vibration, and this in turn sends out sound waves through a horn (Fig. 131). Before the days of electric (orthophonic) recording, the sound waves were gathered up by a large horn, a diaphragm at the end was made to vibrate, and this mechanical vibration was cut into a wax disc. The artists were grouped very close to the horn, and the weaker instruments, such as violins, which could not get very near because of limited space, had special sounding diaphragms and horns for intensifying the sound. With the electrical recording, artists may play under normal conditions. A sensitive, high-quality microphone turns the sound waves into electrical pulsations. These are amplified and then sent into an electromagnetic recorder which transforms the electrical pulsations into mechanical vibrations and cuts a record into a revolving wax disc. The electrical and mechanical systems are designed so that the wave forms of the complex tones are faithfully recorded in the wax disc, something which the old method did not do. This new method records especially the low- and high-pitched tones with greater exactness. In recording, the microphone is sufficiently far from the artists so as to receive the sound properly blended by reflections from the walls of the room, and so-called "atmosphere" or "room-tone" is obtained.

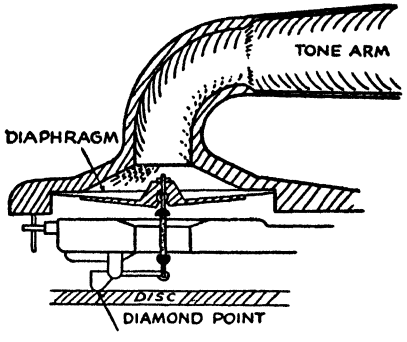


Fig. 131. Hill-and-Dale Phonograph Reproducer.

The Radio

Because of the common expression "on the air," one might suspect that sound waves are responsible for the

transfer of sound from the broadcasting station to the home. But this is not so. *Sound waves are sent out from the performer to the microphone in the studio,* but the wave form is there transferred from moving air molecules to a moving diaphragm, and thence to an electric current which carries the sound characteristics from the studio to the broadcasting plant. This plant may be many miles away, as in the case of national hookups, where currents are sent from New York City to the principal cities of the country. At the broadcasting plant, the sound characteristics are sent out on radio waves, which travel with the speed of light. These are not air waves but represent the long-wave-length side of a great family of electromagnetic waves in which X-rays represent the short-wave-length side.

At the home, these now very feeble radio waves are picked up and increased in energy content by the aid of the radio energized by the home lighting service. An electric current bearing the same variations as the current sent from the studio to the broadcasting plant is produced. This current is sent through a loud-speaker, which by means of mechanical vibrations re-establishes the characteristics of the original sound in the form of sound waves. The artist is heard almost without a lapse of time, so rapid is the speed of electricity and radio waves. *Air waves play a part in the studio and the home but not between them.*

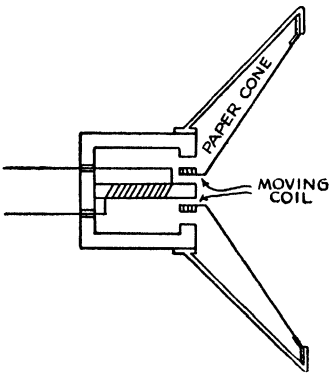


Fig. 132. The Loud-Speaker.

The so-called "dynamic" loud-speaker is shown in Fig. 132. The vibrations evoked in the paper cone by the action of a pulsating electric current are the source of the sound heard. In order to reproduce frequencies covering a wide range, two speakers are often used. One, because of its large size and design, responds to the lower-frequency range; the other, because of its small-sized cone and light

moving parts, responds to the higher-frequency range. If reproduced music or speech is to sound natural, the whole system must be designed so that all important components get through without distortion. Modern improvement in the radio has been due in the main to the widening of the frequency range of faithful reproduction. The need of a wide frequency range is clearly shown in Fig. 133, which gives the frequency range, as determined at Bell Telephone

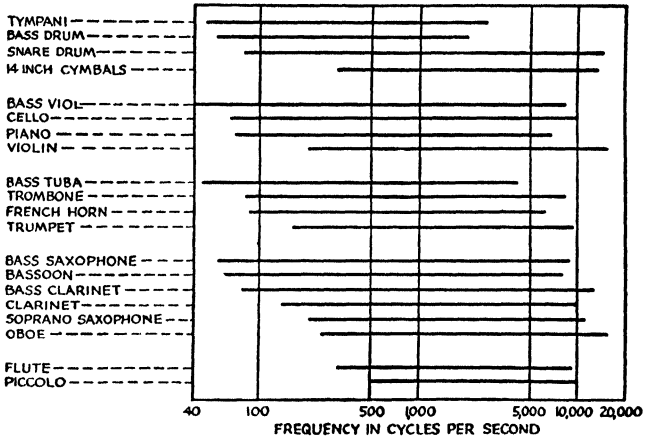


Fig. 133. The Frequency Ranges of Instruments. (*After Snow. Courtesy of Bell Telephone Laboratories.*)

Laboratories, required by various instruments if their tones are to be faithfully reproduced.

The Public Address System

This system consists essentially of a microphone which converts sound waves into electric impulses, an amplifier which feeds in energy, and finally, a loud-speaker which, because of the energy added, is able to reproduce a sound louder than the original. For the reasons explained above, a good address system must have an ability to amplify a wide frequency band without distortion.

The Human Voice

This familiar instrument is composed essentially of three parts: the *stimulator*, a stream of breath produced by the

bellows-like action of the lungs; the *generator*, the vocal cords; and the *modifiers*, the throat, mouth, and nasal cavities (Fig. 134).

The stimulator. For best results, the stream of air from the lungs which sets the vocal cords into vibration should be under perfect control. This requires skill in the use of

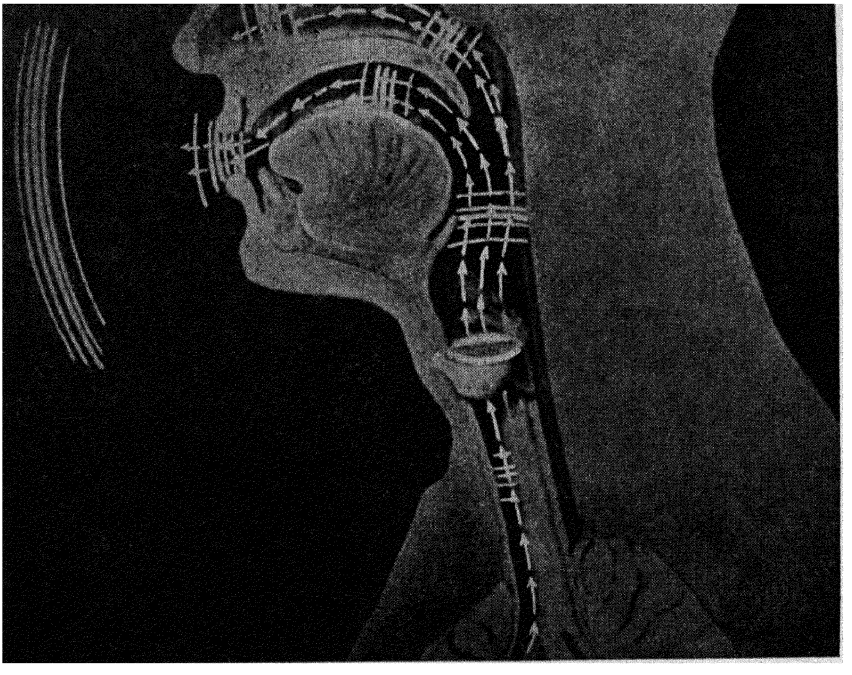


Fig. 134. The Human Voice. (*From "Sound, A Guide for Use with the Educational Pictures 'Sound Waves and Their Sources' and 'Fundamentals of Acoustics,'" by Lemon and Schlesinger. Courtesy of The University of Chicago Press.*)

the muscles of inhalation and exhalation: inhalation should be free and unrestricted, and exhalation adequately controlled.

The generator. The sound produced by the human voice is generated in the main by the vibration of the vocal cords, a pair of muscular shelves mounted on the sides of the larynx (Adam's apple) and jutting out in such a manner as to form a slit (glottis) between the outer edges. A cross section of a muscular shelf would look somewhat like a right

triangle with the angle at the glottis edge rather acute. The top of the shelf is practically flat and perpendicular to the side of the larynx, and the lower surface slopes downward from the edge at the glottis to the larynx wall—the hypotenuse of the triangle. During quiet breathing, the glottis is V-shaped—the shape of a V with its vertex pointing out toward the tip of the chin and the arms toward the spinal column. The closure necessary for vocalization is made by pushing the arms of the V together.

The exact manner in which the vocal cords (shelves) or glottal lips function in the production of sounds is still not clearly known, but it may not be amiss to present the following as a much simplified explanation of the action of the vocal cords as the vibration rate increases. Under small muscular contraction, the glottal lips vibrate nearly as far back as the supporting larynx wall. As muscular contraction increases, the region of vibration is more and more limited to the outer portion (that nearest the glottis) of the shelves; finally, only the very edges next to the glottis vibrate. Thus, the frequency increases with an increase in tension and a decrease in the mass of the vibrating material.

By the aid of X-ray photographs and the laryngo-pariskop (an optical device which permits the visual observation of the vocal cords), Dr. G. Oscar Russell (6) finds that the edges of the glottal lips tend to become round when a very soft tone is sung. On the other hand, when a loud, harsh, high-pitched sound is produced, the false vocal cords and the muscular surfaces which lie immediately above the vocal cords bear down and close in to the extent that only the thin edges of the glottal lips are able to vibrate. These edges are pushed up by the air and probably are made to flap against each other—presumably the cause of the harsh sound. Harshness also results from globules of mucus which gather on the edges of the glottal lips. Thus, the tone quality of the voice seems to depend in some measure upon the mode of vibration of the vocal cords.

The modifiers. Undoubtedly, the various air columns and surfaces of the throat, mouth, and nasal cavities act as modifiers of the tone quality of the sound initiated by the vocal cords. There is much difference of opinion as to the relative importance of these various modifiers. However, it is safe to say that the tone quality depends upon the following factors: (1) the mode of vibration of the vocal cords (discussed above); (2) the nature of the surrounding and overhanging tissue; (3) the "megaphone" action of the mouth; (4) the frictional noises produced by the motion of air through constricted openings of the oral and nasal cavities; (5) *absorption* of the high-frequency components by the soft surfaces of the cavities; (6) the *resonant action* of the head cavities, especially the large cavity between the throat and the lips.

Speech power. The average speech power delivered by a typical speaker is from 100 to 200 ergs per second; if he is talking to a large audience, his speech power may *average* 1,000 ergs per second; if he becomes vehement, the power may *rise* to 20,000 ergs per second. This is a very small amount of power. In lifting a book through the height of one foot, the speaker expends as much energy as he would transfer to sound waves during a four-hour speech. But when we remember that the ear is so sensitive as to be able to detect the flow of one-billionth of an erg per square centimeter per second at high frequencies, and one-ten-thousandth of an erg per square centimeter per second at low frequencies, we understand why a very small energy flow is adequate for hearing.

Frequency distribution of energy in speech. The energy in speech reaches a maximum at about 220 cycles per second, practically all the energy being at frequencies below 500. Since the fundamental pitch of a man's voice is between 100 and 125 cycles per second, and a woman's voice between 200 and 250 cycles per second, it is clear that most of the energy is carried by the fundamental component. Remembering that the ear is

more sensitive to high frequencies than to low and that its response is logarithmic to sound intensities, we find, when approaching this problem from the standpoint of the ear, that the maximum occurs between 500 and 1,000 cycles per second and falls off rather gently for the higher frequencies.

Speech sounds. In general, the vowels are produced when the speech organs modify the oral cavities without obstructing the free passage of the outgoing breath. On the other hand, the consonants are produced when certain parts of the speech mechanism more or less obstruct the breath stream. Thus it follows that the boundary between vowels and consonants is not sharp. For example, the letter **w** may be thought of as a consonant in **w***on* and as a vowel in *n*o**w**, and **y** as a consonant in **y***ard* and as a vowel in *dr***y**.

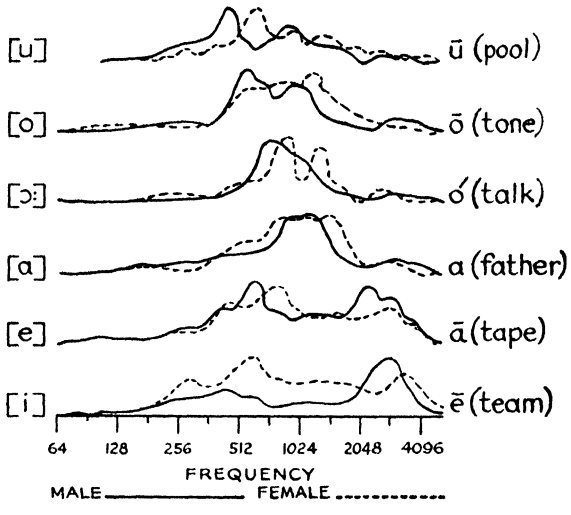


Fig. 135. Energy Distribution in Vowel Sounds. (*After Crandall. Courtesy of Bell Telephone Laboratories.*)

The vowels. The vowels may be spoken or sung at various pitches. The frequency distribution of energy characteristic for each vowel sound, then, in the main must be independent of the fundamental frequency at which it is sounded. Fig. 135 is a set of energy-distribution curves (corrected for the sensitivity of the ear) for six vowels—a small part of the excellent investigation reported by the late I. B. Crandall, of Bell Telephone Laboratories. We note that the vowels here presented are characterized by two or even more frequency bands.

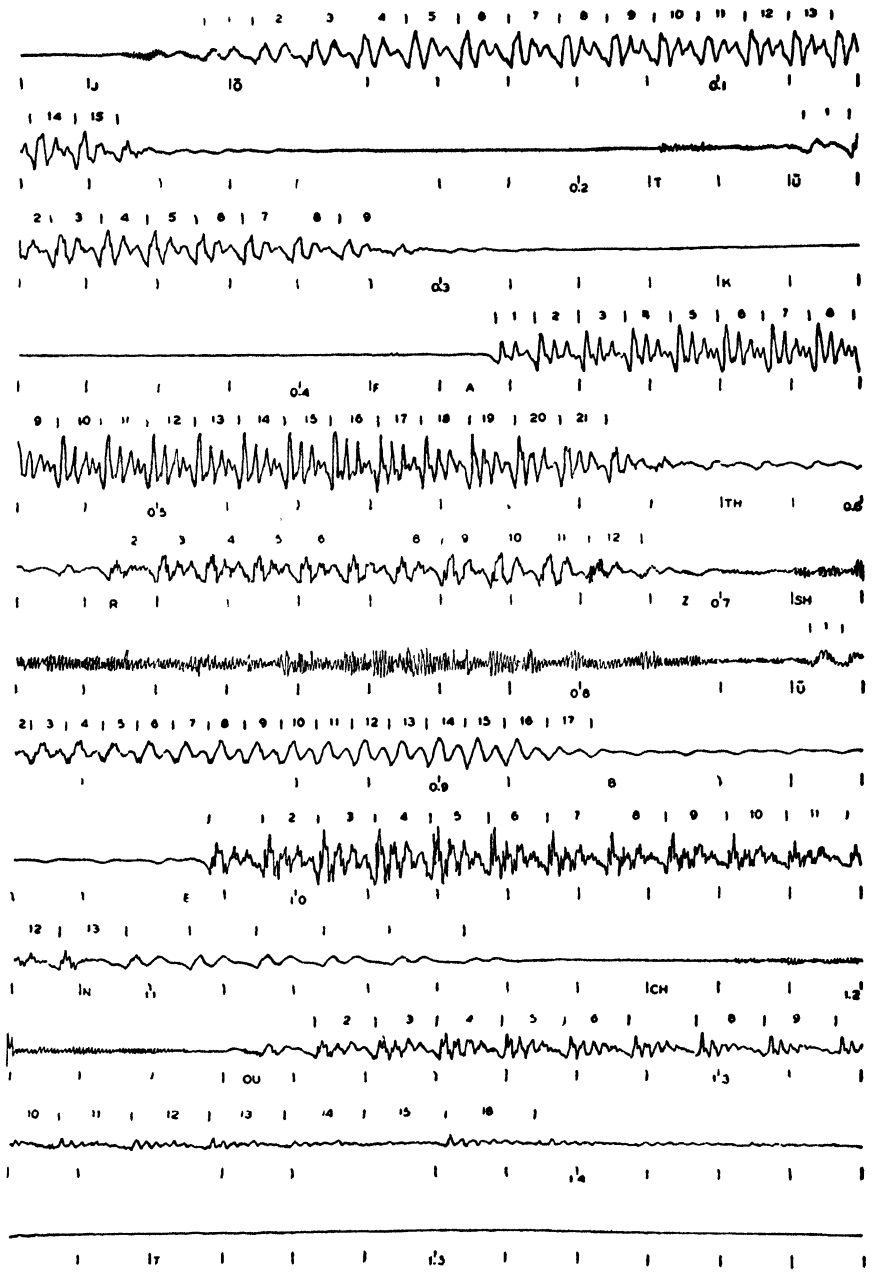


PLATE IV. An Oscillograph Record of the Sentence, "Joe took father's shoe bench out," Spoken by a Man at a Rate Somewhat Faster Than Normal. The time marks just under the waves are intervals of 0.01 seconds. The vowel sounds approximate recurrent waves, as is shown by the marks immediately above the wave tracks. (*Steinberg. Courtesy of Bell Telephone Laboratories. See The Journal of the Acoustical Society of America, Vol. VI, No. 1, p. 16, July* 1934.)

Fletcher, summarizing the work of Stumf, Miller, Paget, and Crandall, lists the characteristic frequencies of certain vowel sounds as follows:

VOWEL SOUNDS

Speech Sounds	*Tone Quality*	*Characteristic Frequency*	
		Low	*High*
pool, boob, truth.....	dullest	400	800
put, foot, cook.....		475	1000
tone, go, pope.....		500	850
talk, law, salt.....		600	950
ton, sun, done.....		700	1150
father, palm, ah.....		825	1200
tap, cat, sat.....		750	1800
ten, get, pep.....		550	1900
pert, her.....		500	1500
tape, they, café.....		550	2100
tip, it, sit.....		450	2200
team, peep, machine..	brightest	375	2400

The components of the characteristic *high*-frequency range of the first six vowel sounds listed in this table are very much weaker than those for the last six vowel sounds. For this reason, the first group is often referred to as being *singly* resonant, while the second group is spoken of as being *doubly* resonant. The vowel sounds of the first group, because of the weakness of high-frequency components, have a dull or mellow quality; those of the second group, because of the prominence of these components, have a bright or even metallic quality.

The positions of the characteristic frequency bands change somewhat from speaker to speaker and even with the same speaker from time to time. Considerable variation is found in the results reported by different investigators. For example, the following data were obtained for the vowel sound in *took:*

CHARACTERISTIC FREQUENCY OF THE VOWEL SOUND IN *took*

Investigator	Resonant Regions			
	First	*Second*	*Third*	*Fourth*
Sir Richard Paget (1923)..	362	966	——	——
J. Q. Stewart (1922).....	420	960	——	——
I. B. Crandall (1925).....	460	960	——	——
J. C. Steinberg (1934).....	420	1650–1300	2400	3200

In 1837, Charles Wheatstone (1802–1875) advanced the hypothesis that the vocal cords generate a complex tone composed of a fundamental and overtones in a *harmonic series*. When the sound waves pass through the throat, the mouth, and the nasal cavities, those components of the original tone which are in resonance or near resonance with the head cavities are radiated into the air much magnified. For example, with the vowel sound of the word *pool*, the lips are rounded and a large resonating cavity is formed in the front part of the mouth and a small one in the throat. As the vowel sound is changed to that found in *father*, the mouth is gradually opened and the tongue is flattened out. In all of these vowels, the throat resonance is playing a minor role. With a further shift to the vowel sound in *peep*, the tongue is raised in the front of the mouth, two resonance chambers are formed, and *both* produce marked effects upon the vibrations originating in the vocal cords. This is considered by many to be an explanation of the frequency bands which seem to characterize the vowel sounds.

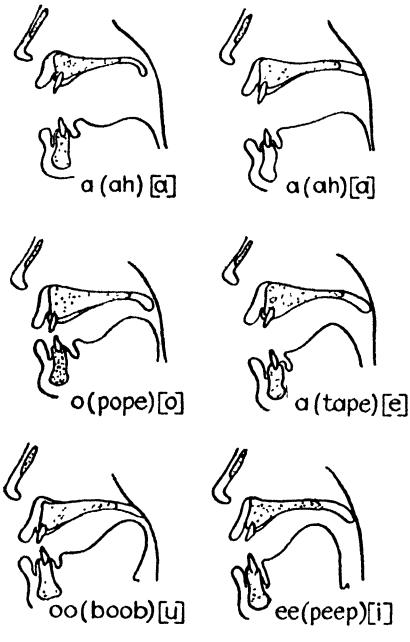


Fig. 136. Oral Cavities for Certain Speech Sounds. (*After Russell.*)

X-ray photographs made by Russell (6) cast some doubt upon this hypothesis. Russell has shown that oral cavities of two greatly different shapes may be used by the same speaker in producing a perfectly normal **ah**. Such cavities are shown diagrammatically in Fig. 136, in which the essential features of the X-ray photographs are preserved. He reports that many of the vowels, considered by all investigators to be doubly resonant, are "regularly manifest

in different subjects with cavities which can hardly be considered as dual." He concludes that "the vocal cord function and the surfaces and walls in the vocal cavities may be quite as important" as the resonant action of the air they contain. Finally, he considers that the vowel sound in *father* is practically a pure glottal tone, a tone heard very much as it is produced by the vocal cords; and he shows by means of X-ray photographs (Fig. 136) that the vowels with brighter qualities (those in *peep* and *tape*) are formed by constrictions made against hard surfaces (hard palate), while the vowels with duller qualities (those in *boob* and *pope*) are formed with constrictions between soft surfaces. This explanation may account, in part at least, for the lack of high-frequency components in the dull tones.

Thus it is clear that the theory of vowel production is in the making. Much excellent experimental work has been accomplished, but more needs to be done before an adequate understanding of vowel production may be evolved.

Diphthongs. A diphthong is the combination of two vowel sounds in the same syllable. The following are examples: a*isle*, h*ouse*, n*oise*, f*eud*. Digraphs, two letters which stand for one sound, should not be confused with diphthongs. The following are examples of digraphs: s*aid*, f*ruit*, h*eart*.

Consonants. In the formation of the consonants, the lips may be parted, brought together, spread, rounded, or protruded; the jaws (and the teeth) may be brought together or parted; the tongue may be drawn forward or back, or it may be raised or lowered in the front, in the middle, in the back, or at the sides; the velum (soft palate) may be raised or lowered; and the vocal cords may vibrate, producing a voiced consonant, or a voiceless sound may be produced by the frictional noises originated by the motion of air through the constricted openings of the oral cavities. For example, the only difference between **p** and **b** is that the former is voiceless and the latter is voiced.

The breath stream may be checked for an instant and then allowed to issue as an explosive puff, as in the consonants **p, t,** and **k.** The air current may be checked in the mouth but allowed to pass out through the nose, producing **m, n,** and **ng.** If the stream of air is only partially blocked, so that it flows out in a continuous stream through the constrictions formed by the obstructions, such consonants as **f, s, l,** and **z** are formed. Obstructions may be made

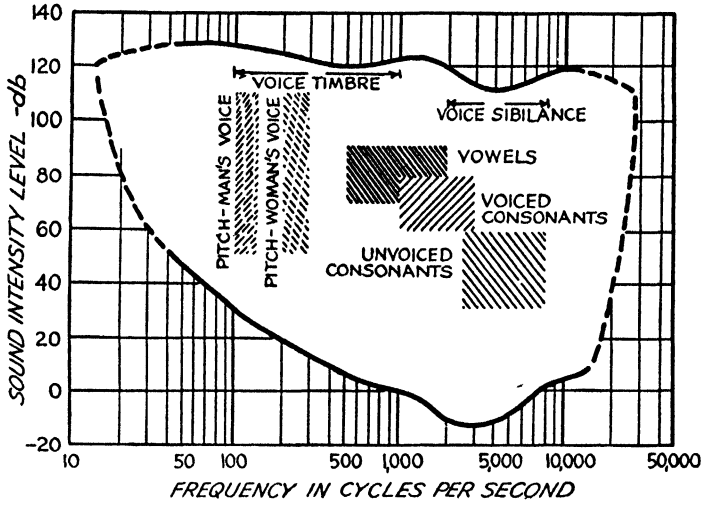


Fig. 137. Frequency and Intensity Level Ranges for Vowels and Consonants. (*After Steinberg. Courtesy of Bell Telephone Laboratories.*)

with the lips (**p, m**), with the lip and teeth (**f, v**), or with some part of the tongue against the teeth or against some part of the hard or soft palate (**th, t, l, k**).

As we have explained, certain consonants are voiceless, the sound being produced largely by vibrations set up by the passage of air between lips, between tongue and teeth, or between tongue and palate. Such sounds are composed of relatively high frequencies and have an energy content very small as compared with that present in vowels. This small energy content of certain consonant sounds is a handicap, since the accurate recognition of speech depends much more upon the consonant sounds than upon the vowel

sounds. Over 50 per cent of the errors made in recognizing speech syllables are due to **th, f,** and **v,** no matter what the sound intensity; while at low intensities, **z, h,** and **s** are especially difficult to understand.

The approximate frequency and intensity level ranges for vowels and consonants, as determined by J. C. Steinberg, of Bell Telephone Laboratories, are given in Fig. 137. Roughly, we may conclude that most of the energy of speech is carried by frequencies below 2,000 cycles per second, but the essential characteristics which determine its recognizability are carried by frequencies above 2,000 cycles per second.

The phonetic alphabet. The English alphabet contains 26 letters, which must represent about 40 sounds. Obviously, certain letters or combinations of letters must stand for more than one sound. In spite of this double duty, there are instances in which the same sound is represented by more than one letter. All this means that a phonetic alphabet is much to be desired. To meet such a need, Paul E. Passy devised the International Phonetic Alphabet now used by the International Phonetic Association. Such an alphabet (not complete in all details) is here shown.

PHONETIC ALPHABET

[ɑ]	father	[aɪ]	aisle	[f]	foe
[a]	ask	[aʊ]	house	[v]	vow
[æ]	cat	[ɔɪ]	noise	[θ]	thin
[ɛ]	get	[ju]	pure	[ð]	this
[ɛɪ]	there	[p]	pa	[s]	so
[e]	café	[b]	be	[z]	zero
[ɪ]	it	[t]	to	[ʃ]	she
[i]	machine	[d]	do	[ʒ]	azure
[ɔ]	cloth	[k]	key	[l]	lie
[ɔɪ]	law	[g]	go	[r]	rest
[o]	go	[m]	me	[ɹ]	are
[ʊ]	put	[n]	no	[j]	yes
[u]	truth	[ŋ]	sing	[h]	he
[ə]	sofa	[ʍ]	why	[tʃ]	chin
[ʌ]	cup	[w]	wet	[dʒ]	joy

Good singing voice quality. Consulting the studies made by Bartholomew (2) and by Wolf, Stanley, and Sette (10), we may list the following four main characteristics of a good-quality singing voice:

Vibrato. The periodic change in the three variables: pitch, loudness, and tone quality, known as *vibrato*, should be smooth and even and should occur about *six* or *seven* times per second. In good voices, the louder the tone, the more prominent is the variation. The vibrato used by Giovanni Martinelli has a frequency of about 5.7 per second, the variations in intensity being smooth and of an amplitude of from one to four decibels. Inartistic voice "tremolo" results from unco-ordinated muscular action, from extreme variation of one or more of the variables, and from a failure to make the variations like a smooth wave motion, and often it is faster than six or seven times per second.

Maximum intensity. The person rated as having good voice quality is usually able, when asked to sing as loudly as possible, to produce a tone of considerably greater intensity than a person with a poor voice quality.

Low-frequency band. Good voices show a decided tendency to strengthen components of the complex tone which have frequencies in the neighborhood of 500 cycles per second.

High-frequency band. Good male voices show the presence of high-frequency components between approximately 2,000 and 3,300 cycles per second. The high-frequency band of the female voice centers still higher, perhaps around or above 3,300 cycles per second. The coloratura type of female voice often shows no high-frequency band at all. Such a voice is said to be good because of its agility and the purity of the tone produced.

Percentage articulation tests. Teaching and preaching usually take place in classrooms and auditoriums. A method of measuring the recognizability of speech sounds for purposes of checking the ability of the speaker to articu-

late or of determining the acoustic quality of the auditorium is much to be desired. Such methods have been thoroughly investigated and developed by Fletcher and Steinberg (3). One of their methods of obtaining percentage articulation, which may be used successfully by the student with little training in acoustics, is summarized as follows:

Simple English words containing the vowel and consonant sounds are spoken by a caller and the words heard are written down by the observer. The following is a list of the words used and the corresponding sounds, both vowel and consonant, which they represent.

VOWEL WORD LIST

Sounds to Be Graded	*English Words in the List*	
a	bat	back
ā	bait	bake
e	bet	beck
ē	beat	beak
i	bit	bit
ī	bite	bike
o	but	buck
ó	bought	balk
ō	boat	boat
u	book	book
ū	boot	boot

CONSONANT WORD LIST

Sounds to Be Graded	*English Words in the List*		*Sounds to Be Graded*	*English Words in the List*	
b	by	by	p	pie	whip
ch	which	which	r	wry	wry
d	die	die	s	sigh	sigh
f	fie	whiff	sh	shy	wish
g	guy	wig	th′	thy	with
h	high	high	th	thigh	thigh
j			t	tie	wit
k	wick	wick	v	vie	vie
l	lie	will	w	why	why
m	my	whim	y		
n	nigh	win	z	whizz	whizz
ng	wing	wing	st	sty	whist

Note: The **h** following **w** is not pronounced in words such as *whim*, *whip*, etc.

Each word is listed on a separate card and the cards are thoroughly shuffled, the vowel list being kept separate from

the consonant list. The particular order of words thus established forms a word list for one trial. Each additional shuffling permits the establishment of a new arrangement of words and a so-called new list. In calling, each word is preceded by an introductory statement such as "the first group is," "can you hear," "I will now say," etc. The vowel word list is called first, then the consonant word list, and the articulation for each is obtained as the percentage of words correctly observed in each case. The syllable articulation is obtained by multiplying the vowel articulation by the square of the consonant articulation. This extra consideration is given to the consonant articulation because often a syllable is composed of a consonant sound, a vowel sound, and a consonant sound. For further information, the student will find complete details of such tests in references (3) and (4), Suggested Readings.

Sound in Auditoriums

Speech in auditoriums. When we gather to listen to speech or music, we usually congregate in some hall or auditorium where we are adequately protected from wind and rain, from scorching heat or biting cold. Often we are provided with comfortable seats, oftener the walls are tinted to our liking, but seldom do we find the enclosure an ideal place in which to hear. An auditorium with ideal acoustics, some persons believe, is attained only by chance —a sort of gift of the gods that comes once in a long while and then very unexpectedly.

The late Wallace C. Sabine, of Harvard, undertook the study of rooms that were built as auditoriums but which failed to function as such because of poor listening conditions. After a long and careful study, he found that there could be built a science of architectural acoustics, the use of which would enable architects and engineers to build ideal auditoriums. Since this monumental work, investigators throughout American and European countries have further developed this science, and any builder who has in mind the

erection of a room in which the listening conditions are to be good may succeed by applying known principles of acoustics.

The rating of an existing auditorium in terms of speech intelligibility may be made by articulation tests, the approximate scale for judging the quality of the room being as follows:

Percentage Articulation	*Hearing Condition*
85 and above.	Very good
75.....	Satisfactory, but requiring attentive listening
65.....	Barely acceptable, very fatiguing
Below 65...	Unsatisfactory

The acoustics of existing auditoriums may be improved or new ones with good acoustics designed by a proper procedure involving the following factors: (1) the shape and size of the room, (2) the loudness of the sound, (3) the intensity of extraneous noise, and (4) excessive persistence of sound, or reverberation. We shall discuss these factors briefly.

Shape and size of room. When sound waves strike hard surfaces, such as plastered walls, with flat areas that are large as compared with the wave length, reflection takes place according to the simple law: The angle of incidence i is equal to the angle of reflection r (Fig. 138). By means of this law, a curved reflector with a shape somewhat similar to that used in the headlights of automobiles, but very much larger in size, may be constructed of hard plaster, metal, and wood and used to reflect speech sounds or music from speakers or performers properly located within the curved surface to the audience straight out in front. Directed reflection from such a reflector is needed and successfully used at Hollywood Bowl, a large open-air theater in which the usual sound reflection from ceiling and walls may not be utilized.

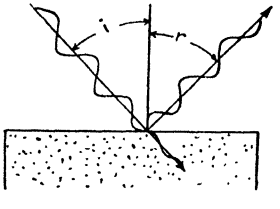


Fig. 138. Reflection of Sound.

In large enclosures, it is better practice to make use of reflection by placing *large, flat, reflecting surfaces* behind, above, below, and on both sides of the stage or speaker's platform. This is an efficient means of utilizing reflected sound and of evenly distributing it without the detrimental focusing effects of curved surfaces. Such a design is illustrated in Fig. 139. Notice also how curved surfaces, which might concentrate the sound, are carefully avoided, and how a splay is introduced at the rear of the balcony instead of letting the ceiling meet the rear wall at right angles, thus making impossible the return of the sound from this region to the front seats as a distinct echo. Further notice that the ceiling under the balcony slopes downward toward the back, thus permitting a reflection of sound down upon the seating area.

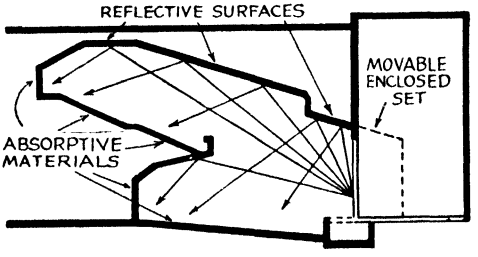


Fig. 139. Auditorium with Good Design. (*After Knudsen.*)

Pronounced focusing, echoes, and poor utilization of reflecting surfaces, defects which should not be tolerated, are shown in Fig. 140.

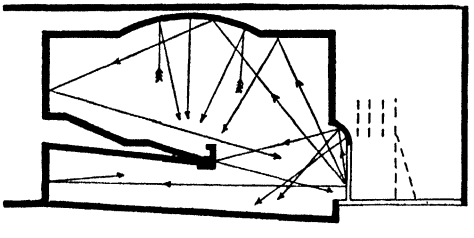


Fig. 140. Auditorium with Poor Design.

Intensity of sound. In the open, the intensity of sound (interpreted as loudness) falls off very rapidly with the distance from the source, approximately inversely as the square of the distance. In such an environment, only speakers with large speech power may be heard well by more than a very small group. The Greeks were fully aware of the inability of the average speaker to be heard in their theaters, and actors were at times supplied with masks designed to serve as megaphones in order that the sound of their voices might be concentrated upon the

audience. They also set about the stage large bronze vessels which they thought would amplify by resonance the speech power of the performers.

As the auditorium has evolved from a gentle slope on a hillside through the Greek theater to the modern enclosure, more and more of the sound energy which otherwise would have escaped has been pent up. In the open, the sound flows out into space and is lost. When a wall is placed behind the source, some of the energy is returned by reflection and the sound intensity in front of the wall is increased. The addition of another wall makes possible multiple reflections with an attendant increase in sound intensity. In an enclosure such as an auditorium, a sound wave is returned many times by multiple reflection before it is completely absorbed, and the sound intensity is correspondingly augmented. The loudness of sound in an auditorium, then, depends upon (1) the speech power of the speaker, (2) the distance from the speaker, and (3) the reflecting power of the walls and fixtures of the room.

Extraneous noise. Noise is always detrimental to articulation. The deterioration is partly overcome if the noise is masked by very loud speech. But a person unaided by a public address system may not always be able to produce sounds sufficiently high to overshadow the noise successfully. An auditorium should be located in a quiet place, therefore, and should be so constructed that noises are not easily set up in its structure. Soundproofing to eliminate extraneous sounds produced in and around the auditorium may need to be a part of the building specifications.

The noise which originates in the room itself may be reduced by the use of well-designed, upholstered seats and a carpeted floor. Outside noises enter through openings such as windows, cracks around doors and windows, and ventilating ducts. Weather-stripping doors and windows and adding absorbing material to air ducts, as illustrated in Fig. 141a, will greatly reduce air-borne noises. The other main source of outside noise is that transmitted through

partitions by the diaphragm action of the walls. Such transmission is especially noticeable between small studios separated by lath-and-plaster walls. Insulation in such cases is achieved by building two or three relatively thin partitions and separating them by blankets of absorbing material (Fig. 141b).

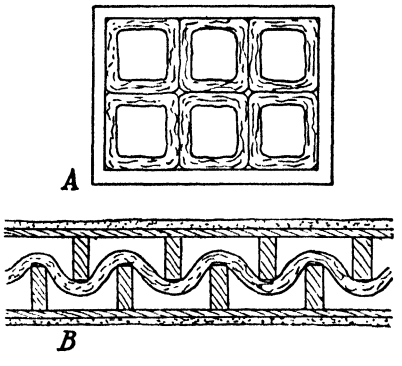


Fig. 141. Insulating Against Noise.

Reverberation. Sound waves will be reflected back and forth many times in an auditorium before they are completely absorbed. Because of this persistence of sound or reverberation, hearing may not be good, even with the speech sufficiently loud and the noise adequately reduced. If the reverberation is excessive, the syllables previously spoken will interfere with those just being spoken, and the articulation will be greatly impaired. This defect may be eliminated if the walls, floors, and ceiling are treated with *absorbing* material, such as carpets, drapes, and acoustical plaster, and if the volume of the room is reduced as much as is consistent with the seating capacity of the auditorium.

Reverberation increases the loudness of sound. From this point of view, it is desirable. But too much has a vitiating effect upon the intelligibility of speech. This means that the optimal reverberation time for various situations should be known. Curves representing the values considered by Dr. Vern O. Knudsen (4) to be optimal for speech and music are shown in Fig. 142.

It is often desirable to predict the extent of reverberation before an auditorium is built, in order that absorbing material may be specified, if needed, and especially in order to determine whether there have been included sufficient surface areas which may be covered with absorbing material without spoiling the artistic features of the room design. For this purpose, a reverberation time formula has been

worked out. It permits the calculation of the time required for sound to fall 60 decibels in intensity level after the source of sound stops. In the calculation, the volume and surface of the room must be known and the absorbing power of the materials covering the ceiling, floors, and walls (plaster, wood, glass, and acoustic plaster, a plaster constructed to absorb sound) must be obtained. By use of this predicted reverberation time, the quantity of absorbing material needed to bring the reverberation time to the

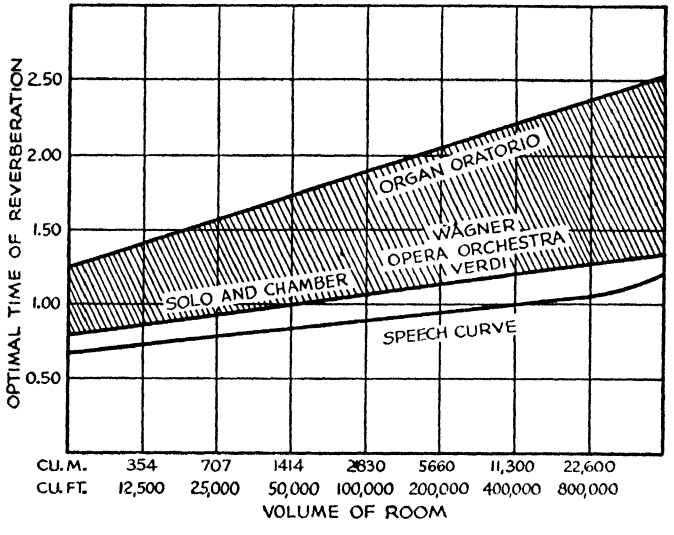


Fig. 142. Optimal Reverberation Times. (*After Knudsen.*)

optimal value for the *size* and *purpose* of the building being constructed may be determined.

Music studios and halls. The musician needs a persistence of sound, a reverberation, in order that he may hear one tone faintly blend into a succeeding one. Thus, a room with optimal reverberation time for speech will usually be too dead for music. This is clearly shown in Fig. 142.

In general, music halls should be built with a generating end composed of large areas of wood, such as a wooden floor and wood paneling, and a listening end which is relatively dead because of sound absorption at the walls,

seats, and floor—care should be taken not to carry the arrangement to an *extreme*. The wood paneling should be about one-fourth of an inch thick and be nailed to the wall on furring strips with variable spacing. Such a mounting gives to the thin boards many natural low frequencies of vibration. These panels, if varnished, will reflect the high-frequency tones with very little absorption; but owing to resonance, low-frequency waves will be absorbed, part of the energy dissipated as heat, and part radiated as colorful, rich tones. The plastered walls connecting the generating and listening ends should be painted so as to reflect high-frequency tones.

With such an arrangement, the generating end is sufficiently reverberant and resonant to give the performers plenty of opportunity to blend the tones produced; the varnished surfaces preserve the high-frequency notes which are easily absorbed; wood paneling absorbs and gives rich quality to the low-frequency tones which are not easily absorbed by hard walls and common absorbing materials; the audience, the seats, and the walls of the listening end, by absorption, keep down excessive reverberation.

The student should keep clearly in mind that reverberation is not resonance nor forced vibrations, but a persistence of sound due to *reflection* of sound waves from the walls of a room. The optimal reverberation time for music rooms is given in Fig. 142. If a room must be used for both speech and music, which is very often the case, a compromise will need to be made between the short reverberation time best for speech and the longer reverberation time needed for music.

Questions and Problems

1. Describe two distinct methods which a violinist might use to play on a given string the double octave of the tone given out by the open string.

2. If the pitch of a string is 256 cycles per second, what change (*a*) in length and (*b*) in tension will be required to change the pitch to 512 cycles per second?

3. Explain why air columns (tubes) are often placed under the bars of the xylophone.

4. If the pitch of an open organ pipe is 435 cycles per second at 20° centigrade, what will be its pitch at 0° centigrade?

5. Explain how you could test the improvement of a person's ability to articulate speech.

6. Make a list of ten words which might be expected to be difficult to understand.

7. List the factors which should be taken into account in the design of an auditorium and explain how each factor is controlled.

8. List the defects which should be avoided in the designing of an auditorium.

9. By means of a drawing, illustrate how a well-designed musical studio should be built.

Suggested Readings

(1) Barrows, S. T., and Cordts, A. D., *The Teacher's Book of Phonetics*, Ginn and Company, Boston, 1926.

(2) Bartholomew, "Physical Definition of 'Good Voice-Quality' in the Male Voice," *Journal of the Acoustical Society of America*, Vol. VI, July, 1934, pp. 25–33.

(3) Fletcher, Harvey, *Speech and Hearing*, D. Van Nostrand and Company, Inc., New York, 1929, Parts I and II.

(4) Knudsen, V. O., *Architectural Acoustics*, John Wiley and Sons, Inc., New York, 1932, Parts II and III.

(5) Miller, D. C., *The Science of Musical Sounds*, The Macmillan Company, New York, 1916, Lectures VI, VII, and VIII.

(6) Russell, G. O., *Speech and Voice*, The Macmillan Company, New York, 1931, Parts I, II, and III.

(7) Sabine, W. C., *Collected Papers on Acoustics*, Harvard University Press, Cambridge, Mass., 1922, pp. 3–274.

(8) Watson, Floyd R., *Acoustics of Buildings*, John Wiley and Sons, Inc., New York, 1923, Parts II and III.

(9) ——, *Sound*, John Wiley and Sons, Inc., New York, 1935, Chaps. XIV–XVI.

(10) Wolf, *et al.*, "Quantitative Studies on the Singing Voice," *Journal of the Acoustical Society of America*, Vol. VI, April, 1935, pp. 255–266.



CHAPTER XII

Energy on Light Waves

How do we see things? If we close our eyes, we shut out the sight; if we are blind, we cannot see at all. It must mean that the eye is a specialized sense organ of sight. So far, we have discovered that the senses of warmth and cold can be used in a very rough way to measure the mean kinetic energy of molecules, as the stimuli of molecular movements are interpreted in terms of degrees of "hotness" or "coldness." We have also discovered that the sense of hearing may be used to analyze the complex vibrations of air molecules in sound waves, as such stimuli are interpreted in terms of loudness, pitch, and tone quality. What stimulates the eye? Light.

What Is Light?

Plato (427–347 B. C.) and his followers claimed that three elements are necessary in vision: First, a visual stream of light or divine fire emitted by the eye itself; second, the union of visual light and the light of the sun; third, this union meeting a third emanation from the object itself. Pythagoras (580?–500? B. C.) and his followers taught that vision was caused by particles continually projected from the surface of objects into the pupil of the eye. Aristotle maintained that light is not a material emission but a mere quality or action of a medium called the *pellucid.* Out of these beginnings developed two theories as to the nature of light. Pythagoras pointed the way for the corpuscular theory, and Aristotle seemed to anticipate the wave theory.

During the seventeenth century, these two theories were in great conflict, the "corpuscles" being championed by Newton and the "waves" being defended by Huygens (1629–1695). The eighteenth century, the century of materialism, saw the "corpuscles" in the ascendancy, but at the dawn of the nineteenth, the "waves" were gaining ground under the able leadership of Young (1773–1829) and Fresnel (1788–1827). At the beginning of the twentieth century, the "corpuscles" seemed doomed forever, but then Planck, Einstein, and Compton came forward and "corpuscularized" the "waves." In the field of electricity, Millikan definitely "corpuscularized" the electron; later Davisson found an electron acting like a "wave"; and finally Bohr pointed out that certain experiments evoke the "corpuscle" aspect and others the "wave" aspect, but that it is never possible to have both aspects evoked at the same time.

In 1927, Heisenberg proposed the uncertainty principle, which states that it is impossible to determine at the same time the exact position and velocity of an electron. Either property may be determined accurately, but the measurement involved in getting one property spoils the accurate measurement of the other. We have been wrong, according to this principle, in the simple assumption that we can measure anything as accurately as we please. This follows not because of our failure to be omnipotent, but because of the nature of the universe itself. If we remember that all our scientific concepts emerge from patterns inferentially woven out of data obtained from apparatus, then it is not surprising that our concepts of the elements of the microcosm, which because of their ultimate nature are the first entities to reveal unambiguously the principle proposed by Heisenberg, should take under certain measurements the form of one sort of thing and under other measurements the form of another sort of thing, but never the form of both things at the same time. The scientist has designated these two things as "corpuscle" and "wave," and no contradiction

exists in fact. (For further information on this subject, refer to Chapter XVII.)

The aspects of light presented here may all be interpreted in terms of waves. Therefore in the following pages we shall consider light a form of wave motion.

Source of light. We receive light during the day from the sun and during the night from the moon (reflected from the sun) and the stars. But we also observe that things on the earth may give out light if burned or if heated to incandescence. Thus we conclude that light comes from a source of energy, a place of high temperature and great molecular activity. One may roughly estimate temperature by using this color scale:

First visible red....	525°C.	Dull orange.....	1,100°C.
Dull red.....	700	Bright orange.....	1,200
Turning to cherry .	800	White.....	1,300
Cherry proper.....	900	Brilliant white....	1,400
Bright cherry.....	1,000	Dazzling white....	1,500

Incandescence. A body that emits light because it is maintained at a high temperature is *incandescent.* The light of a flame is due to the presence of incandescent carbon particles. The light from an electric lamp is due to the incandescence of the carbon or tungsten filament.

In such bodies, more and more of the energy radiated becomes visible as the temperature increases. This is why the nitrogen-filled tungsten electric lamp is more efficient than the carbon lamp. The increase in temperature is made possible by adding a gas pressure to keep down the evaporation of the metal at high temperatures.

Luminous vapors. Luminous vapors and gases, produced in flames or in vacuum tubes, have characteristic colors. For example, the sodium vapor produced by burning salt in a flame is yellow; the vapor of lithium is red. Vacuum tubes containing neon, hydrogen, or a mixture of gases are used as electric signs and have characteristic colors.

Luminescence. There are bodies which give off light at temperatures far below that required for incandescence.

They are said to be *luminescent.* For example, crystals of fluorite when held in a heated iron spoon give off *white light* before the spoon shows *any trace of red.*

Fluorescence and phosphorescence. Certain bodies have the ability to absorb radiations, such as invisible ultraviolet light and X-rays, and afterward emit the energy as visible light. If they emit visible light only during the time of absorption, they are said to be *fluorescent;* if they emit visible light also immediately after absorption, they are said to be *phosphorescent.* Calcium sulphide, for example, is phosphorescent. A luminous paint is made by the mixture of zinc sulphide with a trace of radium salt. The zinc sulphide absorbs the invisible radiation from the radium and emits it as visible light. Thus, a steady fluorescence is maintained.

Luminous animals. Glowworms and fireflies produce light by a chemical reaction at low temperatures—an oxidation of some substance produced in their cells. The firefly has a very high luminous efficiency estimated at 99.5 per cent; that is, this per cent of the radiated energy is visible light. On the other hand, a four-watt carbon glow lamp has a luminous efficiency of 0.43 per cent, most of the energy, therefore, being emitted as radiant heat.



Fig. 143. A Candle.

Intensity of a source of light. *Candle power.* A generation or so ago, candles were the main source of artificial illumination (Fig. 143). As newer and better sources of light were discovered and introduced, it was natural, therefore, to think of them in terms of the candle. Today we measure the intensity of a light source and assign it a certain candle power, meaning that it is equivalent to that many standard candles. A candle made of paraffin, melting at 55° centigrade, having a cotton wick of 24 to 25 threads and a flame 5 centimeters high, and burn-

ing 7.7 grams of paraffin per hour, has an intensity as seen in the horizontal direction of 1.11 candles.

Velocity of light. After years of careful and painstaking effort, the late Professor A. A. Michelson, of the University of Chicago, made the most precise determination of the velocity of light ever obtained. The most accurate base line ever measured is that stretching between Mt. Wilson and Mt. San Antonio, California, a distance of approximately 22 miles. This measurement was made by the United States Government in order that Professor Michelson might have a very exact distance through which to reflect light. A beam of light reflected from an eight-sided

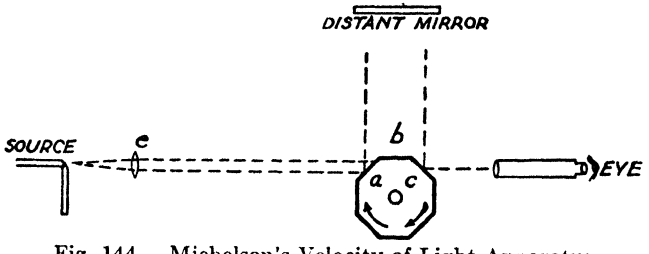


Fig. 144. Michelson's Velocity of Light Apparatus.

mirror spinning at the rate of about 30,000 revolutions per minute was sent from Mt. Wilson to Mt. San Antonio and back, and then observed (Fig. 144). The spinning rate was adjusted so that the mirror turned through an angle of 45° while the light made this round trip. From the distance between stations and the rate of spin of the mirror, the speed of light was obtained from the relation,

$$c = 16nD, \tag{12.1}$$

where c is the velocity of light, n is the number of turns per second of the eight-sided mirror, and D is the distance between the two stations. Such measurements made in air when reduced to a vacuum yield the precise value of 299,796 ± 4 kilometers per second, or 186,285 miles per second.

Light travels a little faster in a vacuum than it does in air. Its speed in water is about three-fourths and its speed in ordinary glass about two-thirds of its speed in a vacuum.

Diffraction

A casual observer is impressed with the ability of water waves and sound waves to bend freely around corners, while light waves seem to travel ahead in straight lines. For example, a person in an adjoining room is heard with ease when the door is left open, but it is impossible to see him unless he is in the observer's field of vision, a region determined by the straight lines touching the edges of the door and entering the observer's eye. Because of this first impression that light travels in straight lines, while commonly observed waves bend around corners, much opposition was lodged against the theory that light is a wave motion. But in 1815 Augustin Fresnel (1788–1827) presented to the Paris Academy his first paper on the diffraction of light, the bending of light around small objects and out from small holes and slits. One observes this phenomenon by looking at a light through nearly closed eyelids, by viewing the moon or a street lamp through a screen door, by looking through a slit and gradually closing it, or by viewing the outline of a wood screw or fine needle against the sky.

The diffraction of light remained long undiscovered because the effect is not detected unless the openings through which the light passes are very narrow, the objects viewed very thin (at least as small as a few hundred wave lengths of light, approximately 0.01 centimeters), or the source of light small enough so that the shadow boundaries are sharp.

Because of diffraction, one may not see anything clearly through a small hole or a narrow slit. When the diameter of the hole or the width of the slit is gradually decreased, the things viewed through the opening become more and more blurred (Fig. 145). Finally, when the opening is small as compared with the wave length of light (0.000058 centimeters for yellow light), the hole or slit gives out light in all directions exactly as if it were a point source. The objects

formerly viewed are completely obliterated. In all optical instruments, including the eye, we view objects through holes, and because of this a certain amount of blurring results. Because of diffraction, therefore, these instruments are limited in their ability to make distinctly visible objects which are very close to each other. This limitation, spoken of as *resolving power*, is discussed further in Chapter XIV.

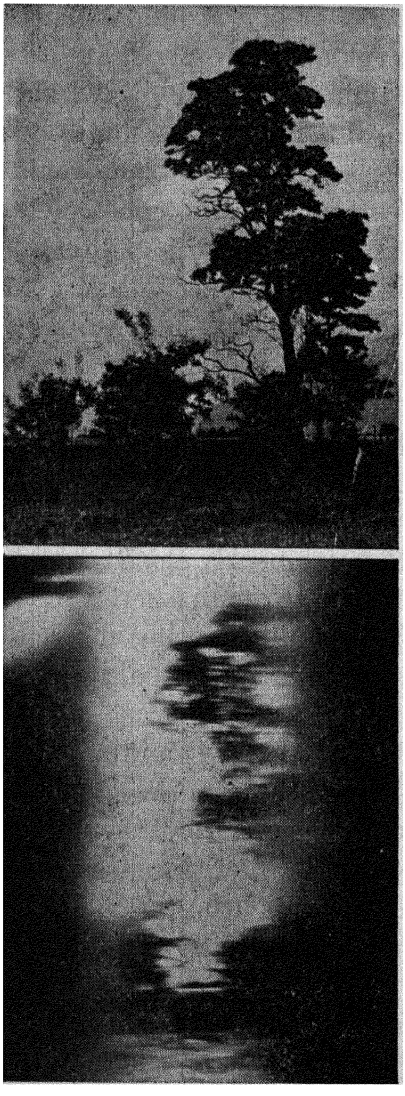


Fig. 145. The Blurring Effect of a Very Narrow Slit. (A slit having a width of 0.016 centimeters was placed in front of the lens of a camera when the lower photograph was taken.)

Since diffraction is a phenomenon which appears in association with openings and objects of very small dimensions, for practical purposes we may consider that light travels in straight lines or rays when we investigate its action in connection with objects of ordinary size.

Reflection

Reflected light. Objects such as trees, clouds, mountains, the moon, and persons are not sources of light. The light just seems to come from them as a source after it is reflected. They are illuminated bodies. The glare of reflected light from a placid lake, from a plate glass window, from a varnished surface, or from the highly glazed paper of a book may be a hindrance

to good vision. What causes the glare from some surfaces, and why are other surfaces so soft and easy to look at?

The law of reflection of light. Let a beam of sunlight be reflected from a mirror. The path of the light is made visible by the brightly illuminated dust particles suspended in the air. Chalk dust or smoke may be added if necessary. When the mirror is perpendicular to the incident beam, the ray is reflected directly back on itself; when the mirror is turned through a 45° angle, the incident and reflected beams are 90° apart. Other positions may be tried. In every case it is found that the *angle of incidence* (measured with reference to the perpendicular to the mirror) is *equal* to the *angle of reflection* (measured with reference to the same line). This is the law of reflection of light, the same law found to be true for the reflection of sound (Fig. 146).

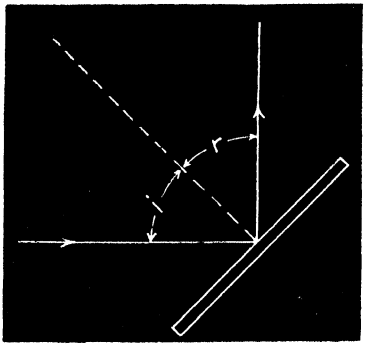


Fig. 146. The Reflection of Light.

Regular reflection. When a surface is perfectly plane, that is, without roughness, a beam of light on reflection keeps its form because all parts are reflected in the same direction. This is called *regular reflection*, and is that produced by a mirror, a smooth water surface, and a surface of plate glass. *A surface that produces perfectly regular reflection cannot be seen.* If a plate glass mirror is carefully cleaned and its edges skillfully camouflaged, its presence may be completely overlooked, with a resulting optical illusion. The regular reflection from the sun, or from any other bright source, produces a glare. The effect is as if one looked directly at the source. Often one wonders why he becomes sunburned while fishing on a lake in spite of the use of a large hat as a sunshade. The reason is now clear.

Diffuse reflection. When one covers a glaring surface with a piece of soft, white cloth, the light is scattered

through the room and the glare is removed. When a varnished table top is rubbed down with pumice-stone and oil, the sheen is greatly reduced and a velvet-like surface is obtained. To further show this scattering of light, let a small amount of whiting or white calcimine be mixed with water and a letter painted with it on a plane mirror. When a spot of sunlight is reflected on a screen, the figure is clearly visible; but a dark letter, instead of the original white letter, is seen. The white letter reflected light, and *each portion of the beam obeyed the law of reflection;* but owing to the uneven nature of the surface, the light was sent out in all directions. Thus there was a scattering or a diffusion of the light, and the amount reflected in the direction of the screen was greatly reduced, causing the letter to appear darker than its surroundings (Fig. 147).

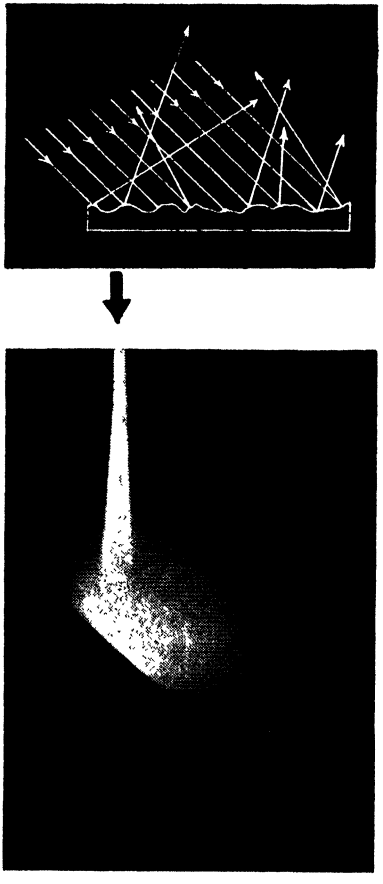


Fig. 147. Diffuse Reflection. (*Courtesy of General Electric Company.*)

When a surface reflects diffusely all the light falling upon it, all visible aspects of the source are obliterated; in the process of seeing, the surface *appears* to be a source of light. A calcimined wall, a wall decorated with flat paint, clouds, and the sky are fine examples of surfaces of this kind. One has but to look through a window to realize that most of the light of the room comes from the sky, and to look about the room to discover that a good deal of this light is again reflected from the walls. It is perhaps correct to say that one does not see light, but simply the objects that emit or diffusely reflect it. Thus, one sees the

sun when looking along the beam of sunlight. Looking across it, one sees dust particles, trees, houses, mountains, clouds, and the sky, all of which receive light from the sun and reflect it diffusely.

The Sky

Very small particles, such as air molecules and very light dust particles suspended in the air, scatter light according to its color (wave length), the violet and blue being scattered more than the red and orange. If it were not for this scattering, there would be no blue sky, no red sunset, no twilight or dawn. The sky would be black, and the stars could be seen both night and day. We would be compelled to place large white screens near our homes to reflect into them the light which now comes from the sky. Since the blue light is scattered more than the red, the sun and stars look a little more red than they actually are. Just before sunset, the sunlight passes through a greater depth of air than during the middle of the day, and this means that more of the blue light will be scattered. *Looking in the direction of the sun*, one will see red light either in the sky or reflected from clouds. An increased dustiness adds to the coloring. Who has not seen the sun set as a red ball of fire at the end of a dusty day or looked with awe at its red glow as it dropped behind the smoke of a forest fire? On the other hand, when we look in directions nearly at right angles to the sun's rays, we see the scattered blue light which gives the sky its characteristic color and distant mountains a purple hue. If the air is especially free from dust and water vapor, the sky takes on a deep blue or violet color. The coloring of different sunsets, then, is due to differences in the quantity and size of dust particles present, and the varying hues of a certain sunset are due largely to the changing angle at which the sun's rays are viewed.

Illumination

Illumination and distance from the source. Experience has taught all of us that we may place a surface under

increased illumination by bringing it nearer to a light. If the light from a candle illuminates, first, the inside of a spherical shell of one-foot radius, and then a similar shell of two-foot radius, the same light energy is spread over different areas (Fig. 148). The smaller shell has one-fourth the surface of the larger. Therefore, each square foot of the smaller surface receives four times as much energy, and so four times as much illumination, as a square foot of the larger surface. The surface of the small shell would be

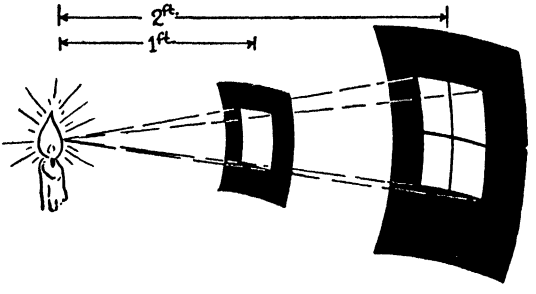


Fig. 148. Illumination Varying Inversely as the Square of the Distance from the Source.

under nine times as much illumination as that of a shell with a three-foot radius, and sixteen times as much illumination as that of a shell with a four-foot radius. Thus, *the illumination produced by a source radiating light equally well in all directions varies inversely as the square of the distance from the source.*

The foot-candle. Illumination is measured in foot-candles. Let us imagine a standard candle sending out light in all directions, none of it being concentrated by a reflector. A foot-candle represents an amount of illumination equal to that produced by this standard on a surface placed perpendicular to the rays of light and located one foot away (Fig. 149). The same illumination may be obtained in a number of ways. For example, surfaces placed *one* foot from *one*

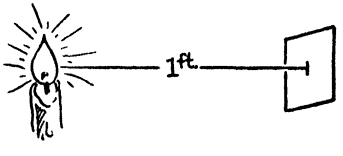


Fig. 149. The Foot-Candle.

candle, *two* feet from *four* candles, and *three* feet from *nine* candles are under the *same* illumination.

Thus, the illumination in foot-candles produced at a given distance by a light source with intensity measured in candle power may be calculated by the simple relation,

$$\text{Foot-candles} = \frac{\text{Candle power}}{(\text{Distance})^2}$$

For example, the illumination 2 feet from a 16 candle-power lamp is 4 foot-candles. A shade may increase this illumination if by reflection it keeps the light from going out in all directions and in doing so does not absorb too much of it.

Brightness of a surface. A gray surface under the illumination of one foot-candle will not appear as bright as a white surface under the same illumination. The gray surface does not reflect as much of the light falling on it as does the white one, part of the light being absorbed. Hence the brightness of an object depends, not only upon the illumination under which it is placed, but also upon the fractional part of the light reflected from it.

Since diffusely reflected light is used in seeing objects, increased illumination should be used when a person is working with dark materials. The ratio of the foot-candles reflected to the foot-candles applied is here shown for a number of colored surfaces:

	Per Cent		*Per Cent*
White	80	Sky Blue	35
Ivory	70	Olive Green	20
Buff	65	Cardinal Red	20
Sage Green	40	Black	5

Foot-candles required for proper lighting. The illumination required for proper lighting depends upon the degree of accuracy needed in an operation, the fineness of detail to be observed, the color of the materials used, and the attractiveness desired. Various approximate illuminations are attained as follows: 10,000 foot-candles in the sunlight on a tennis court, 1,000 foot-candles in the shade

of a tree, 100 foot-candles indoors near a window, 10 foot-candles in a store well-lighted artificially, and 1 foot-candle in a shop poorly lighted by artificial means. In a school building, the illumination measured in foot-candles should be much as follows:

Auditorium	8
Classrooms, Library, and Office	12
Corridors and Stairways	5
Drawing and Drafting Rooms	25
Laboratories	12
Manual Training Shops	12
Sewing Rooms	25
Study Rooms—Desks and Blackboards	12

Fig. 150 shows a foot-candle meter which measures the illumination in foot-candles at any place by the action of a photoelectric cell (explained in Chapter XVI).

Glare. Any brightness within the field of vision which causes discomfort, annoyance, eye fatigue, or in any way interferes with good vision is called *glare*. The interference may be caused by a direct glare from a naked incandescent electric light filament or by a reflected brightness from a varnished wall or desk top. The extent of the glare depends in part upon the contrast in brightness between the light source and the background. The headlights of an automobile are blinding on a dark night but are scarcely noticed if left burning during the day.

Direct glare from a bare lamp is one of the greatest hazards of vision (Fig. 151). Lamps should be properly shaded. If an electric lamp must be used without a shade, it should by all means be of the frosted type. To be free from glare, the intensity of a lighting unit, hung so as to be in the central portion of the field of vision, should not exceed 2 or 3 candle power per square inch of apparent area. A 40-watt electric light with frosted globe has an intensity per square inch of about 5 candle power (candle power 30 and effective area 6 square inches), a 100-watt light of this type an intensity per square inch of 11 candle power, and a 100-watt clear globe an intensity per square inch of at least 70 candle power.

Shades or diffusing globes have the advantage of either shielding the light source from the field of view or giving it a larger equivalent surface. If a 100-watt clear lamp

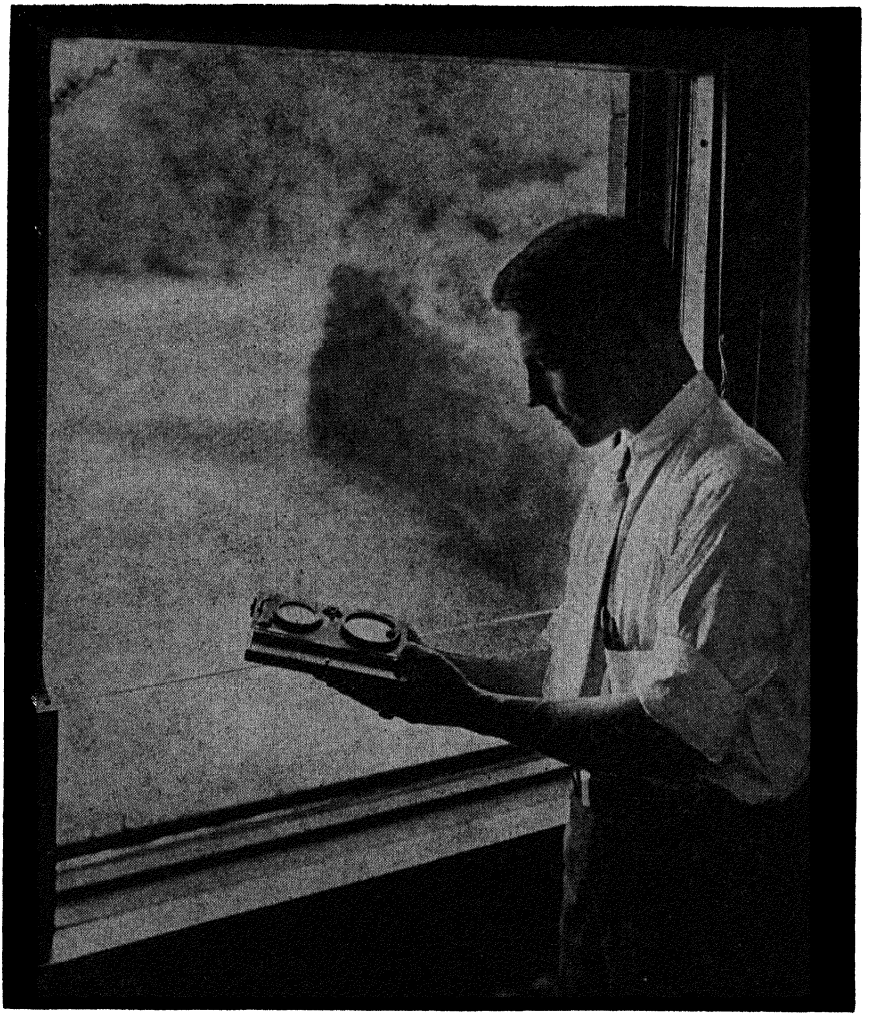


Fig. 150. A Foot-Candle Meter. (*Courtesy of General Electric Company.*)

were placed in a diffusing globe of 10-inch diameter, its intensity per square inch would drop from 70 to 1 candle power. As higher candle-power lamps are adopted, indirect or semi-indirect methods of lighting should be used. By such means, part of the light may be reflected over the

ceiling area and the brightness may be kept within the limits of comfort. The dangers attendant upon artificial

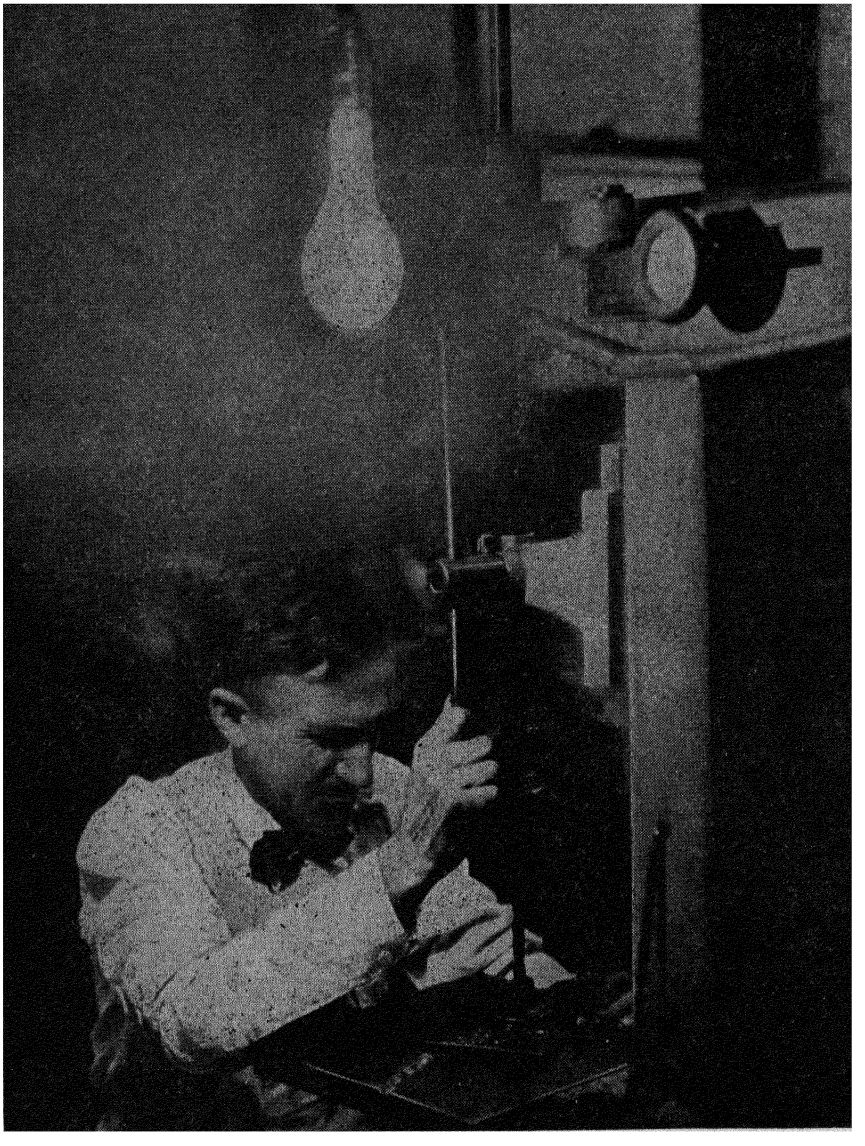


Fig. 151. Glare. (*Courtesy of General Electric Company.*)

illumination, then, are not due to the supply of too much light, but rather to the installation in the field of vision of small areas of intense brightness.

Reflected glare. Reflected glare, which comes from some light source to the eyes by way of regular reflection, is often more harmful than direct glare. It comes often from a direction below the horizontal, a zone in which the eye has no natural protection; because of this, its insidiousness may not be realized. Glossy paper, polished metal or wood, and glass-top desks may thus contribute to eye fatigue which in the end may lead to permanent injury.

Shadows. Soft shadows are helpful in distinguishing form and detail; a proper blending of light and shade is of great importance in photography. When shadows are sharply defined and black, they are harsh and unpleasant and may even be dangerous in shops where they hide moving machinery.

Sharp and black shadows are produced by small intense light sources, such as exposed electric light filaments and arcs. You may have noticed this as you have passed under certain street lighting fixtures. Large light sources, such as diffusing globes, do not create sharp shadows. You may easily discover the reason for this by performing the following simple experiment:

Close one eye and gradually move the hand across the face so as to cut out the sight of a large globe. There is first a position where the globe just starts to be covered, then a region where a smaller and smaller part of the globe is seen. The intensity of the light gets less and less until finally the last trace of the globe is gone and a complete shadow covers the eye. Thus, a shadow produced by a large globe has a hazy edge. If, however, the hand is passed across the face in front of an electric lamp with naked filament, the distance which the hand may move with the filament partially visible is short, and hence the shadow is sharp. Now let the hand be moved as far as possible from the face. If the globe is large enough, or if you use a single finger as a screen, you will find it impossible to hide the source completely. This means that at this distance no complete shadow of the hand

or finger is formed. Thus, a large globe makes hazier-edged shadows than does a small light source, and the same procedure that eliminates direct glare also helps to make soft shadows. Conversely, sharp shadows are indicative of a glare which may be annoying and dangerous (Fig. 151).

Refraction

When light strikes a sheet of glazed white paper, part of it is regularly reflected, part is diffusely reflected, part is absorbed, and a very little is transmitted. If light strikes a piece of glass, part of it is reflected when it enters, another part when it leaves, some little is absorbed in the glass, and the remaining portion is transmitted on through the air. At the surfaces of separation of air and glass, light is not only reflected, but the transmitted portion is bent or *refracted.* This may be demonstrated as follows:

Let a ray of light pass obliquely through a parallel plate of glass. At the air-glass surface, the beam is bent. It continues in a straight line in the glass; but when it reaches the glass-air surface, it is bent again. The first bending is toward a line drawn through the glass perpendicular to its surface; the second bending is away from this line (Fig. 152). The parallel plate did not make a change in the general direction of the beam, but simply shifted it to one side. This action may be illustrated vividly by placing a pencil back of a piece of thick plate glass and observing the lower portion of it through the glass and the upper portion directly. If viewed from one side, the displacement of the light coming through the glass makes the pencil seem completely severed.

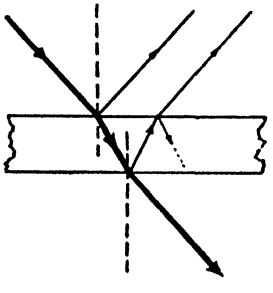


Fig. 152. Refraction.

If light is sent through a prism (Fig. 153), it bends toward the perpendicular *N* drawn in the glass and away from the perpendicular *M* drawn in the air. This bending toward

the base of the prism means a definite change in direction. Thus, refraction as well as reflection may be used to change the direction of light.

Next imagine two prisms placed base to base. The upper one will cause light to bend downward and the lower one will cause it to bend upward. It is easy to imagine the prism-pair smoothed down to the characteristic shape of the lens shown in the figure. When direct sunlight is sent through this double convex lens, it is brought to a focus, and a sufficiently high temperature to burn paper may be produced. Thus we have discovered an easy means of concentrating light.

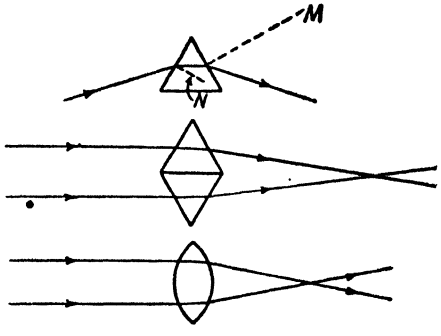


Fig. 153. Refraction Through a Prism and a Lens.

What causes the bending at the air-glass and glass-air surfaces? The refraction may be explained by the use of the fact that the speed of light is less in glass than in air. As a ray of light passes obliquely through the air-glass surface, the lower edge meets the surface before the upper, and as it passes into the glass, its speed is slowed up. The upper edge thus takes the lead, and by the time it reaches the glass, the ray will have swung around a little toward the slower edge of the beam—toward the perpendicular in the glass (Fig. 154). When the beam reaches the glass-air surface, one of the edges will in general emerge first, which one depending on the shape of the glass and the direction of the ray. The edge of the beam that gets out first will speed up and swing around toward the slower edge (the edge that leaves the surface last) and away from the perpendicular to the surface erected in the air.

Refraction is also found to be present at an air-water surface. This explains why the depth of a tall jar of water seems shortened as one looks at the bottom through the

water, and why an oar seems broken at the water surface (Fig. 154).

Finally, we conclude that refraction takes place at the *surface of separation of any two media* in which the *speed of light is different.* If at the point of *entrance* a perpendicular to the surface is drawn in the medium *being entered,* then the bending will be *toward* this line if the speed is *slowed up* and away from it if the speed is *increased.*

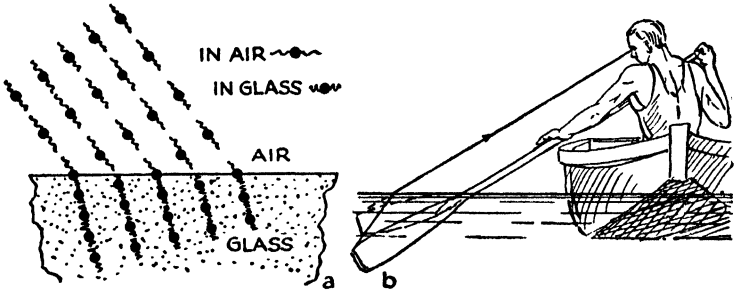


Fig. 154. (*a*) Illustrating the Cause of Refraction. (*b*) The "Broken" Oar.

In Chapter XIV we shall explain how reflection and refraction are used in optical instruments.

Questions and Problems

1. List five sources of light.

2. A surface that produces perfectly .... cannot be seen.

3. Illuminated objects are seen by .... light.

4. The intensity of a light source is measured in ....; the illumination under which an object is placed is measured in .....

5. What is the illumination in foot-candles at a surface 4 feet from a 32 candle-power lamp?

6. ..... shadows are produced by .... light sources; ..... shadows are produced by ..... light sources.

7. Refraction takes place at the surface of separation of any two media in which the ..... of light is .....

Suggested Readings

(1) Bragg, W. H., *The Universe of Light,* The Macmillan Company, New York, 1933.

(2) Dietz, David, *The Story of Science*, Sears Publishing Company, Inc., New York, 1931, Chaps. I, II, and III.
(3) Harris, F. S., and Butt, N. I., *Scientific Research and Human Welfare*, The Macmillan Company, New York, 1924, Part IV.
(4) Keene, F., *Mechanics of the Household*, McGraw-Hill Book Company, Inc., New York, 1918, pp. 274–288.
(5) Lenard, P., *Great Men of Science*, The Macmillan Company, New York, 1933, pp. 67–83 and 129–213.
(6) Luckiesh, Matthew, *Lighting the Home*, D. Appleton-Century Company, Inc., New York, 1920.



CHAPTER XIII

Color

Colors in Nature

The rainbow. The rainbow with its colorful hues is always a source of admiration. It often appears in the eastern sky in the late afternoon during a rainstorm and may appear in the western sky during the early morning. It is seen in a waterfall, in the spray produced by a garden hose, or whenever sunlight plays at the proper angle upon falling water droplets. This fine display of coloring is evidence that sunlight when returned from water droplets is separated into its colors. These data of common experience may be explained and interpreted as follows:

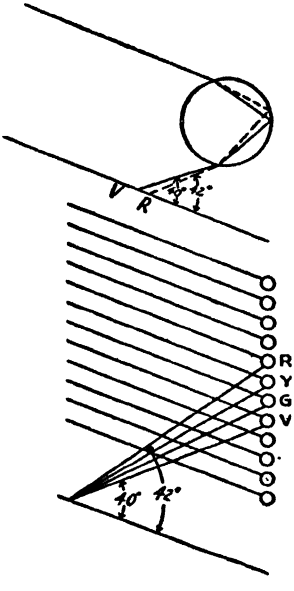


Fig. 155. Sunlight and Raindrops Producing a Rainbow.

As a ray of sunlight enters a water drop, it may be reflected at least once inside the drop and refracted as it enters and leaves. If the violet component of sunlight is refracted more at each surface than is the red component, this might account for the fact that the violet color of the rainbow appears at an angle of 40° and the red color at an angle of 42° with a line parallel to the sun's rays. In Fig. 155, the path of a ray of light through a drop is traced, the violet component being shown bent more than the red. The angle which each emergent beam makes with the direction of sunlight is noted. The violet component, which we have assumed is bent the most, makes

the smallest angle with the direction of the direct rays of the sun. Therefore, our assumption is justified.

When an observer is standing with his back to the sun, light coming from drops of water located in the sky, in a waterfall, or in a spray and making an angle of 42° with the rays of direct sunlight will appear red. The light from drops which make an angle of 40° with the direct rays of sunlight will appear violet. These angles may be obtained by looking up, to the right or to the left, and in some rare cases by looking down. Thus, the rainbow is always a part of a circle. Each falling drop contributes to the red color seen by a particular person when it happens to be at the correct angle of observation. As a drop moves down and away from the proper angle, others take its place, and the color is still seen in the definite direction, a great host of drops contributing to the picture. It is certainly true, therefore, that each person's rainbow is his very own.

When light is reflected a second time within the drop, the violet color, being refracted most, will be seen on the outer side, and a so-called secondary bow, which is often observed surrounding the regular rainbow, is formed.

Colors of thin film. Who has not admired the spectral colors of thin soap films, oil films on water, and the iridescence of mother-of-pearl? We shall seek the cause of this color.

Let two pieces of plate glass be placed together so that a very thin air wedge is formed between the two adjoining surfaces. When the light from a sodium flame is reflected from these surfaces, a series of yellow and black bands will be seen. The explanation is as follows: The light from the flame is partly reflected from the glass-air surface and partly from the air-glass surface of the *wedge*. The train of light waves reflected from the second surface combine with the train reflected from the first. If it happens that the second train exactly matches the first by being an exact number of wave lengths behind, the light will be re-enforced. But

if the matching is not perfect, and especially if the waves are in opposite phase, destructive interference will result and the intensity will be weakened or completely destroyed. Thus, at the position of re-enforcement, a bright band is seen; at the position of destructive interference, a dark band appears. The result is a set of light and dark bands known as *interference fringes* (Fig. 156).

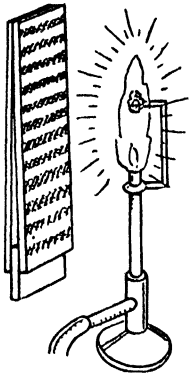


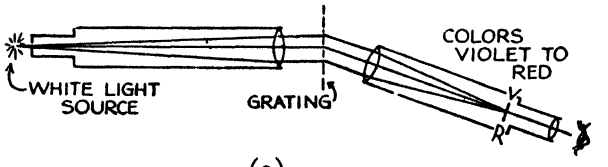
Fig. 156. Interference Fringes.

The difference in the thickness of the air wedge corresponding to one dark band and that of an adjacent similar band must be one-half wave length, because the light traveling *across the air wedge and back* at one band must go a wave length farther than the light at the adjacent band. Were this not so, destructive interference could not occur at both positions. By counting the number of dark bands in a given length and measuring the difference in wave path involved, it is possible to calculate the wave length of the light studied.

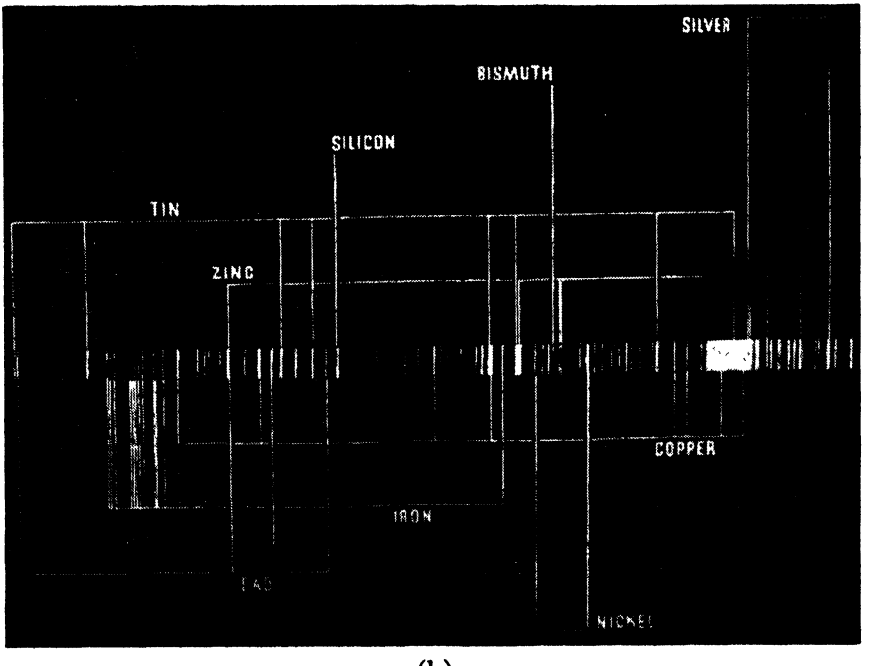
With the same air wedge the fringes of the interference pattern formed with red light are farther apart, and with violet light closer together, than those of the pattern produced by the yellow light from a sodium flame. This means that red light has a longer wave length than yellow, and yellow light a longer wave length than violet. The Michelson interferometer, an instrument making use of interference to measure wave length, permits the accurate measurement of the wave length of light of any color. Such measurements are listed below; but since there is a

Color	*Wave Length in Centimeters*	*Frequency in Trillions of Cycles per Second*
Just visible red	0.000076	390
Red	0.000068	440
Orange	0.000065	460
Yellow	0.000058	520
Green	0.000052	580
Blue	0.000046	650
Violet	0.000042	720
Last visible violet	0.000038	780

gradual blending of colors, it must be clearly kept in mind that the values given are for a certain red, a certain orange, and so on. The frequency is obtained by dividing the speed of light by the wave length (see Equation 9.1).



(a)



(b)

Fig. 157. (*a*) Diffraction Grating. (*b*) Spectrum of Bronze with Important Lines of Various Constituents Identified. (*Courtesy of Bausch and Lomb Optical Company.*)

Under the action of gravity, soap bubbles become thinner at the top than at the bottom, and the very thin wedge thus formed should give rise to interference phenomena. When the wedge is thick, as when the soap bubble is just formed, the bright bands of all the interference patterns produced by the various color components overlap and white light is still observed. However, as the wedge gets thinner

and thinner, the bright bands of each color get farther and farther apart until a bright band of orange, for example, falls upon a dark band of red, and all the colors of the spectrum are seen vividly displayed. Thus white light is separated into its many colors by the process of light interference.

If as many as 20,000 lines to the inch are ruled on a piece of glass, each little glass space between will act as a very narrow slit, so narrow as to send light in all directions because of its diffraction. When the light waves from these tiny sources cross, interference results and white light is separated into its many component parts. Such a so-called *diffraction grating* is used to study the color composition of light (Fig. 157). It will be shown in Chapter XVII that light has its origin inside the atom, and that the frequency (color) distribution emitted gives information concerning the unique nature of the atom itself, not, of course, by direct vision, but by scientific inference. The light from the sun and stars has been analyzed and compared with that produced in the laboratory, and when the "messages" have been deciphered, we find that the same elements which compose the earth also make up the sun and stars.

Natural colors due to absorption and reflection. Under daylight, one sees a green color whether he looks at a leaf or through it. In the first case, reflected light reaches the eye; in the second case, transmitted light enters it. When a landscape is viewed through a red glass filter, the deep blue sky and the green trees look black and the white clouds look red. On closer observation, one sees shifting spots of red light as the leaves flutter in the breeze.

These simple observations may be interpreted as follows: A portion of the sunlight striking the leaves is reflected with little or no penetration of the surface, and comes off as white light, appearing through the filter as a red spot. Another portion penetrates the leaf to small depths, but is finally returned as reflected light after all the components of white light except green have been selectively absorbed. Such

light appears green when viewed directly and black when viewed through a red filter. The last portion is transmitted and emerges with all the components except green absorbed, thus adding to the green appearance of the leaf. This means that a leaf is green whether viewed with reflected or transmitted light and that its color in both cases is due to *selective absorption.* The green leaves and the deep blue sky appear black through a red filter because green and blue light are absorbed by the glass—only red may pass through this filter.

Finally, an object appears white, or more properly, a certain gray, because all the components of white light are returned to the eye equally well; an object appears black because all light which strikes it is absorbed; and the various colored objects under daylight get their color by the process of selective absorption, that is, by absorbing certain components of white light and not others. Many colored objects have a glossy appearance. This is because a portion of the white light falling on the object is reflected from the smooth surface without penetrating deep enough into the surface to have its components selectively absorbed.

A striking display of what happens to objects of various colors when illuminated by light of a pure color may be shown by darkening a room and placing various colored objects in the presence of a sodium flame. All objects appear yellow or black. Those which naturally reflect yellow light or varying quantities of white light appear yellow. Those which reflect neither yellow nor white must appear black. When an electric light is turned on, a spectacular change in color is observed.

When a red glass filter is placed in a beam of sunlight, the transmitted light is red, but the reflected light is white (a very slight red tint will appear owing in the main to the light returned from the opposite surface after having traversed the glass). Thus, the color of red glass is due entirely to selective absorption (Fig. 158). But when light strikes a piece of gold foil, yellow light and red light are

reflected, blue is absorbed, and green is transmitted. Thus, as one looks at such a thin film, he sees yellow; as he looks through it, he sees green. This means that the reflected light is not yellow because of absorption but because of a *selective reflection* which takes place at the very surface (Fig. 158). Such selective reflection is also observed if a large, dry spot of analine red ink is viewed by both reflected and transmitted light. The reflected light is green while the transmitted light is red. The red color of the ink, therefore, is due not to selective reflection at the surface, but rather to selective absorption as white light penetrates through the ink probably as far as the paper before being returned as diffusely reflected red light. A "metallic" color due to selective reflection is seen in certain brilliantly colored insects. Other insects, and feathers, are colorful as a result of the interference action in thin films explained above.

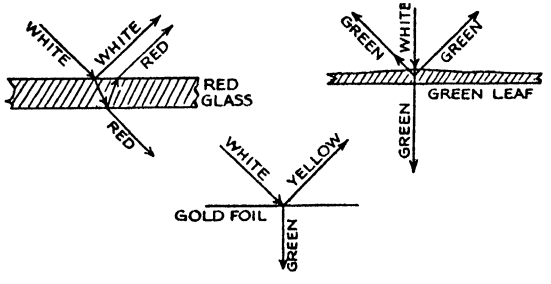


Fig. 158. Selective Absorption and Selective Reflection.

Dispersion

We were able to interpret the formation of the rainbow by assuming that sunlight is separated into its colors as the beam bends on passing in and out of a raindrop. Newton was the first to make a scientific study of this action, known as *dispersion*. He substituted a glass prism for water droplets and let a beam of sunlight pass through a narrow slit and fall upon the prism. In such an experiment (Fig. 159) the beam is bent toward the

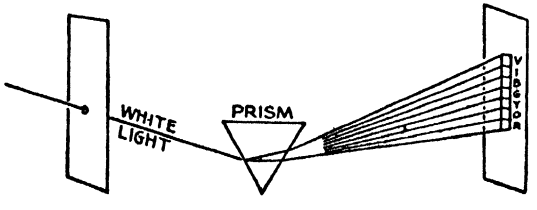


Fig. 159. Dispersion by a Prism. (*See Plate I.*)

base of the prism and on emerging is found separated into the colors of the rainbow, the violet being bent most and the red least. The general colors observed, listed in the order of increasing refraction, are red, orange, yellow, green, blue, and violet. Each color merges into the next, and the band of color, consisting of many more than six distinguishable hues, is called a *spectrum*. We have already listed the wave lengths of various colors, and it is obvious from this information that light of the shortest wave length is bent or refracted most on passing through a prism. Making use of dispersion, we have a rather easy means of obtaining light waves of many frequencies. We simply select the portion of the spectrum we wish. To determine the color sensation or sense datum produced by light waves of a single frequency, we look at a very narrow portion of the spectrum. If a mixture of frequencies is desired, the light from various narrow portions of the spectrum may be reflected on a white screen.

Mixing Colored Lights

Under no illumination, a white screen appears black; but when white light of increasing brightness falls upon it, the illuminated spot, which first appears nearly black, moves up in brightness through a series of grays toward white. Thus, under the stimulus of *white* light of varying degrees of brightness, the eye senses a long series of grays between the darkest black and the whitest white.

In the study of the spectrum of sunlight (white light), we observed that certain patches of color are associated with light waves of a definite frequency range. For this reason a color is often designated by a certain frequency of vibration or wave length. Conversely, light of a certain frequency is often given the name of a color. With this manner of speaking, therefore, the word *yellow* at one time means that particular stimulus (light) from a narrow portion of the solar spectrum which gives rise to the sense impression yellow. At another time, it means

the sense datum resulting from this particular stimulus or from certain other sensory experiences. Because of this freedom of usage, the stimulus may easily be confused with the sense datum. Whenever ambiguity might arise, the author will speak of the *stimulus* as a *certain colored light.*

We shall now be interested in mixing colored lights and noting just what color sense data result from the various mixtures of stimuli.

Complementary colors. Let the arrangement of apparatus shown in Fig. 159 be modified by a lens of large diameter being placed in the path of the refracted beam at a position between the prism and the screen, and let the lens be adjusted back and forth till a pure white (gray) spot appears on the screen (Fig. 160). Here we have definite evidence that a proper mixture of the spectral colors (vibrations) produces white light and a sense datum gray.

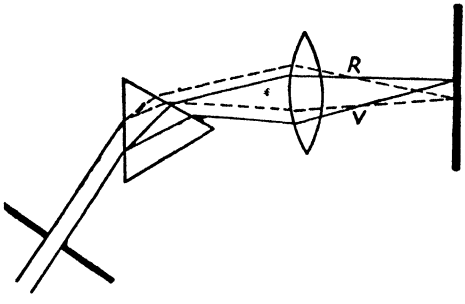


Fig. 160. Recombining Spectrum Colors to Form White Light.

When a card is moved back and forth between the lens and the screen, a vivid spectrum of colors is seen at some position RV. When just the edge of the card is slipped into the beam at R so that only the red portion of the spectrum is cut off, the spot which appeared gray now appears bluish green. This compound color (bluish green) which is observed after red light is removed from white light is called the *complement* of red. When the card is placed in the beam at V so that the violet portion of the spectrum is removed, a greenish yellow tint is observed. Thus, violet and greenish yellow are complementary. When a thin pencil is held in the central part of the beam and the green and yellow portions of the spectrum are cut off, the remaining red, blue, and violet components unite, and a brilliant purple is observed.

You will find the following experiment interesting and fascinating. Fix your gaze upon a bright mass of red color for twenty or thirty seconds. Then look off at a white or gray background and hold the gaze there for a few seconds. You will see a spot of bluish green—a color which is the complement of red. It is supposed that the nerves of the eye which respond to red light become fatigued. Later, when white light comes from the background, the eye is unable for a time to respond as readily to red as to the other colors. Hence, the sensation is that which would be produced if red light were subtracted from white light. Thus the color seen is the complement of red. When colors are investigated by this method, the following are found complementary:

Red	Orange	Yellow	Violet
Bluish Green	Greenish Blue	Blue	Greenish Yellow

We must constantly bear in mind that persons with so-called normal color vision differ in the names they apply to different hues. For example, one person calls blue that which another calls greenish blue; one person calls yellow that which another calls orange-yellow; and one person calls violet that which another calls blue. It is probably safer to say that a certain red is complementary to a certain bluish green, that a certain orange is complementary to a certain greenish blue, and so on.

By recombining the colors of the spectrum in their original proportions, we have found that such a mixture of vibrations is perceived as gray. In the study of complementary colors, we have found that when white light is divided into two portions, each portion serves as the stimulus for a complementary color. When produced in this manner, a complementary color results from a stimulus made up of a mixture of frequencies. Is it possible to produce the same color sense datum with a stimulus composed of a single frequency? If so, it will be possible to arouse the sense datum *gray* by the use of two properly

chosen light vibrations, two single-frequency (monochromatic or pure) colored lights; and the whole gamut of vibrations found in the white light spectrum will not be needed. This is found to be the case. For example, a certain red stimulus characterized by the wave length 0.0000654 centimeters and a certain bluish green stimulus characterized by the wave length 0.0000492 centimeters will produce the same sense datum (gray) as does the whole array of frequencies found in the solar spectrum. The wave lengths of other characteristic pairs are as follows:

$$\begin{matrix} 0.0000607 \text{ cm.} \\ 0.0000489 \text{ cm.} \end{matrix} \qquad \begin{matrix} 0.0000585 \text{ cm.} \\ 0.0000485 \text{ cm.} \end{matrix} \qquad \begin{matrix} 0.0000564 \text{ cm.} \\ 0.0000462 \text{ cm.} \end{matrix}$$

Finally, we conclude that two *colors* (sense data) which fuse to form *gray* are *complementary*. Also, two *colored lights*, whether pure or of the mixed-frequency kind, are also complementary if their mixture appears *gray*, that is, if the mixture of stimuli arouses gray as a sense datum.

Mixing red, yellow, green, and blue light. We shall now seek the hues (sense data) produced by the mixture of pure red, yellow, green, and blue light in various proportions. By an arrangement of small mirrors, any color or combination of colors found in the spectrum may be reflected and mixed upon a white screen, or this mixing may be accomplished by the use of a set of projection lanterns and color filters. With sufficient equipment, elaborate color combinations may be obtained.

(*a*) *Adjacent combinations.* When *yellow* light is mixed on a screen with *red* light in increasing proportions, *reddish orange, orange,* and *yellowish orange* are observed. This means that the stimulus of red light when combined with the stimulus of yellow light in proper proportions is interpreted by the eye as the equivalent of the stimulus of orange light. In terms of vibrations, this means that a slower vibration mixed with a faster vibration produces a stimulus equivalent to that of an intermediate vibration.

In a similar manner, when *green* is added to *yellow* light in increasing proportions, *greenish yellow* and *yellowish green* colors are perceived; when *blue* is added to *green* light in increasing proportions, *bluish green* and *greenish blue* are observed; and finally, when *red* is added to *blue* light in increasing amounts, *violet*, *purple*, and *purplish red* or *crimson* are seen. Purple is not found as a spectrum color; hence, by adding the low frequency of red with the high frequency of blue, an effect is produced on the eye which *the light of a single frequency cannot achieve.* From these results we conclude that when the eye is stimulated by the proper dual color blends of adjacent colored lights in the series: red, yellow, green, and blue, all the hues of the spectrum, with purple in addition, are seen.

(*b*) *Alternate combinations.* Of the possible dual combinations of red, yellow, green, and blue lights, we have remaining the alternate combinations: red with green and yellow with blue. When *red* and *green* lights are mixed, *yellow* is seen; when *yellow* and *blue* lights are combined, *gray* is perceived. These facts may be shown symbolically as follows:

$$R + G = Y,$$
$$Y + B = R + G + B = W;$$

and from them we conclude that a mixture of red, green, and blue stimuli in proper proportions gives rise to the sense datum gray.

As we have already shown, dual mixtures of the stimuli of adjacent colors in the series: red, yellow (red + green), green, and blue give all the hues of the spectrum with purple in addition. But a mixture of red and green stimuli gives a sensation of yellow, and thus a yellow stimulus is not needed if red and green stimuli are available. This means that all color sense data may be aroused by the proper mixture of *three fundamental colored lights*, a certain red, a certain green, and a certain blue (some would call it a blue-violet), which when mixed in proper proportions

will arouse, as sense data, gray, any one of the many hues of the spectrum, or even colors not found in the spectrum. A vivid demonstration of these facts may be shown by the mixture of these three colored lights in varying proportions on a white screen in a darkened room (Fig. 161). Projection lanterns provided with proper light filters and dimming resistances in the electric circuits will serve as excellent light sources.

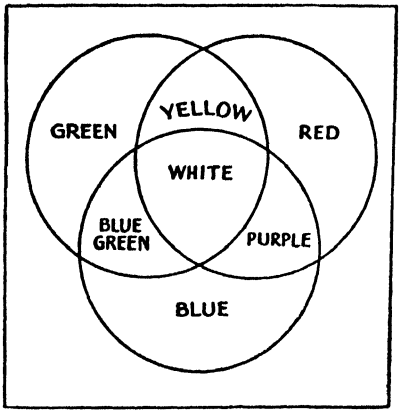


Fig. 161. Mixing Colored Lights.

Newton's color wheel. By cutting color discs along their radii and uniting them by turning a portion of one under another, it is possible to make their visible portions form the sectors of a disc. A rapid rotation of a strongly illuminated disc of this type produces the same effect on the eye as do colored lights projected on a screen. Thus the results of the experiment outlined above may be duplicated on the color wheel. With a proper selection of sector colors and sizes, gray and a very large number of hues may be made to appear as the wheel reaches a high speed (Fig. 162).

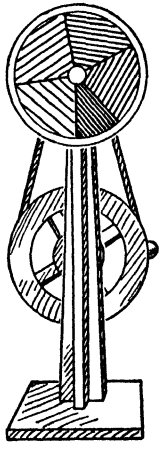


Fig. 162. Newton's Color Wheel.

The law of mixing colored lights. We may summarize the results of mixing colored lights as follows: (1) Every colored light has its complement, and if these two colors are mixed in proper proportions, gray is seen. If the proportions are not correct, the hue observed will be that of the stronger component, but it will have a washed-out appearance. (2) A mixture of two colored lights which are not complementary gives an intermediate hue. (3) If a mixture of two colored lights produces the same color sense

datum as another mixture, then a blending of the two mixtures produces the same sensation. For example, a certain yellow and a certain blue light when mixed in correct proportions appear gray; likewise a certain red and a certain bluish green light give the same sense datum. When these four colors are mixed in the original proportions, gray is still perceived.

Color Vision

Color zones of the retina. The retina of the eye has three zones with the spot of clearest vision (*fovea centralis*, Fig. 176) at the center. The zones differ somewhat in contour for different individuals, but their relative positions are always as shown in Fig. 163. The outermost zone gives only colorless or achromatic sense data no matter what the nature of the stimulus. Thus, when the image of an object is focused in this zone, the object will appear gray. The middle zone is sensitive not only to white and black, but also to yellow and blue. The innermost zone, which lies immediately about the *fovea*, is sensitive to red and green as well as to white, black, yellow, and blue.

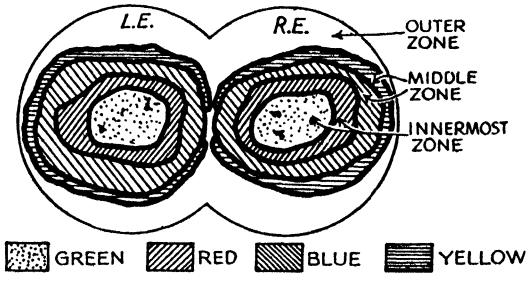


Fig. 163. Color Zones of the Retina.

A student may verify the presence of color zones by fixing the eye on some point ahead and then bringing a small red object with a shaking motion into the field of vision. It will appear a very dark gray, then a lighter gray, then yellow, and finally red as its image moves from the periphery of the retina to a point near the fovea. Ordinarily we pay attention only to the objects with images located on the innermost zone of the retina.

Color blindness. Approximately 3.5 per cent of the male and 2 per cent of the female population are color-blind.

Many of these have partial color blindness of the red-green type, in which all colors are seen as yellows, blues, and grays. Such persons would probably see the solar spectrum with the reds of longest wave length replaced by grays; the other reds, all the oranges and yellows, and part of the greens replaced by a band of yellows; the other greens and neighboring blues replaced by a band of grays; and the rest of the blues and the violets replaced by a band of blue. Such persons confuse red and green because both these colors appear to be either yellow or gray.

A very small portion of those suffering with color blindness live in a colorless world. To such persons the solar spectrum appears as a gray band lighter at the center than at the ends.

In many cases of partial color blindness, the red-green color zone of the retina is practically missing. In total color blindness, the middle and central zones are not developed.

Ladd-Franklin theory of color vision. The late Dr. Christine Ladd-Franklin has proposed a genetic color theory in which she suggests that total color blindness is evidence that the most primitive eye has colorless or achromatic vision, and that our present color vision is the result of an evolutionary development divided into three stages (Fig. 164).

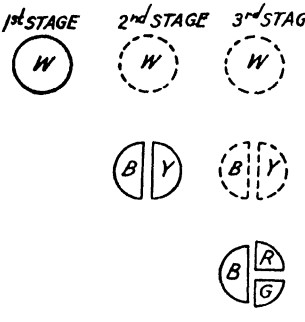


Fig. 164. Representing the Ladd-Franklin Color Theory.

In the first stage, a photochemical substance which would break down under any light stimulus was developed in the rods and cones. Thus, achromatic vision was achieved, and this is the condition which has persisted in the few cones of the outer zone of the retina and in all the rods. Under stimulation, these receptors still give the sensation of gray.

During the second stage, a differentiation of function and probably of structure took place among or within the mole-

cules of the photochemical substance of some of the cones. One type now responded only to the long wave lengths of the spectrum, and the sensation of yellow was added to vision. The other type responded only to the shorter wave lengths, and the sensation of blue was evolved. This situation is still found in the middle and inner color zones of the retina and represents the only vision possible by many of the persons with red-green color blindness. In the interest of simplicity, we shall call this new receptor the *yellow-blue* cone.

Finally in the third and last stage, a further differentiation took place in the "yellow" part of some of the *yellow-blue* cones, and again two new types of function and structure were evolved. One type was affected only by red light and the other only by green light, and thus a *red-green-blue* cone was developed. This condition is now found in the inner color zone of the normal eye and makes possible our present chromatic vision. This region, then, is composed of *yellow-blue* and *red-green-blue* cones.

According to the theory, complete decomposition of the photochemical substance is necessary to give the sensation gray, while anything short of this gives rise to some color sense datum. For example, the "yellow" part of a *yellow-blue* cone is decomposed by the action of yellow light, and the "red-green" part of a *red-green-blue* cone is decomposed by the action of red and green light mixed in proper proportions. Hence, either yellow light or a mixture of red and green light decomposes all but the "blue" portion of the photochemical substance, and yellow is sensed in either case. When a proper mixture of blue and yellow light stimulates the *blue-yellow* cones, a complete decomposition of at least some of the molecules of their photochemical substance takes place, and gray is perceived; also, when a proper mixture of red, green, and blue light stimulates the *red-green-blue* cones, the same action takes place and gray is sensed. Any color stimulus which does not completely decompose the photochemical molecule gives a color sensa-

tion. Thus, the results we obtained from mixing various color stimuli are adequately interpreted by this theory.

But a complete theory must include more than the retina in order to interpret and co-ordinate all the available experimental facts. For example, when a piece of apparatus designed so that one eye may look at a distant white light through a red window and the other at the same light source through a green window, care being taken to prevent any possible criss-crossing of colored rays, is tried out, a yellow light is seen. This must mean that the necessary color mixing has been performed not at the retina but in the brain itself. For further information on this important subject and a discussion of the Young-Helmholtz and the Hering theories of color vision, the student is referred to the books listed under "Suggested Readings."

Three Characteristics of Color

We often speak of the amount of light, meaning the energy content or flux, and the kind of light, meaning the frequency of vibration or wave length. As we pointed out in the last chapter, the amount of light which may be obtained from a source depends upon its intensity (candle power). The amount of light given out by an illuminated body depends upon the illumination (foot-candles) under which it is placed and the ability of its surfaces to reflect. Light may be classified roughly into two kinds: the single-frequency kind (monochromatic) found in very narrow patches of a well-spread-out solar spectrum, and the mixed-frequency variety reflected from green trees, red roses, and blue violets and found in a "perfect" mixture in white light itself.

We shall be interested to find the relationship between the two characteristics—amount and kind—of the light stimulus and the three characteristics—brightness, hue, and saturation—which form the data of vision.

Brightness. When a light source shines on a white screen, the brightness of the surface depends on the illumi-

nation. With no illumination whatsoever, the surface looks black; but when the illumination gradually increases from nothing to the greatest possible, a series of neutral grays ranging from the darkest black to the brightest white is observed. This series of grays ranging from black to white is used as a standard by which to judge the brightness of color. For example, the brightness of a certain color is said to be that of the neutral gray which appears to have the same brightness. In the color pyramid (Fig. 165), the scale of brightness, extending from the darkest black to the brightest white, is laid off on the dotted axis between *B* and *W*.

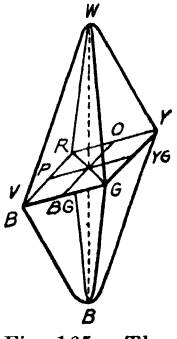


Fig. 165. The Color Pyramid.

Hue. If a hue is a spectral color, its stimulus may be characterized by a definite frequency. Most natural colors, however, are aroused by a mixture of light vibrations. Even so, any one of them may be matched with a spectral color characterized by a definite frequency, or with a color resulting from a measured mixture of frequencies. Thus, although purple is not found in the spectrum, it may be matched with a color resulting from a mixture of measured amounts of blue and red lights of definite frequencies.

In Fig. 165 the various hues are placed around the edge of the square which serves as a base for each of the two pyramids. Red, yellow, green, and blue are placed at the corners, and a few of the intermediate hues, including purple produced by a dual mixing of colors, are located in between. Thus, we interpret the drawing to mean that a mixture of green and yellow gives a yellow-green, a mixture of red and yellow gives orange, a mixture of red and blue gives purple or violet, depending on the quantity of red used, and finally, that a mixture of blue and green gives blue-green. The colors placed at the opposite ends of the lines drawn diagonally through the square are complementary, and the central point through which all the lines pass is located on the axis of neutral gray. This location

on the axis is necessary because two complementary colors produce gray when fused. Yellow is shown at a higher vertical position than blue because it appears brighter in the spectrum.

Saturation. The perimeter around which the colors are placed represents the position of complete saturation. This means that when a color is found on this line, its stimulus is not diluted with white light and that the color sense datum is not fused with gray. When neutral gray of the same brightness is added to a saturated color (white light added to the color stimulus in a correct amount), the brightness does not change, but the color becomes washed out and the color position moves from the perimeter toward the axis representing neutral gray. Finally, if the addition of neutral gray continues, a neutral gray having the same brightness as the original saturated color is observed.

With increased brightness, saturated colors take on a washed-out appearance; and at a high enough brightness, all colors appear white. To represent this decrease of saturation with an increase of intensity, the upper part of the figure is shown as a pyramid with the apex at *W*. On the other hand, when the brightness of the saturated colors is reduced, they become darkened and finally look dark gray or black. However, under decreasing illumination, red is known to preserve some of its color until it disappears entirely. To illustrate this darkening effect, the figure is made to converge to the apex *B*, and the double pyramid is formed. The darkening hues seen as night comes on are examples of colors decreasing in brightness and saturation.

Thus, any point within the double pyramid represents a hue (also gray, if on the axis between *B* and *W*) of definite saturation and definite brightness. One may recognize about 230 spectral hues. These would be located around the base of the double pyramid. It is estimated that the normal person may distinguish from 500,000 to 600,000 colors, when possible differences in hue, brightness, and saturation are taken into account. This number of

points, each representing a distinct color, can be located within the color pyramid.

Colors in Art

The immediate problem of the artist is pigment mixing, not light mixing. Of course, illumination and vision are involved in the production and final appreciation of a colorful painting, and the physical properties of light and the data of vision must be understood and utilized by him; but for practical reasons, the artist's theory of color utilization must be based on the action of pigments as a source of color stimuli. It simplifies matters, therefore, if a color is thought of, in this field, as the sensation produced by a certain commonly used, standardized pigment rather than a sense datum aroused by light of a certain wave length. But in the mixing of pigments, selective absorption is increased. Therefore, the range of the wave lengths emitted by the mixture is reduced. Thus, the net effect on the eye of mixing two colored lights is very different from that produced by mixing two pigments of the same colors. It becomes clear, therefore, why the so-called primary and complementary colors of art may be very different from the primary and complementary colors of physics and psychology.

Hue, value, and intensity. The artist uses three terms to describe a color: (1) hue, to determine its color; (2) value, to determine its relation to white and black; (3) intensity, to designate its strength of hue as compared with colorless gray (Fig. 166).

These characteristics of color, resulting from the stimulation of the eye with light from illuminated pigments, are analogous to but not identical with the characteristics of color listed above. The analogy may be indicated as follows:

Pigment Colors:	hue	value	intensity
Colored Lights:	hue	brightness	saturation

A color stimulus from a pigment is always mixed with white light. Thus, the colors used in art are never saturated, and they never have as great a color intensity (saturation) as do colored lights. A change in brightness, as we have seen, is due to the change in energy content of the light used as a stimulus. But a painting must be viewed under a more or less standard illumination. Brightness variations needed in a painting, therefore, may not be achieved by a change in the energy content of the colored light emitted by a pigment. But if we remember (Fig. 165) that with increasing brightness all saturated colors appear washed out, and that with decreasing brightness they appear darkened, then we see a solution of the problem. The artist may accomplish the analogous effect of an increase in brightness by adding a white pigment, or an analogous effect of a decrease in brightness by adding a black pigment or its equivalent. This modification he calls a change in *value*, the value being increased if white is added and decreased if black is introduced. This change of value, accomplished really by adding or subtracting the quantity of white light reflected from the pigment, decreases the color saturation. But the hue is said to have its full intensity for *this new value* unless neutral gray, in addition to the white or black needed to change its value, is added. The intensity may be decreased without a change in value if neutral gray of the same value is added (Fig. 166).

Thus, we see that the use of different terms in art and science to designate the three characteristics of color is justified because, although the characteristics described are analogous, they are not identical.

Mixing pigments. When very small patches of red and green pigments, placed in juxtaposition, are viewed from a distance, the *red light* and *green light* become mixed, and yellow is perceived. But if these pigments are thoroughly mixed, gray is seen. Red pigment is red because it absorbs yellow, green, and blue from white light; green pigment is

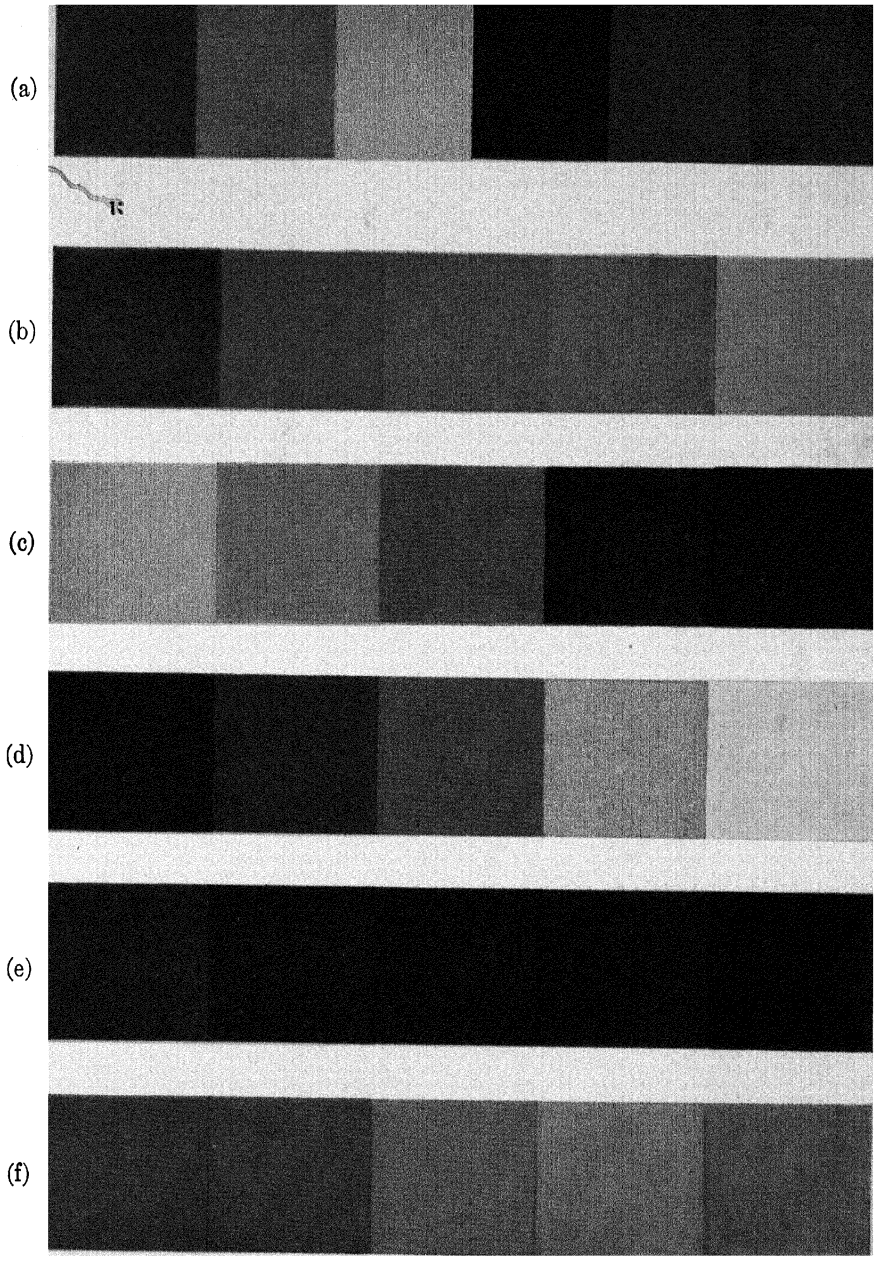


FIG. 166. COLOR CHART

(a) **Hues.**
(b) Mixing two adjacent hues (red and orange).
(c) Mixing yellow and its near complement, red-violet.
(d) Tones of green showing an increase in value.
(e) Tones of green showing a decrease in value.
(f) Illustrating a decrease in intensity without a change in value—neutral gray of the same value has been added in increasing amounts.

(Facing p. 280)

green because it absorbs orange, red, and violet from white light. When thoroughly mixed, all colors are absorbed, the selective absorption is complete, and the mixture would appear black were it not for the small amount of white light reflected from the surface. The final result is a sensation of gray (Fig. 167). Mixing pigments always increases selective absorption and thus eliminates certain of the color stimuli which were present before the mixing. In the mixing of colored lights, no color stimuli are thus subtracted. We see clearly, therefore, why the mixing of red and green pigments *is not the equivalent of adding red and green stimuli.*

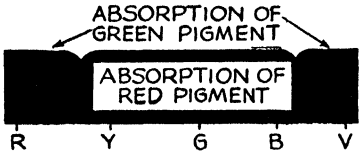


Fig. 167. Mixing Pigments Increases Selective Absorption.

Primaries, secondaries, and complements. In Fig. 168 the colors are arranged in their order of appearance in the spectrum. Red, yellow, and blue are called the *primary* pigments. Orange, green, and violet—the *secondary* pigments—are obtained by mixing, respectively, red and yellow, yellow and blue, and blue and red pigments. The intermediate hues are obtained by mixing the adjacent pigments in proper proportions. Gray is placed at the center, because when all the pigments are mixed, complete absorption is achieved and a small amount of white light reflected from the surfaces is perceived as gray. The pigments at opposite positions in the color circle, when mixed in proper proportions, also produce gray, and they are said to be *complementary.* It is possible to decrease the intensity of any hue by mixing with the pigment a little of its complement. The gray manufactured by this mixing process always produces in the dominant hue a

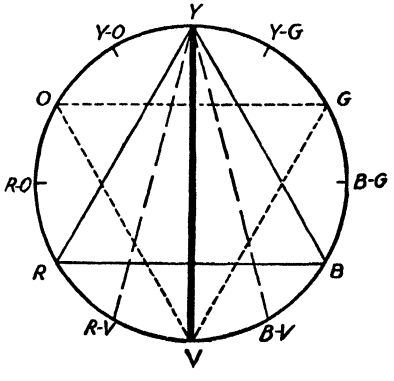


Fig. 168. The Color Circle.

decrease in intensity and often a change in value, if its value is different from that of the dominant hue.

Color groups. In general, color groups may be divided into two classes: (*a*) similar groups; (*b*) contrasting groups.

(*a*) *Similar color groups.* A single hue may be used effectively if it is combined with neutral gray. The necessary contrasts are achieved by the use of the various values and intensities of the hue. Also, a color and two adjoining hues may produce a pleasing effect. The differences in color are sufficient to add variety without destroying the color tonality (Fig. 166b, d, e, and f).

(*b*) *Contrasting color groups.* A color and its exact complement may be used to give a maximum of contrast. A color and its near-complement may be used to give a contrast that is softened by intermediate hues (Fig. 166c).

When two exact complements are mixed, as, for example, when small amounts of violet are added to yellow, the complement added (the violet) has the effect of decreasing the intensity of the other (the yellow) without changing its hue. With such a combination, therefore, only the two pure colors and all their degrees of intensity, including neutral gray, may be obtained. However, if yellow is mixed with red-violet, its near-complement, a range of colors varying in hue as well as in intensity may be created (Fig. 168). A slight amount of red-violet changes yellow to orange-yellow with a slight decrease in intensity. A greater proportion of red-violet changes yellow to orange with a still further decrease in intensity. Equal proportions of red-violet and yellow produce a reddish orange of a rather low intensity. A still further increase of red-violet carries the mixture with increasing intensity through red to red-violet. Thus contrast is achieved with a very effective blending of hues (Fig. 166c).

With care and skill, triads such as the primaries, red, yellow, and blue, and the secondaries, orange, green, and violet, may be united in a very effective manner.

Questions and Problems

1. Explain why two colors matched in artificial light may seem different in hue when viewed in daylight.

2. What color is produced (*a*) when yellow and blue lights are mixed? (*b*) When yellow and blue pigments are mixed?

3. Why does ice turn white when crushed?

4. (*a*) Why is a lake blue on a clear day? (*b*) In a storm, why are *whitecaps* white?

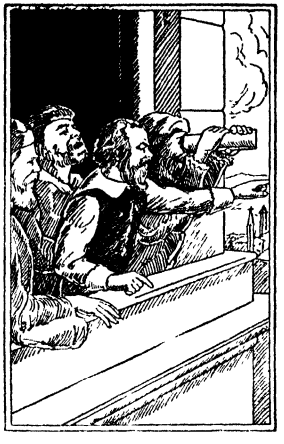
5. Explain why an artist may add white to red pigment to increase its apparent brightness.

6. Contrast mixing colored lights and mixing pigments.

7. Yellow pigment and its near-complement, blue-violet, are mixed. (*a*) List the hues produced. (*b*) Compare the relative intensities of the hues.

Suggested Readings

(1) Breese, Burtis B., *Psychology*, Charles Scribner's Sons, New York, 1917, Chap. VIII.
(2) Luckiesh, Matthew, *The Language of Color*, Dodd, Mead and Company, Inc., New York, 1918.
(3) Robinson and Robinson, *Readings in General Psychology*, University of Chicago Press, Chicago, 1923, Chap. VII.
(4) Sargent, W., *The Enjoyment and Use of Color*, Charles Scribner's Sons, New York, 1923, Chaps. I–VII.
(5) Saunders, F. A., *A Survey of Physics*, Henry Holt and Company, Inc., New York, 1930, Chap. XXXV.



CHAPTER XIV

Optical Instruments

The Plane Mirror

Metallic mirrors were probably manufactured by the ancients. Looking glasses are referred to in Exodus, and they have been found in graves with Egyptian mummies. Certainly they play an important role in modern times. One learns early that the object seen is not behind the mirror but in front of it. That which appears behind we call the *image* of the object which is in front. Light does not actually come from the image as it seems to do. Images of this type are called *virtual* to distinguish them from the real ones seen projected on a screen in a moving picture, for example. Through the use of the law of reflection of light—the angle of incidence is equal to the angle of reflection—the paths of rays of light are traced as shown in Fig. 169. Triangles *ABD* and *EGH* are isosceles, making the line *AC* equal to the line *CD*, the line *EF* equal to the line *FH*, and the line *AE* equal to the line *DH*. The image in a plane mirror, therefore, has these four characteristics: (1) it appears *behind the mirror;* (2) it is a *virtual image;* (3) it seems to be just as *far behind the mirror* as the *object is in front of it;* (4) it appears to be the *same size as the object.*

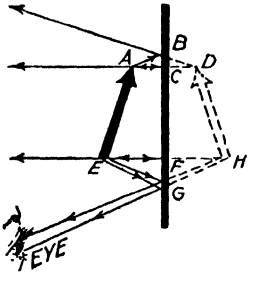


Fig. 169. The Plane Mirror.

The Convex Lens

The focal length and principal focus. In a previous chapter, we showed that a convex lens has the ability to

focus light. Let a beam of sunlight be focused on a screen. The distance from the lens to the screen is called the *focal length* of the lens. The image of a distant mountain will also be located at this same distance. Light which comes from such distant objects is made up of plane waves or parallel rays. Therefore, the focal length f of a lens may be defined as the distance from the center of the lens to the point where parallel rays are brought to a focus. This point of focus is called the *principal focus* (F, here).

Secondary foci. When an object is placed twice the focal length from a lens, the image is found on the opposite side at exactly the same distance from the lens. These points where the image and the object are the same size and at the same distance from the lens are called the *secondary foci* S.

The nature of an image. If an image is larger than the object, it is said to be *enlarged;* if smaller, it is said to be *reduced.* An image may be upright or it may be inverted. If it can be projected upon a screen, it is said to be *real;* if the light just appears to come from a point of focus, the image is *virtual.*

Series of images produced by the lens. The object may be located, in general, in one of three regions: farther from the lens than a secondary focus S, between a secondary focus and the principal focus F, and between the lens and the principal focus. *If the object is beyond S, the lens is essentially a camera,* and produces inverted, real, reduced images located on the other side of the lens between S and F. If the object is between S and F, the lens is essentially a projection lantern, and produces inverted, real, enlarged images located on the other side of the lens beyond S. If the object is between the lens and F, the lens is essentially a reading glass or simple magnifying glass and produces upright, virtual, enlarged images which can be seen only by looking *into* the lens. Fig. 170 illustrates these three typical cases. In making such drawings as these, we remember that parallel rays pass through the principal focus F, and that the general direction of a ray passing

through the center of the lens is not changed, the air-glass and glass-air surfaces encountered being parallel (Fig. 152).

Relative size of object and image. A study of the isosceles triangles (Fig. 170) formed by the two rays which pass from the extremities of the object to the image through the center of the lens leads to the conclusion that the object is as many times larger or smaller than the image as its distance from the lens is larger or smaller than the corresponding image distance.

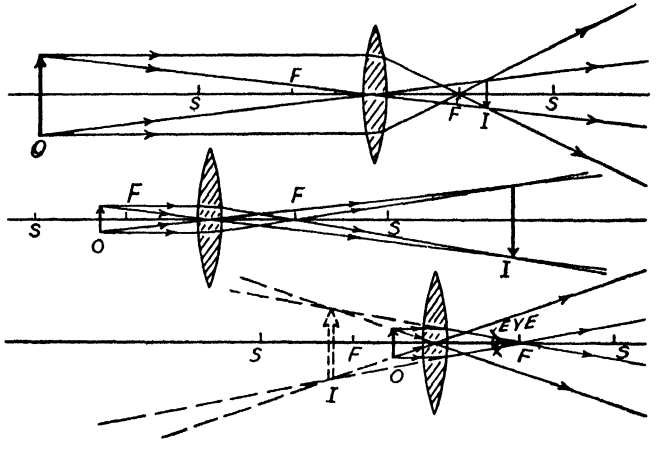


Fig. 170. Images Formed by a Convex Lens.

Numerical aperture of a lens. The ratio of the diameter of a lens to its focal length is called its *numerical aperture.* This number is an index of the amount of light which the lens gathers and puts into a unit area of image, because the larger the diameter, the greater the quantity of light which enters, and the shorter the focal length, the smaller the image (for a given object distance) over which this quantity of light is spread.

Defects of a Convex Lens

Spherical aberration. Light rays from distant objects passing through a lens with spherical surfaces do not come to the same focus unless the lens has a small numerical

aperture. This defect, known as *spherical aberration*, is shown greatly exaggerated in Fig. 171. The rays passing through the edge are brought to a focus nearer to the lens than those passing through the central position. The blurred image produced may be shown by projecting the image of a piece of wire cloth upon a screen by means of a projection lantern in which the projection lens is replaced by a plano-convex lens with large numerical aperture—large diameter and short focal length. The blurring is decreased when the curved surface is directed toward the illuminated wire cloth.

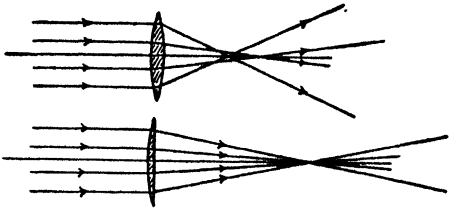


Fig. 171. Spherical Aberration.

This defect, spherical aberration, caused by the use of spherical surfaces, may be avoided or at least reduced in three ways. First, the lens may be specially ground—usually a process done by hand and hence expensive—so that the surfaces are not quite spherical. Second, the defect may be greatly reduced by the selection of the lens shape best adapted to the proposed use. For example, in a telescope, where the light used comes from distant objects, the parallel rays passing through all parts of a plano-convex lens will be brought nearly to the same focus if the *convex surface is toward the distant object viewed* (Fig. 171). In the case of a microscope, on the other hand, where the object is very close to the lens, the plane surface should be placed toward the object so that spherical aberration will be reduced. Third, the spherical aberration may be diminished if the effective diameter of the lens is reduced by being "stopped down." The opaque diaphragm used for this purpose permits light to pass through only the central portion of the lens, and thus the amount of light transmitted is reduced. Often this shutting out of part of the light places an undesirable limitation upon the optical instrument using the lens, and it is necessary to resort to one or both of the other methods of correction. Also,

under certain conditions, a stop may produce an undesirable distortion of the image, as we shall now see.

Distortion. When a stop (a cardboard with a round hole at its center) is placed between the object and the lens in the experiment outlined above, the image shows a barrel-

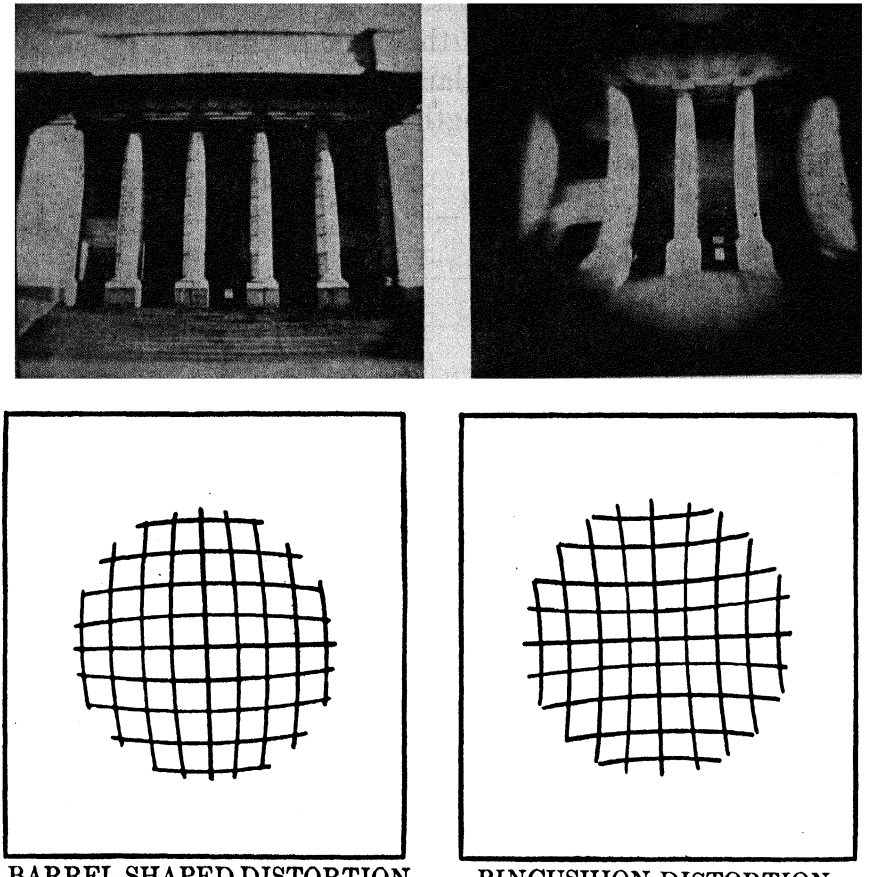


Fig. 172. Distortion.

shaped distortion; when placed between the lens and the image, a pincushion-shaped distortion is observed (Fig. 172). When two lenses are used and the stop is placed between them, the distortion is eliminated. By this arrangement, a stop may be used to reduce spherical aberration without resulting distortion.

Astigmatism. When the lens is now placed so that light passes obliquely through it, the horizontal and vertical lines of the wire cloth image are not at the same focus (Fig. 173). This defect is called *astigmatism* and is always present at the extreme edges of the image formed by an uncorrected lens because the light which forms these image parts passes obliquely through the lens. A lens which is corrected for this defect by proper grinding and a proper combination of different kinds of glass is called an *anastigmatic* lens.

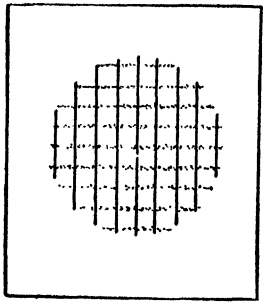


Fig. 173. Astigmatism.

Chromatic aberration. We have already pointed out the fact that red and blue light are refracted differently on passing through a prism. Because of this, red light will be brought to a focus at a greater distance from the lens than will blue light, and a coloring is seen at the edges of the projected image of each element of the wire cloth.

Newton assumed that dispersion—separation of light into colors—is directly proportional to the refraction—bending of the rays—at a surface, and saw no means of preventing dispersion without also eliminating refraction. However, this assumption is not justified. It is found that *crown glass* refracts light more for a given dispersion than does *flint glass*. Hence, a *dispersion* produced by crown glass may be annulled by flint glass without all the *bending* in the crown glass being eliminated. Thus, a lens combination such as that shown in Fig. 174 causes light to be bent but not dispersed. Such a lens is said to be *achromatic*.

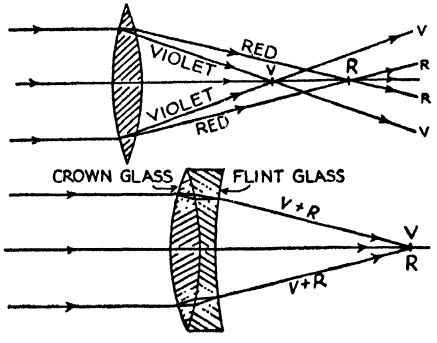


Fig. 174. Achromatic Lens.

The Camera

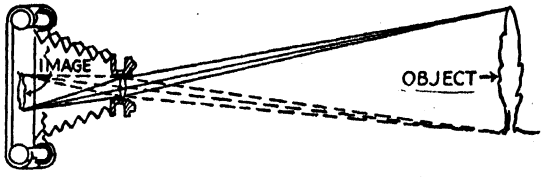
Operation. A camera consists essentially of a lens, a stop, a shutter, a light-tight box, and a light-sensitive film opposite the lens (Fig. 175). The stop controls the size of the opening through which the light enters and determines how much of the area of the lens is used. The shutter controls the time during which light may pass into the light-tight box. The sensitive film is the screen upon which the lens forms the image of the object. The chemicals placed on the film are affected by the intensity of the light received. When the film is developed, dark areas appear where the light was most intense and transparent areas where the intensity was the least.

The following is a list of various films in common use, together with their color ranges and sensitivities measured in Scheiner degrees—units such that each increase of three degrees gives a doubling of sensitivity and a halving of the time needed for proper exposure.

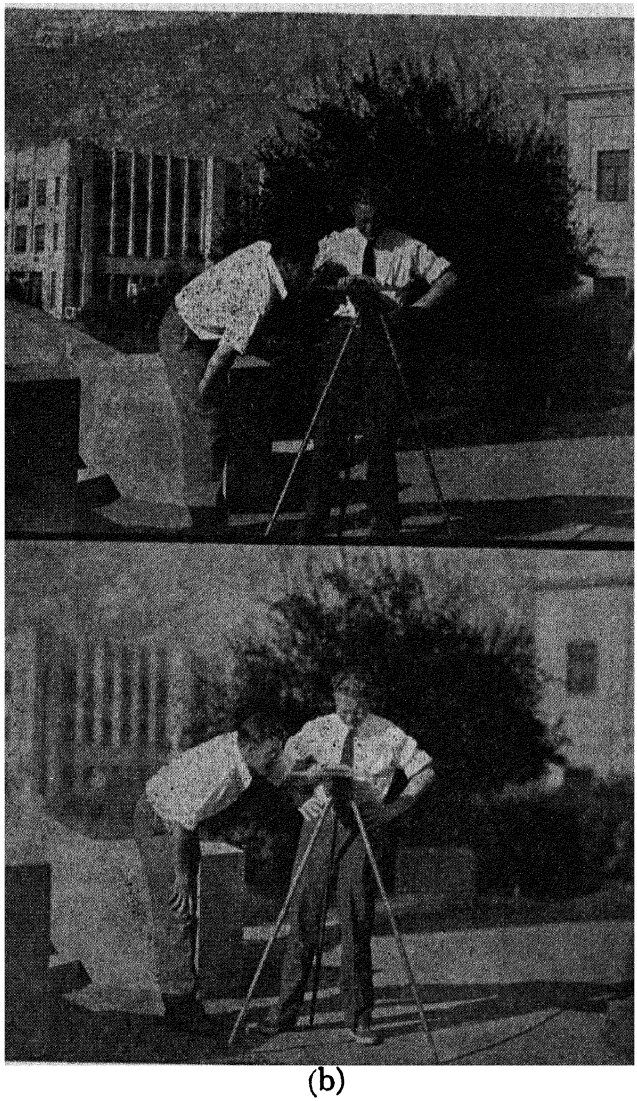
COLOR RANGES AND SENSITIVITIES OF COMMON FILMS

Film	*Color Range*	*Sensitivity in Scheiner Degrees*
Commercial Nitrate	Violet to yellow	12
Commercial Orthochromatic	" " yellow	15
Regular Roll Films	" " yellow	17
Commercial Panchromatic	" " red	18
Verichrome	" " orange	19
Plenechrome	" " orange	20
Supersensitive Plenechrome	" " orange	23
Supersensitive Panchromatic	" " red	23

Brightness of image. The brightness of the image in a camera depends upon three factors: (*a*) the brightness of the object photographed, (*b*) the area of the stop opening, and (*c*) the area covered by the image. The brightness of the image varies directly as the brightness of the object, directly as the area of the stop opening, and inversely as the area of the image. The brightness of the object may be changed by artificial illumination, by a flashlight, or by simply



(a)



(b)

Fig. 175. (*a*) Camera. (*b*) Depth of Focus Increases When a Small Stop Is Used. (Notice that the distant mountain is not in focus with stop set at f.4.5, but is in focus with the stop at f.32.)

moving objects which are in shadows into the light. Very often a change in illumination is not possible or practicable. This means that the brightness must be increased either by reducing the size of the image or by increasing the area of the stop. The size of the image for a given camera is fixed by the focal length; hence the only variable left is the change in the size of the stop. If the brightness needs to be reduced to permit proper exposure, the problem is simply one of closing the stop. But if, because of action in the scene photographed, the slowest shutter speed permissible is not slow enough to give the film a proper exposure under the illumination available, even with the stop wide open, then picture taking must be postponed, a more sensitive film installed, or a faster camera secured.

What controls the speed of a camera? The shutter speed used for proper exposure increases as the brightness of the image increases. For distant objects, the image is formed at the principal focus of the lens; hence, the length or width of the image is directly proportional to the focal length of the lens, and the *area* over which the light spreads is directly proportional to the square of the focal length. For a given illumination, the amount of light which passes through the lens in one second is directly proportional to the area of the lens, or to the square of the diameter. This means that for constant illumination, the brightness of the image (the light energy concentrated on a unit area) is directly proportional to the number obtained by dividing the square of the diameter of the lens (the amount of light entering) by the square of its focal length (the area of the image), or is simply proportional to the square of the numerical aperture. But since the shutter speed varies directly with the image brightness, the speed of a camera is also proportional to the numerical aperture. Finally, we have for the speed of a camera the equation

$$\text{Speed} = K\frac{D^2}{f^2}, \tag{14.1}$$

where K is a constant of proportionality, D the diameter of the lens or stop used, and f the focal length of the lens.

The so-called "f"-*value* of a lens is defined by the relation,

$$f\text{-value} = \frac{f}{D}. \tag{14.2}$$

This means that the speeds of two lenses with the same f-value are the same even though the lenses differ in focal length or diameter, and though one is a large portrait lens and the other the lens of a small camera. It also means that the speeds of two lenses with different f-values bear the inverse ratio of the square of their f-values. For example, a single achromatic lens in camera construction has a maximum value of $f.14$, a so-called rapid rectilinear lens has a maximum value of $f.8$, and an anastigmat lens of moderate cost a value of $f.6.3$. Obviously, the anastigmat lens is the fastest and the single achromatic lens the slowest in the group. To make a quantitative comparison, we use the relations given above and find the speeds to be in the ratio $(\frac{1}{14})^2 : (\frac{1}{8})^2 : (1/6.3)^2$, or in whole numbers, in the ratio 51 : 156 : 252. Thus, the speed of the rapid rectilinear lens is over three times that of the single achromatic lens, and the speed of the anastigmat is 61 per cent greater than the speed of the rapid rectilinear. It is possible to obtain a speed less than the maximum by closing the stop, since the speed of a given stop varies directly as the area of the opening or as the square of the diameter.

When just the center of a lens is used, the defects outlined above are practically eliminated; but if the edges are used, the lens must be corrected. Hence, high-speed lenses are corrected lenses and therefore expensive.

Depth of focus. If the opening into a camera were the size of a pinhole, no lens would be needed; all objects, near-by and distant, would be focused on the film. Thus, in this case the depth of focus—the distance range between objects which may be focused well at the same time—is at a maximum. Such a camera would be very slow indeed

because of the smallness of the stop opening. (Minutes instead of seconds would be required to make a proper exposure.) For this reason, a lens is used. But with its use, near-by and distant objects are not brought to the same focus, and the range between objects which may be focused at the same time is very much shortened and the depth of focus is reduced. However, when the lens is stopped down, the depth of focus is increased because the pinhole camera is more nearly approximated (Fig. 175).

The images of near-by and distant objects produced by a short-focus lens are nearer together than when formed by a long-focus lens. Thus, by the action of a short-focus lens, objects over a rather wide range are brought into good focus at the same point; in other words, such a lens has a greater depth of focus than that of one with a long focal length.

Fixed focus. It is impossible for a lens to be *exactly* focused at the same time on objects at different distances; but its depth of focus may be so great as to make it possible to set the lens at a fixed distance from the film and obtain an exact focus for objects at intermediate distances while the images of distant and near-by objects are also rather sharp. This "fixed focus" camera does not have speed because of the small stop used to get depth of focus, but it is a very splendid instrument for the use of the amateur.

General suggestions. Remember that the best part of an uncorrected lens is its center. Thus, when using an inexpensive camera, have the stop as small as can be used successfully. This procedure adds to the depth of focus and in a landscape gives good detail to all parts of the picture. If there is action in the picture, a snapshot must be made; but the shutter of the slowest speed consistent with the type of action should be used so that the stop may be closed as much as possible. When the light is poor and the time of exposure needs to be short, the stop must be opened wide and greater care must be taken in focusing since the depth of focus is decreased. Under poor light

do not expect a slow lens to give the same results as a fast lens; but remember that under excellent light, where a small stop is used, a slow lens may give just as good pictures as a fast one. Always keep in mind that it takes a certain amount of light energy to expose a film properly, and that the amount of light energy that gets into a given camera depends upon the illumination of the objects photographed, the size of the stop, and the length of exposure.

The Eye

How the eye functions. Of all optical instruments, the eye is the most important. As an optical instrument, it is similar to a camera. The image is reduced, real, and inverted, and forms on the retina, the light-sensitive portion of the eye. The iris is a pigmented membrane which gives color to the eye and serves as a stop, reducing the spherical aberration and controlling the amount of light which enters. When the pupil is dilated by the use of some drug such as atropine, the effectiveness of the iris in eliminating spherical aberration is clearly detected.

As light enters the eye, it must pass through the cornea, the aqueous humor, the crystalline lens, and the vitreous humor before it reaches the retina (Fig. 176). The speed of light in the cornea and the aqueous humor is practically the same, so together they form a rather thick lens. The crystalline lens lies just back of the iris and is in contact with the aqueous humor in front and the vitreous humor behind. This lens is not a homogeneous body, but consists of a series of concentric layers, the inner ones being more dense and therefore having a greater refracting power than the outer ones. This helps to reduce spherical aberration. Focusing is accomplished by a change of the curva-

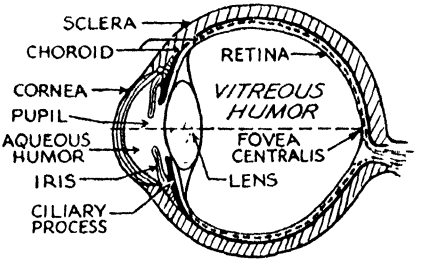


Fig. 176. The Eye.

ture of the front surface of the lens, the curvature being greater when near-by objects are viewed. With the increased curvature comes a slight decrease of the diameter of the pupil, which tends to offset the increased aberration produced by the added curvature. Chromatic aberration is corrected by the combination of the front thick lens (cornea and aqueous humor) and the crystalline lens.

Defects of the eye. The principal optical defects of the eye are: astigmatism, near-sightedness, and far-sightedness.

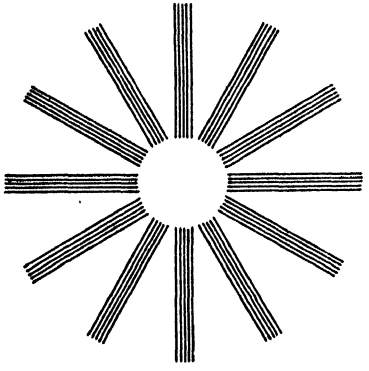


Fig. 177. Lines to Test Astigmatism.

Astigmatism is due to the unequal curvature of the cornea. In looking at a series of heavy black lines forming the radii of a circle (Fig. 177), the eye with astigmatism will see some of the lines well defined and others blurred. This defect is corrected by the use of cylindrical lenses.

Near-sightedness, or *myopia*, is usually caused by an abnormal length of the eyeball; yet it may be caused by too great a curvature of the crystalline lens. A person with such eyes cannot see distant objects clearly. The defect is corrected by the use of concave lenses (Fig. 178).

Far-sightedness, or *hypermetropia*, is caused by the abnormal shortness of the eyeball. Even the image of distant objects may fall behind the retina, and no object may be seen distinctly without an effort of accommodation. This defect is corrected by the use of a convex lens (Fig. 178).

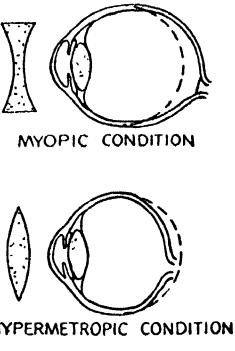


Fig. 178. Eye Defects.

Another type of far-sightedness comes to most people in later life. The hardening of the crystalline lens makes accommodation impossible. Distant objects are clearly

seen, but near-by objects cannot be seen without the aid of convex lenses.

Resolving power of the eye. As one views the headlights of an automobile a half-mile or more away, they appear as one light. When a distant brick building is looked at, the separate bricks are not seen. As a result of diffraction, discussed in the last chapter, the image of a point source, even in a completely corrected lens, is not a point but a bright spot. It is the overlapping of such spots that makes it impossible to distinguish the two headlights until the images are sufficiently separated. Two points will appear as one unless the angle separating them, as measured with the position of the eyes as vertex, is greater than one minute.

Other Instruments Using Lenses

The projection lantern. This instrument consists essentially of a lens which produces upon a distant screen an

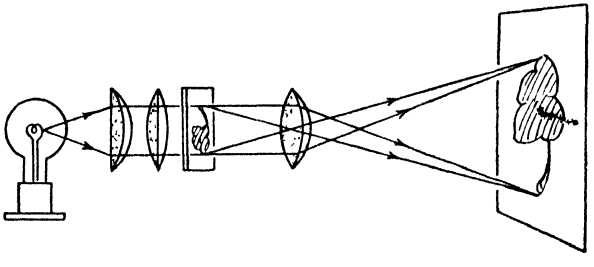


Fig. 179. Projection Lantern.

enlarged, real, inverted image of an illuminated object (Fig. 179). The object is placed between the principal focus and the secondary focus, and the image is far beyond the secondary focus on the other side of the lens. The object, usually a so-called lantern slide, is illuminated by a strong electric light, the light being concentrated upon the slide by a large condensing lens. The magnifying power of the lantern is equal to the distance v of the screen (image) from the lens divided by the distance u of the slide (object) from the lens, or simply v/u.

If one wishes to fill a screen completely with the image and at the same time locate the lantern at a definite distance from the screen, he may do so by the selection of the proper focal length of the lens. For example, suppose the dimensions of the screen are 50 times those of the slide; then the distance of the screen from the lens must be 50 times the distance of the slide from the lens. If the lantern is to be operated at a distance of 40 feet from the screen, what should be the focal length of the lens? By experiment it is found that the focal length of a lens is related to the distance of image and object as follows: $\frac{1}{f} = \frac{1}{v} + \frac{1}{u}$, where f is the focal length of the lens, v is the distance of the image from the lens, and u is the distance of the object from the lens. In the suggested problem, $v/u = 50$ and $v = 40$ feet. At once, $u = \frac{4}{5}$ of one foot, and

$$\frac{1}{f} = \frac{1}{40} + \frac{5}{4} = \frac{51}{40}$$

and $f = \frac{40}{51}$ of one foot, or 9.4 inches. Were the lantern to be 100 feet from this screen, a focal length of 23.5 inches would be required; at a distance of 20 feet, the focal length required would be 4.7 inches. Thus, a lens of long focus should be used far from a screen and a lens of short focus near a screen if the same magnification is expected. Unless a wide variation of distance is required and an exact filling of the screen is demanded, a lens of a certain focal length may be used for a considerable variation of the screen distance, owing to the large relative change that may be made in the *distance between the lens and the object.* Focusing is brought about by the variation of this distance. The moving picture machine is a complex form of projection lantern.

Sound motion pictures. The moving picture machine is a complex form of the projection lantern. The film serves as the object. It is made up of a series of still pictures which are jerked, one at a time, through the machine at a

rate of from 16 to 24 per second—the same rate at which the pictures were photographed, unless distortion of the action depicted is desired. A revolving disc cuts off the light while the film is jerked and once in between jerks, but permits light to fall on the screen twice while the film is momentarily at rest, the time of illumination being approximately four times the time of darkness. The eye fails to notice the shifting from light to darkness at this speed, and the observer fills in the action between the still images (Fig. 180). On sound picture films, a sound track is printed alongside the still pictures. Consulting the figure, you will note that this track is made up of variable density lines, tiny strips with varying degrees of opacity. During reproduction, a narrow beam of light is focused on the track and varying amounts pass through and into a photoelectric cell (Chapter XVI) which changes the varying light intensity into a changing electric current. The fluctuating electric current, after amplification, evokes mechanical vibrations in a loudspeaker and the recorded sound is heard. During recording, the sound waves, changed to variable electric currents by means of a micro-



Fig. 180. Sound Picture Film (Variable-Density).

phone, cause a "light valve" to open and close, thus modifying the exposure as the film moves very steadily at the rate of 90 feet per minute. This is the origin of a variable-density sound track. A variable-area sound track is also used effectively. Such a track has a saw-tooth appearance, thus permitting varying amounts of light to pass into a photoelectric cell, and sound is the end product.

The simple magnifying glass. As an object comes nearer and nearer, the angle which it subtends at the eye increases, and it appears larger and larger. However, when it is brought nearer than about 25 centimeters, its image ceases to be distinct. Thus, for close inspection, we bring an

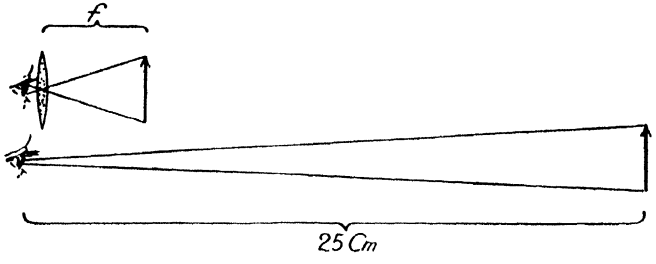


Fig. 181. Simple Magnifying Glass.

object to within about 25 centimeters of the eye. However, the enlargement due to a close-up view, which may not be used effectively by the unaided eye, may be utilized if a simple convex lens, which renders distinct the image of the very close object, is used. Such a lens is placed immediately before the eye, and the object is placed just within the principal focus of the lens. As the rays pass through the lens, they are rendered nearly parallel and may be focused by the eye with little accommodation.

At this new position, the object subtends a greater angle at the eye than it did when placed at the position of unaided distinct vision. Hence, there has been a magnification equal to the ratio of the two angles. This ratio, as may be determined from the geometry of Fig. 181, is $25/f$, where f is the focal length of the lens.

The compound microscope. This instrument consists essentially of an objective lens and an eyepiece. These lenses as used in practice are complex because corrections for defects have been made. We shall present the simple theory of image formation. The object is placed between the principal focus and the secondary focus of the objective lens. As in the case of the projection lantern, an inverted, enlarged, real image is formed. This image is located in the tube of the microscope and is viewed by a simple magnifying glass, the eyepiece. The eyepiece forms a virtual, enlarged, upright image of the real, enlarged, inverted image produced by the objective lens (Fig. 182). As in the case of the projection lantern, the magnification of the objective is v/u. But u is nearly the focal length of the objective lens, and v is approximately the length of the tube; hence, if we call the focal length of the objective F and the length of the tube L, the magnification due to the objective lens is L/F. But the magnification of the eyepiece is $25/f$, where f is its focal length. At once the magnification of the microscope becomes $25L/Ff$. This means that a large magnification is obtained by means of making the tube a convenient length and then making F and f as small as is consistent.

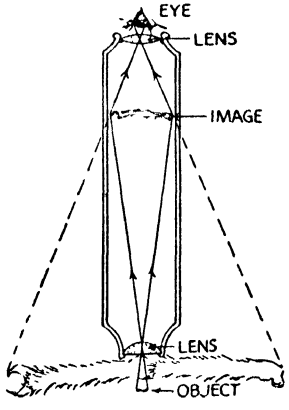


Fig. 182. Simple Compound Microscope.

Resolving power of a microscope. The smallest distance two objects may be apart and still be distinguished by the aid of a microscope as being separate entities is approximately one-half the wave length of the light used. Since the eye is not sensitive to wave lengths below about 0.00003 centimeters, we shall never be able to see minute corpuscles having a diameter smaller than 0.00001 centimeters. By the use of ultraviolet light and a photographic plate, photographs may be made of particles somewhat smaller than those which the eye can see directly.

The astronomical telescope. This instrument has two corrected lenses, the objective with long focal length and an eyepiece of short focal length. The objective forms an image at the principal focus which is inverted, reduced, and real. The eyepiece forms an enlarged, virtual, upright image of this real image (Fig. 183). The magnifying power of such an instrument is equal to the focal length of the objective divided by the focal length of the eyepiece

Offhand, it would appear that one could make the focal length of the objective so long and the focal length of the eyepiece so short as to permit one to see the activity of life on Mars. But here again, diffraction comes in to keep us from such detailed star gazing. The resolving power of a

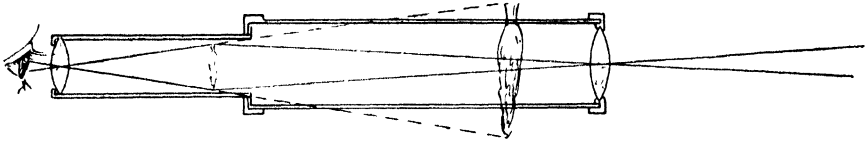


Fig. 183. Telescope.

telescope varies directly with the diameter of the objective lens, and there is a limit to the diameter of lenses which we are able to construct. Lenses up to a diameter of 40 inches have been constructed. One of such is located at the Yerkes Observatory in Wisconsin.

Telescopes with a concave mirror as an objective may be used. The largest of this kind in use is the 100-inch Hooker reflector at the Mount Wilson Observatory near Pasadena, California. This instrument has a magnification so great as to cause the moon to appear at a distance of about 50 miles. It collects 150,000 times as much light as the eye; hence, it discloses very faint and distant stars which were never known before.

Owing to diffraction, the greatest practical magnifying power of a telescope is never greater than 40 times the diameter in inches of the objective, and the minimum useful magnifying power is 3 times this diameter. This means that the maximum practical magnifying power of the

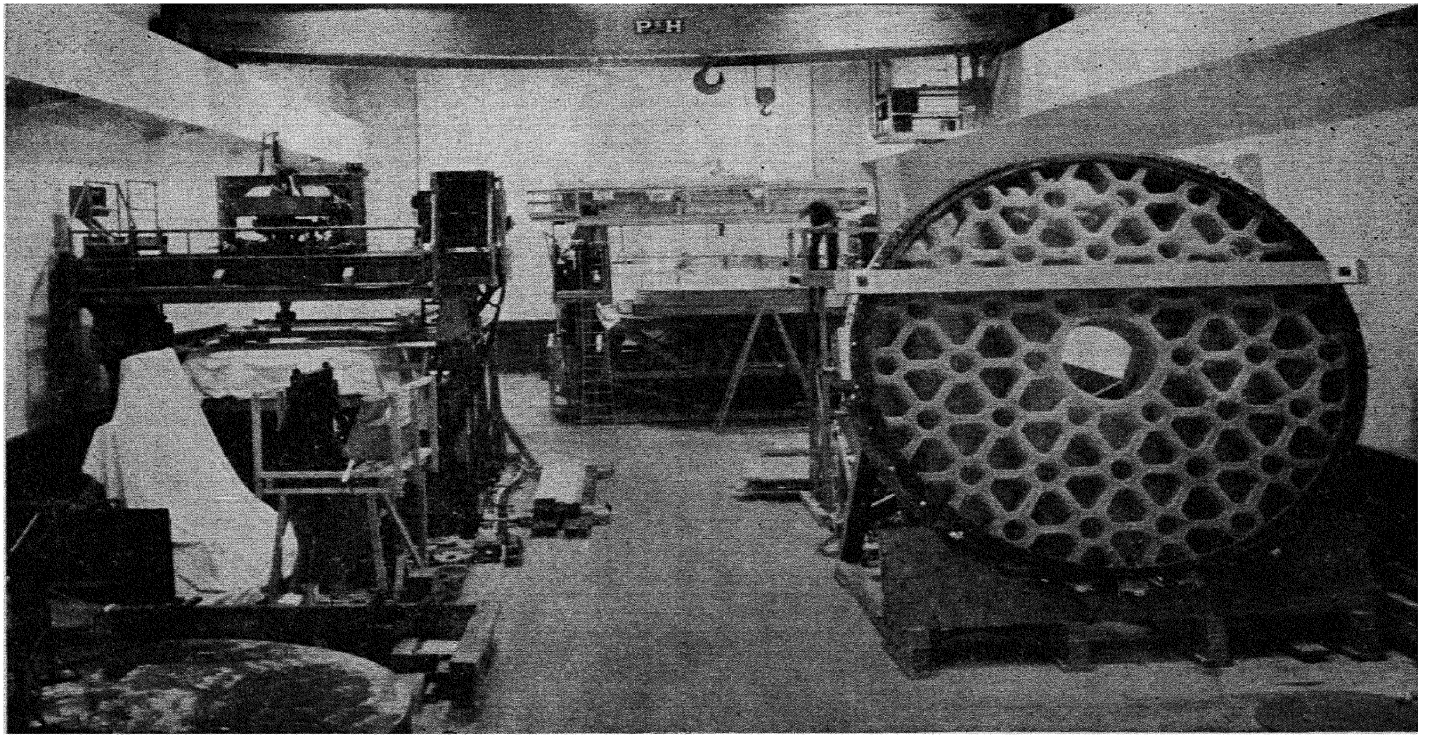


PLATE V. The Disc Which Will Serve As the Concave Mirror of the 200-Inch Telescope As It Appeared Standing on Its Rack Soon After Being Delivered (April, 1936) to the Optical Shop, California Institute of Technology, Pasadena, California. The disc itself with its back surface showing is at the right. The machine to be used in grinding and polishing the disc is behind and a little to the left. The polishing machine used to finish the 120-inch "flat"—a perfectly flat surface required for test purposes in the accurate grinding of the 200-inch concave mirror—is at the extreme left. (*Photograph through the courtesy of Dr. I. S. Bowen, California Institute of Technology.*)

Hooker reflector is 4,000; and if a magnification lower than 300 is desired, a smaller instrument can be used to better advantage. At the present time (1936) the glass for a 200-inch reflector has been cast, carefully cooled, and is now being ground and polished at the California Institute of Technology, Pasadena, California. When completed and installed, the telescope using this mirror will have twice the resolving power of the 100-inch one now in use at Mount Wilson Observatory. Man will then be able to reach more successfully for the detailed information to be found in the depths of the universe.

Questions and Problems

1. Show by constructing an image that a person may see his entire length in a mirror half his length.

2. List the characteristics of images.

3. What is the magnifying power of a simple magnifying glass with one-centimeter focal length?

4. How tall is a tree located 200 feet away if the image formed in a camera having a lens with four-inch focal length is one inch long?

5. Explain why a reflector with a large diameter is so much to be desired in an astronomical telescope.

6. List the common eye defects and give the cause of each.

Suggested Readings

(1) Bell, L., *The Telescope*, McGraw-Hill Book Company, Inc., New York, 1912, Chaps. I–XI.

(2) Hale, G. E., *The Depths of the Universe*, Charles Scribner's Sons, New York, 1924, Chap. I.



CHAPTER XV

Electrical Manifestations

We have no specialized sense organ uniquely adapted to respond to electrical stimuli, but all the senses co-operate in aiding us to detect the presence of electricity. Some of us think of wires, motors, electric lamps with their white-hot filaments, street cars, spark plugs, batteries, power plants, lightning, and thunder when electricity is mentioned. Others, realizing that these mechanical things are not electricity, speak of the "juice" which operates them. Undoubtedly there is a subtle something which acts behind the scenes.

During the materialistic eighteenth century, electricity, along with caloric, was considered to be an imponderable, subtle fluid. Today we believe that electricity is atomic in structure, that its invisible corpuscles are smaller than atoms, and that they are the "building stones" out of which atoms are made. Some of the evidences for this belief are discussed in Chapter XVII. We shall now point out some of the manifestations by which the presence of these invisible corpuscles may be detected.

Nature of an Electric Charge

Two kinds of electricity. Thales of Miletus (640–546 B. C.), one of the "seven wise men" of early Greece, is given credit for knowing how to generate electricity by rubbing amber with silk. The experiment, slightly modified in form, may be performed by anyone as follows:

Let a hard rubber rod be rubbed with a cat's fur (or let a fountain pen be rubbed on a coat sleeve) and then be brought near small bits of tissue paper. The bits are attracted to the rod. But when the hand is passed com-

pletely over the rod and it is again brought near the paper, the bits are not attracted.

The something appearing on the rod when rubbed with fur was called *electricity* by William Gilbert (1540–1603), and he is known as "the father of the modern science of electricity and magnetism." This name was chosen because the Greek word for *amber* is *elektron.*

We may continue our study of electrical manifestations by making use of a pith ball electroscope (electricity indicator) shown in Fig. 184 and constructed by the suspension of

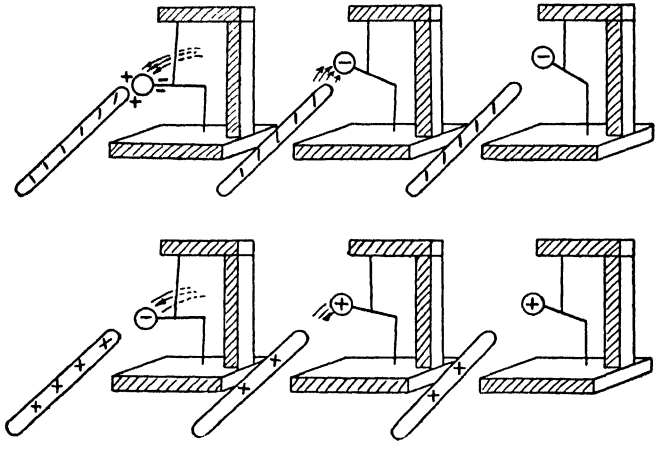


Fig. 184. Pith Ball Electroscope.

a ball of cornstalk pith by means of a silk thread. Let the rubber rod be rubbed with fur and then brought in contact with the pith ball. As soon as the ball becomes electrified, it moves away as if repelled. Now let a glass rod be rubbed with silk and brought near the pith ball; an attraction is observed. But when the charged rubber rod is again brought near, a repulsion is noted. Obviously, the electricity on the glass is different from that on the rubber. Now let the pith ball be charged by contact with a charged glass rod. At first the ball is attracted and then it is definitely repelled. But when a charged rubber rod is brought near, an attraction is observed. This simple experiment reveals the fact that the kind of electricity

on the rubber attracts the kind on the glass, but each repels its own kind. The electricity appearing on the rubber rod is called *negative;* that on the glass *positive.*

After years of research, Dr. Robert A. Millikan and others have found that negative electricity is made up of tiny corpuscles named *electrons.* Positive electricity has been found to be composed of corpuscles named *protons.* A proton has a mass about 1800 times as great as the mass of an electron. As we know, atoms are composed of protons, neutrons, and electrons. The protons and neutrons make up most of the atomic mass and are seldom found alone outside the atom. The electrons, on the other hand, often free themselves from the atom, and being small of mass, are agile in their movements. *A negatively charged object has an extra supply of electrons; a positively charged object has a lack of them.*

Law of attraction and repulsion. The law of electrical attraction and repulsion, first formulated by Charles Francois du Fay (1698–1739), may now be stated as follows: *Like charges of electricity repel, unlike charges attract; or electrons repel electrons, protons repel protons, but electrons and protons attract each other.* The charge (the quantity of electricity) making up the electron is exactly equal to that of the proton, but their kinds are different. When an equal number of electrons and protons are present in a body, it is neutral or uncharged. This is the condition present in most of the objects of our common experience. The electron charge has been measured very accurately by Millikan by the use of the now famous oil drop apparatus, which we shall describe later.

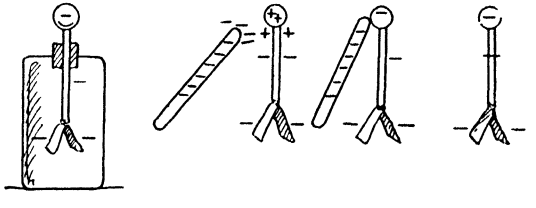


Fig. 185. Gold Leaf Electroscope.

A gold leaf electroscope may be used even better than a pith ball to show the presence of an electric charge

(Fig. 185). This instrument consists of an insulated metal rod with a pair of gold leaves attached at one end and a round knob mounted at the other end. When the ball is touched by an electrified rubber rod, the leaves move apart. This action illustrates the law of repulsion of like charges, because each leaf receives by conduction a part of the charge put on the knob.

Conductors and Insulators

Let a charge be placed on the gold leaf electroscope. When the knob is touched with a metal rod, the charge disappears suddenly. This is evidence that metal is a good conductor of electricity. Again let the electroscope be charged. But now when the knob is touched with a stick of sealing wax or sulphur, the leaves do not collapse. These substances are insulators. When a stick of wood is made to touch the knob, the leaves slowly fall. Thus, wood is a poor conductor of electricity.

In general, the metals are good conductors, their conductivity being better at low than at high temperatures. Carbon, however, conducts better when hot. Most insulators conduct somewhat when the atmosphere is humid. The film of moisture on the surface is responsible for this increased ability to conduct an electric charge. The best insulators for average conditions are amber, sulphur, and sealing wax. Glass and porcelain are commonly used as insulators in electric wiring and power lines (Fig. 186).

The condenser. In 1746, Pieter van Musschenbroek, a resident of Leyden, Holland, attempted to electrify a bottle of water. A friend, Canaeus, held the bottle in one hand, and after the charging had proceeded for some time, attempted to remove the wire which connected the water to the electric machine. He received an electric shock, a very new and unexpected experience. Musschenbroek then wished to have his chance, and no doubt he got it, for later in a letter he wrote that he "would not take another shock for the kingdom of France." This is the

beginning of the so-called "Leyden jar." Today such a jar is usually constructed with tin foil pasted on the inside and outside of the glass.

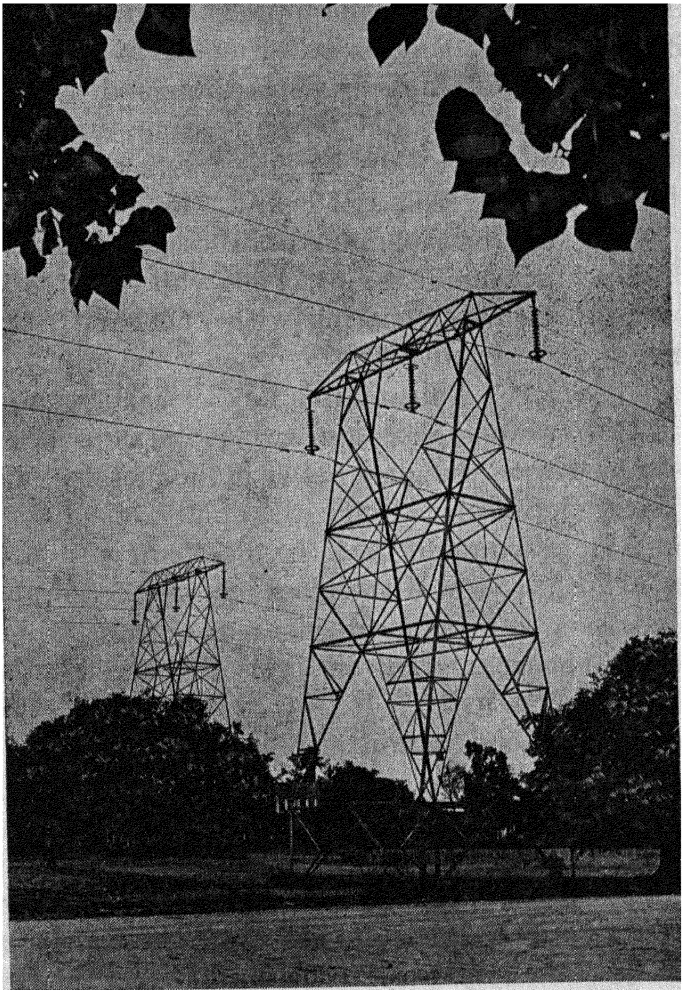


Fig. 186. Conductors and Insulators on a High Tension Line.

When a Leyden jar (Fig. 187) is held with its knob near an electric machine (static machine), sparks jump across the short air gap. When the charging is complete, let the jar be placed on the table. If the outer and inner coatings are connected through a very short air gap by the use of a dis-

charging rod, as shown in Fig. 187, a powerful spark will be produced.

Again let the jar be charged and then placed on a piece of glass so that the outer metal is insulated. If one touches the knob with the finger, no appreciable discharge will be felt; if one touches the outer covering, again no appreciable discharge will be detected. But if the inner and outer coatings are connected with a discharging rod, a thick spark will be seen and heard. If you wish to have the experience of Musschenbroek, *partially* charge a jar, and while contacting the outer coating with one hand, touch the knob with the other.

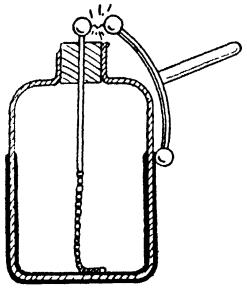


Fig. 187. The Leyden Jar.

A dissectible Leyden jar is composed of one glass and two metal cups which may be assembled with the glass between the metals. If such a jar is charged and placed on a glass plate, one may remove the inner metal cup and then the glass cup without experiencing a shock. Now if the glass jar is placed in the outer metal covering and the inside cup is dropped into place, one may see and hear a spark as the jar is discharged in the usual manner.

What is the explanation of this storage of electricity? Suppose electrons are placed on the inner coating. They will at once spread out over the whole covering because the metal foil is a good conductor. But glass is an insulator; hence the electrons cannot pass through it nor over its surface to any appreciable extent. Thus they remain spread out *at the surface of the glass* and repel the electrons of the outside metal covering, which under this force move from the outside metal and into the earth through the hand. Finally, when the jar is charged, the glass separates two rather large quantities of electricity of opposite kinds. The mutual attraction through the glass crowds more electricity on the surfaces than could be placed there without such an attraction. Thus there is a condensing effect.

Such an arrangement of metals separated by an insulator is called a *condenser.*

In the experiment with the dissectible Leyden jar, the charges which finally came together and produced the electric spark were located on the glass surfaces, and they were not conducted away because of the insulating properties of the glass. However, after the condenser was reassembled, the glass surfaces again made contact with the metal coverings, which by their conductivity aided the discharge from all parts of the glass surfaces when an opportunity was given for the electrons to migrate to the protons by way of a discharging rod.

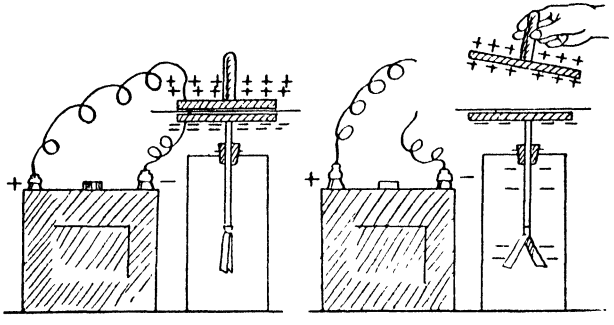


Fig. 188. The Terminals of a Battery Are Electrically Charged.

The terminals of a battery are charged with electricity. Most of us have had experience with electric batteries, and we shall be interested to see just how the electricity produced by them is related to that produced by friction. This relationship may be clearly shown by a simple experiment.

Let the knob of a gold leaf electroscope be replaced by a plate condenser in which a sheet of tissue paper serves as the insulator (Fig. 188). After the plates have been carefully discharged, let them be touched simultaneously with wires leading to the terminals of a B-battery and then be disconnected. When the upper plate is removed, the gold leaves move apart, indicating that they are charged. If the wire leading to the negative terminal of the battery

touched the bottom plate, a negative charge is found on the leaves; if that leading to the positive terminal was in contact with it, a positive charge is found on the leaves. This means that a battery is marked in terms of the kind of electricity found at its terminals. In a B-battery or a flashlight battery, zinc is the negative terminal and carbon the positive terminal.

In order to get a sufficient number of electrons on the electroscope to cause the leaves to separate, it was necessary to use the condensing action of the plate condenser. Had a wire from one of the terminals been touched directly to the electroscope, no charge would have been detected.

How May We Detect an Electric Current?

By an *electric current* we mean a flow of electric charges. The following are some of the manifestations which indicate that such a process is taking place.

Spark. A spark from a condenser, as we have already pointed out, accompanies the passing of an electric charge from one metal covering to the other. Benjamin Franklin (1706–1790) sent a kite into a thundercloud, and by drawing sparks from a key fastened to the kite string, demonstrated that lightning is an electric phenomenon. We shall now describe how a spark forms.

The terminals at either end of the spark gap are highly charged, one having a supply and the other a scarcity of electrons. A few electrically charged molecules called *ions* are present in the air at all times. A positively charged ion will be drawn rapidly toward the negative terminal, where it will take on one or more electrons. If it receives more than enough charges to make it a neutral molecule, it will be forced away toward the positive terminal. Here it will give up at least one or even more electrons. If it becomes positively charged at this contact, it will move back toward the negative terminal, where again it will take on a charge. Thus, the ion has acted as a very rapid carrier of electrons. New ions may be formed during this active process and

finally a rapid and extensive migration of electrical charges may be set up, the electrical current produced manifesting itself in the form of a spark. A further discussion of this phenomenon is found in Chapter XVII.

Heating effect. When a resistance wire (iron or nichrome) is connected to the terminals of a battery, it becomes hot and, with certain batteries and lengths of wires, even red in color. The conduction of electricity through gases, as we have shown, is accomplished by the aid of ions, but it is assumed that electrons (either free or loosely attached to the atoms) carry the current in metals by actually moving through the open spaces between the atoms. On such a trip, many encounters are made with the atoms, and much energy is lost by the electrons. Thus the kinetic energy of the atoms is increased, and in the process electric potential energy is converted into the kinetic energy of heat. The electric incandescent light and the electric stove, flatirons, curling-irons, and toasters are examples of this heating phenomenon. A discussion of these appliances is given in the next chapter.

Chemical effect. The chemical effect of electricity may be vividly shown by experiments in which a current of electricity is made to flow between two metal strips placed in a chemical solution. In such an arrangement, the metal strips are called *electrodes* and the liquid is called an *electrolyte*. If an electric lamp is included in such a circuit, a flow of electricity through the electrolyte may be detected by the glow of the lamp. If distilled water is used, the lamp remains dark; but if a little table salt or sulphuric acid is added to the water, the light glows and a current is known to be passing through the solution. Bubbles appear at the electrodes; this is evidence of a chemical action.

A modified form of this apparatus may be used to break water down into its component parts. In such a piece of apparatus, the electrodes are made of platinum and the electrolyte is a solution of water and a small amount of sulphuric acid (Fig. 189). When an electric current flows

through the apparatus, bubbles appear at the electrodes and the gas evolved may be collected in the long tubes directly above. By a study of the properties of the gases produced, one discovers that water is broken down into oxygen and hydrogen, the hydrogen appearing in an amount double that of oxygen. By checking on the battery connections, it is found that hydrogen appears at the electrode connected to the negative terminal of the battery and oxygen appears at the positive electrode.

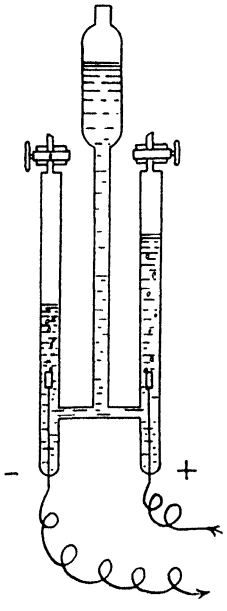


Fig. 189. Electrolysis of Water.

To explain this process, known as *electrolysis*, we assume that the sulphuric acid molecule (H_2SO_4) when in solution divides up into electrically charged parts, two positive H-ions and one negative SO_4-ion. These ions are the carriers of the electric current. When an H-ion gives up its positive charge (takes on one electron), it unites with another similarly neutralized hydrogen atom, and they form a molecule of hydrogen. When a sufficient number of hydrogen molecules are thus formed, a tiny bubble appears at the negative electrode and then rises in the tube. On the other hand, when the SO_4-ion gives up its two negative charges (electrons), it does not leave the liquid, but attacks a water molecule (it may unite first with another SO_4-ion), drives out its oxygen, and unites with the hydrogen to form a molecule of sulphuric acid (or persulphuric acid if it united with an SO_4-ion first). These replaced oxygen atoms combine as molecules, and when in sufficient numbers, they form a bubble of gas which rises in the tube at the positive electrode. By this process water is separated into its components, and the amount of sulphuric acid, although important in the process, remains constant.

Electrolysis may be illustrated further by the use of lead electrodes and a solution of sugar of lead (lead acetate) as

an electrolyte. When a current flows through the mixture, lead goes out of solution at one electrode and into solution at the other. This action is vividly portrayed when an image of the electrodes is projected upon a screen. Fern-like "trees" are seen growing at one of the electrodes.

This is the process used in electroplating. For example, in electroplating spoons, a silver bar is used as one electrode, the spoons as the other, and a solution of silver cyanide and potassium cyanide serves as the electrolyte. By the action of the electric current, silver is taken from the bar and deposited on the spoons. Articles made of iron, steel, zinc, tin, and lead may not be silvered until they have received a thin coating of copper. Chromium-plating is replacing nickel-plating because of the fine wearing qualities of chromium.

Magnetic effect. Hans Christian Oersted (1777–1851) was the first to discover that a current of electricity affects a magnetic needle. At the close of one of his lectures he said to his assistant, "Let us now once, as the battery is in activity, try to place the wire parallel to the needle." Always before, he had placed the wire at right angles to the needle. He had supposed that, if a wire carrying a current were to show a magnetic effect, it would act like a magnet placed in the direction of the wire. A wire placed at right angles to the needle never gave Oersted the effect he was looking for. But with this new position the magnet swung around, much to his astonishment and delight. "Let us now invert the direction of the current," he said to his assistant, and they saw the needle turn in the opposite direction. There seemed to be a definite relationship between the direction of the current and the direction of the needle swing.

Before electrons were known, Franklin in his definition of an electric current stated that the flow takes place from a positively charged body to a body negatively charged. The electrons are the moving charges in an electric current in a wire and actually flow in a direction opposite to this.

But rather than attempt to change a well-established rule, we shall still say that the current flows from the positive to the negative terminal of a battery, remembering that we mean the direction in which protons would move if they could.

With the direction of the electric current defined, we are ready to seek the exact relationship between the direction of the electric current and the magnetic needle swing. A compass needle points toward the north. The north-seeking end is called its *north pole.* If we place a wire *above* the needle, so that the current flows *north,* the swing of the north pole of the needle, we find, is toward the west. If the current is reversed, the north pole of the needle swings toward the east. If we now place the wire beneath the needle with the current flowing north, the north pole of the needle swings toward the east; if the current is reversed, the swing is toward the west. André Ampère (1775–1836) formulated the right-hand rule which will help us correlate these data.

Right-hand rule. This rule may be stated as follows: *Grasp the wire with the right hand so that the thumb points in the direction of the electric current; the fingers point in the direction of the swing of the north pole of a compass needle* (Fig. 190). For example, if the current flows north, the thumb points in this direction; the finger direction above the wire is towards the east; the finger direction beneath is toward the west. With the current flowing north, we found that when the wire was beneath the needle, its north pole moved toward the east; and that when the wire was above it, the north pole moved toward the west. Thus the rule agrees with our observations.

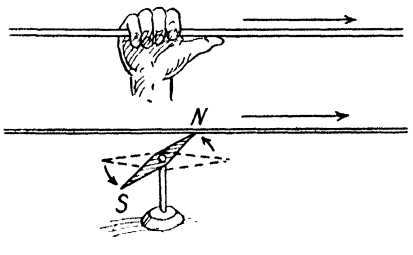


Fig. 190. Illustrating the Right-Hand Rule.

A magnetic effect is detected only when electricity is in motion. *A static electric charge does not produce magnetism.*

Magnets

Electromagnets. An electromagnet is constructed of a large number of turns of wire wrapped around a rod of iron. When a current is sent through the wire, the rod becomes strongly magnetized and attracts other pieces of iron with considerable force. The attraction becomes very much less when the current is broken. Electromagnets are often shaped so that their poles (the ends of the iron core) are separated by only a small air gap. If small nails (a pound or so) are allowed to fall upon such an electro-

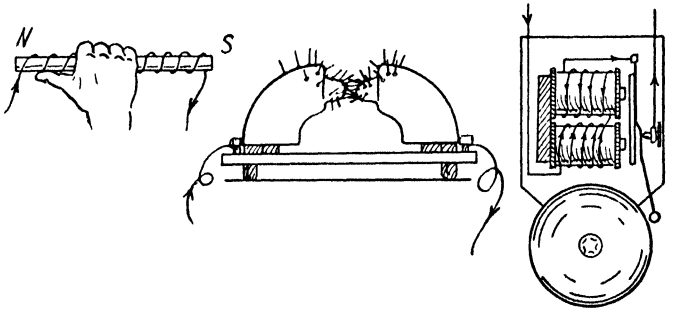


Fig. 191. The Electromagnet and Its Application.

magnet, they bristle out much as do the quills on the back of an angry porcupine. Between the poles, however, they form a direct connection across the gap. The nails, as they arrange themselves in equilibrium, give a concrete picture of the invisible magnetic field which surrounds the electromagnet. They crowd into the space between the poles, indicating that the field strength is greatest there (Fig. 191).

Permanent magnets. When a sewing needle is placed in contact with a strong electromagnet, it becomes a magnet, and when removed from the influence of the strong field, it still retains its magnetism. The needle, therefore, becomes a permanent magnet. When floated on water by means of a cork, the needle swings around and points in a northerly direction (Fig. 192). The

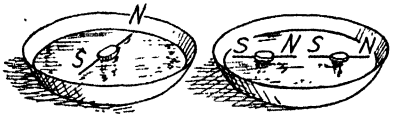


Fig. 192. The Compass Needle.

north-seeking end is called the needle's north pole; the other end is its south pole. When another such needle is placed on the water near the first one, the presence of a magnetic

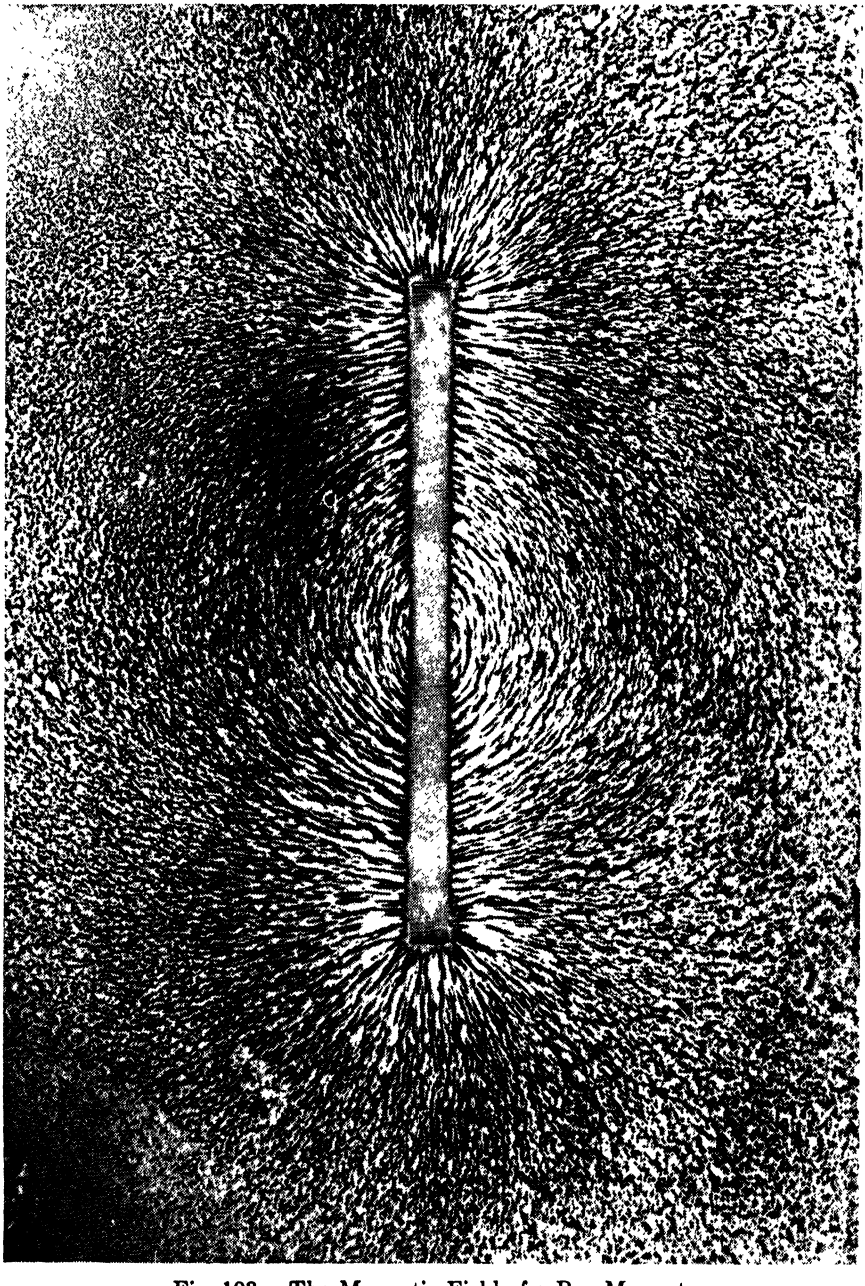


Fig. 193. The Magnetic Field of a Bar Magnet.

force is detected. The north-seeking poles repel each other, but a north pole attracts a south pole. A little experimenting will show conclusively the validity of the rule: *Like magnetic poles repel; unlike poles attract.*

The magnetic field of a permanent magnet may be vividly portrayed by placing a pane of glass over a bar magnet (a bar of magnetized steel), and then sprinkling the glass surface with iron filings (small bits of iron). The filings become tiny magnets and arrange themselves in the direction of the magnetic forces, thus giving a concrete picture of the invisible magnetic field (Fig. 193).

Magnetic compass. This is simply a permanent magnet made of hardened steel and pivoted on a support so that it may turn freely. A compass points true north only at a few positions on the earth. The *declination*, defined as the number of degrees the needle points away from the true north, changes at a given place from time to time. A few values of magnetic declination (1930) are as follows:

San Francisco	18° East	Chicago	2° East
Salt Lake City	17° East	Washington	6° West
Denver	14° East	New York	11° West
Omaha	10° East	Boston	15° West

The earth is a great magnet, with one magnetic pole in northern Canada at latitude 70°N and longitude 97°W, a location almost due north of Omaha, Nebraska. The other pole lies between Australia and the south geographical pole, not far from latitude 73°S. Fig. 194 shows diagrammatically the location of the earth's magnetic poles, and also how the north pole of a needle dips down or up as it is moved north or south of the equator. In 1831 Sir James Ross found a place in the extreme north of Canada where the dip is 89°59′. This region marks the location given above of a magnetic pole of the earth.

In the great majority of cases, ships are navigated by the aid of the magnetic compass. Therefore, it is very important that the magnetic declination be known on all

oceans, seas, and lakes. In 1909 the nonmagnetic yacht, Carnegie, was built by the Department of Terrestrial Magnetism of the Carnegie Institution of Washington. She was built of white oak and Oregon pine, and her fittings, anchors, engine, propeller-shaft, and all metal parts were made from copper or composition-bronze. The vessel made seven voyages in the quest of more and better information concerning the magnetic and electric fields of the earth. The seventh and last cruise ended abruptly and tragically in November, 1933, when the vessel was destroyed by an explosion while in Apia Harbor. Captain James Percy Ault and a cabin boy lost their lives.

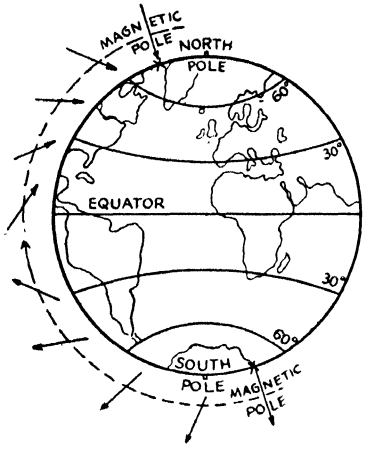


Fig. 194. The Earth, a Great Magnet.

Magnetism is always associated with *moving electricity.* The theory of its cause will be considered briefly in Chapter XVII.

Questions and Problems

1. List the manifestations of electricity as it is used in the home.

2. After one has moved briskly over a carpeted floor, an electric spark is often produced as an electric light switch is touched. Explain.

3. List at least two devices in the home which make use of the magnetic effect of an electric current.

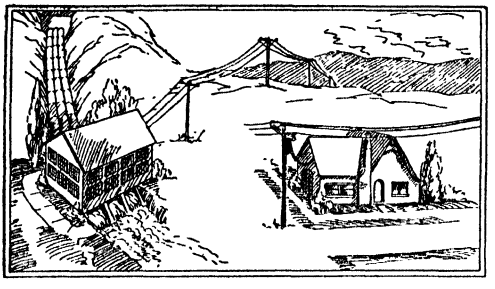
4. A wire carrying a current of electricity is placed over a magnetic needle. The current flows south. In which direction does the north pole of the needle swing?

Suggested Readings

(1) Dietz, D., *The Story of Science*, Sears Publishing Company, Inc., New York, 1931, Part III.

(2) Gibson, C. R., *Our Good Slave, Electricity*, J. B. Lippincott Company, Philadelphia, 1915, Chaps. I, II, and III.

(3) Paul, J. H., *The Last Cruise of the Carnegie*, The Williams and Wilkins Company, Baltimore, 1932.
(4) Tyndall, J., *Lessons in Electricity*, D. Appleton-Century Company, Inc., New York, 1877, Sections I–XXXII.



CHAPTER XVI

How Electricity Is Generated

Energy may be transferred to electrical corpuscles by various methods. In each case we say that electricity is *generated.* We shall now consider the methods used in this energizing process.

Generation by Intimate Contact

We discovered in the last chapter that electrons are removed from a cat's fur and accumulated on a hard rubber rod when these insulators are brought into very intimate contact. In a similar manner, electrons are removed from glass and accumulated upon silk. When an iron rod is rubbed with fur, no charge is detected on the rod. But when this experiment is repeated with the rod well insulated from the hand, the rod is found to be charged. This means that the uninsulated rod was not able to retain its charge when generated because of the excellent electrical connection to the earth through the rod and the hand. It is now obvious why electricity was first discovered when amber was rubbed with silk. Both these materials are good insulators, and the generated charge is not easily conducted away. After experimentation with very many substances, it is found that all bodies become either negatively or positively charged when brought into very intimate contact.

This means that electrons move from one substance to another when the two are brought into very intimate contact. Thus, certain bodies seem to have a greater "affinity" for electrons than do others. We call this

"affinity" which causes the electron migration the *contact electromotive force.* Energy is stored when electrons are "robbed" by the action of the contact electromotive force, and further energy is stored when the person performing the rubbing moves the oppositely charged substances out of immediate contact.

Generation by Induction

We remember that energy is stored when forces lift weights against the pull of gravity. In a similar manner, one might expect energy to be stored if electrons are moved away from protons by the action of an electrical field. The following experiment illustrates such a process:

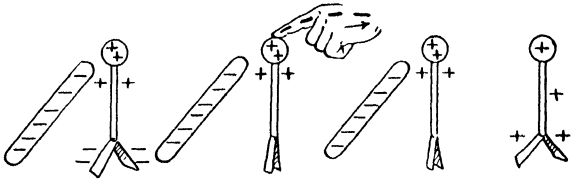


Fig. 195. Charging by Induction.

Let a negatively charged rubber rod be brought *near* (not touching) the knob of a gold leaf electroscope. The leaves move apart. Without removing the charged rod from its position, let the knob be touched. The leaves collapse. Now if the finger is removed first and then the charged rod is taken away, the leaves separate and show a permanent charge. If a positively charged glass rod is brought near, the leaves separate still farther, and if a negatively charged rubber rod is then brought near, they collapse. When we remember that like charges repel and unlike charges attract, and that a widening of the leaves means an increase of charge while a collapsing means a decrease of charge, we have definite evidence that a positive charge is left on the leaves by the charging process. Thus, by the influence of a negative charge, the electroscope has been charged positively. We call this process *charging by induction.*

In terms of electrons, the process of charging by induction may be interpreted as follows: When the negative charge is brought near the electroscope, electrons are repelled into the leaves. When the hand touches the knob, it conducts enough of these corpuscles to earth to make the leaves neutral. The hand is removed, and again the electroscope rod is insulated from the ground. When the inducing charge is then taken away, the force which repelled electrons away from the knob is removed also, and a redistribution of the electrons remaining in the rod takes place. *Electrons go from the leaves to the knob* to compensate partly for the electrons which were forced away and out into the earth by the action of the inducing charge. Thus, the leaves become lacking in electrons and are positively charged (Fig. 195).

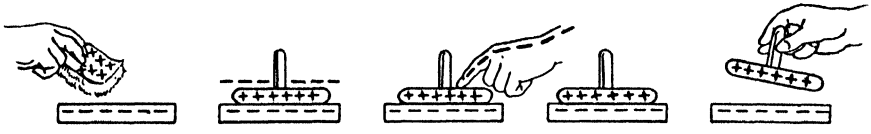


Fig. 196. The Electrophorus.

The electrophorus. This instrument was first invented by Alessandro Volta (1745–1827). It consists essentially of a hard rubber plate and a metal disc with an insulating handle (Fig. 196). When the rubber plate is struck vigorously with a cat's fur, the plate becomes negatively charged. The metal disc is then placed on top of the charged rubber. The surfaces are uneven, contact is made at a few points only, and very little negative charge is taken on by the metal. Under the action of the negatively charged rubber plate, the electrons of the metal are forced into the top part of the disc. If the metal is touched, electrons pass to the earth, leaving protons satisfied under the influence of the charged rubber. When the metal disc is lifted by means of the insulating handle, work is done against the electrical attraction and potential energy is stored. As a finger is now brought near the metal disc, a

spark jumps the air gap, and the energy stored changes into heat, light, and sound.

One may continue thus to energize electricity without using up the charge originally produced on the rubber plate. The energy needed for the generating process is furnished by the person who pulls the metal up and away from the rubber. The electrophorus, then, is a simple electric machine which when operated is capable of generating electricity by induction. The so-called *influence* or *static machine* also makes use of this process and is designed so that the generation may be made to go forward simply by the turning of a crank. Such machines give a spectacular display of electrical sparks. They appeared early in the development of electrical machinery but are not used today as commercial generators.

Generation by Chemical Action

By the use of a condenser and a gold leaf electroscope we demonstrated, in the last chapter, that the terminals of a battery are electrically charged. We might continue to use this instrument here, but another type, called the *galvanometer*, which is constructed to make use of the magnetic effect of an electric current, we shall find a more convenient device to demonstrate that electricity is generated chemically. A galvanometer consists essentially of a coil of wire mounted in a permanent magnetic field. The coil is delicately pivoted so that it may rotate against the action of a coiled spring. When an electric current flows through the coil, a magnetic field is produced; and this field, acting in conjunction with the magnetic field of the instrument, causes the coil to turn and a pointer to move over a scale (Fig. 197).

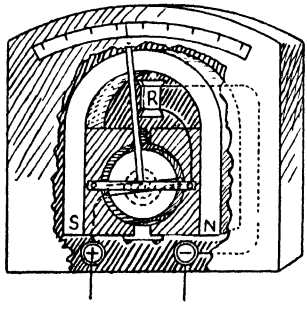


Fig. 197. A Galvanometer.

To illustrate how electricity is generated chemically, the following experiment may be performed: Let a strip of zinc and a strip of copper be dipped in a solution composed of water and a small amount of sulphuric acid, and let them be connected in series with a high-resistance galvanometer (voltmeter) as shown in Fig. 198. A deflection of the pointer of the instrument indicates the presence of an electric current flowing from the copper strip to the zinc strip. (Note: The terminals of the instrument will probably be marked "plus" and "minus," meaning that if the pointer is to turn in the positive direction of the scale, the minus terminal must be connected to the negatively charged metal strip of the battery and the plus terminal to the positively charged strip. This enables one to determine the direction of the current. Of course, the gold leaf electroscope and condenser could be used as a final check.)

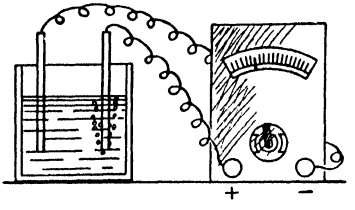


Fig. 198. A Simple Electric Cell.

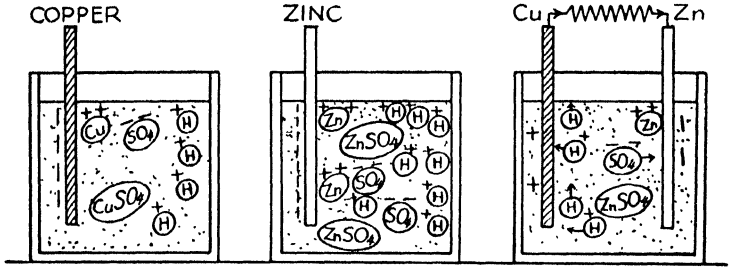


Fig. 199. The Action of a Simple Cell.

The description of the chemical action which generates the electric current is as follows: Sulphuric acid when placed in a water solution separates into positive H-ions and negative SO_4-ions. When a copper strip *alone* is dipped into the liquid, positive copper ions go into solution and unite with the negative SO_4-ions, forming copper sulphate ($CuSO_4$). Each atom of copper leaves behind two electrons as it enters the solution not as a neutral atom but as a charged ion. The copper ions continue to go into solution

until the "solution pressure" which causes them to leave the metal is just balanced by the attractive force of the negatively charged metal strip. When the zinc *alone* is dipped into the solution, the process is very similar, except that the zinc becomes more negatively charged than the copper (Fig. 199).

If copper and zinc strips are dipped into the solution simultaneously, the zinc, because of its greater "solution pressure," is in a position to furnish more positive ions to the solution than copper. This great "army" of zinc ions repel the hydrogen ions of the sulphuric acid toward the copper plate, and the hydrogen ions, in turn, force the copper ions out of solution. But the predominance of zinc ions is such that all the copper ions are forced out of solution and even the positive hydrogen ions appear upon the copper plate. Soon equilibrium is reached with the copper plate positively charged and the zinc plate negatively charged. When a wire is connected between the terminals, a current flows, the plates tend to discharge, and the chemical action proceeds until the chemicals are used up or the circuit is broken. Hydrogen bubbles may be seen forming on the copper plate as the cell delivers an electric current. Whenever electricity is generated by chemical action, one of the plates is eaten up and upon the other something is deposited.

A simple cell has two main defects: (1) Hydrogen bubbles accumulate on the copper and change it in effect to a hydrogen plate, decreasing the ability of the cell to furnish current; (2) impurities in the zinc permit tiny electric currents to be set up between particles of these and the main body of the metal. This local electrical action, in terms of zinc consumption, is equivalent to a closed circuit between the copper and zinc plates. Therefore, such a cell deteriorates when not in use.

Cells formed by the use of dissimilar metals placed in an acid solution are often called Galvanic or Voltaic cells in honor of Aloisio Galvani (1727–1798) and Alessandro Volta (1745–1827), who were pioneers in this field.

The dry cell. The dry cell is a commercial product. It is the unit out of which flashlight batteries and radio B-batteries are built. A can made from zinc serves as the negative terminal, and a centrally located carbon post acts as the positive terminal. Next to the zinc is a layer of blotting paper, and between this and the carbon post is a pasty mixture of sal ammoniac, manganese dioxide, and powdered carbon. If it were not for the manganese dioxide, ammonia gas would collect at the positive terminal. Because of local action at the zinc surface, such cells may deteriorate on the vender's shelves; hence, they should be tested before purchase.

The lead storage battery. The lead storage battery is composed of two types of plates suspended in a strong solution of sulphuric acid. The positive plate is composed of peroxide of lead (PbO_2) and the negative plate is spongy lead. The chemical reactions during charging and discharging are as follows:

Charging:

At positive plate, $PbSO_4 + SO_4 + 2H_2O = 2H_2SO_4 + PbO_2$

At negative plate, $PbSO_4 + H_2 = Pb + H_2SO_4$

Discharging:

At positive plate, $PbO_2 + H_2SO_4 + H_2 = PbSO_4 + 2H_2O$

At negative plate, $Pb + SO_4 = PbSO_4$

During discharge, lead sulphate ($PbSO_4$) is formed at both plates, sulphuric acid is used up, water is produced, and the density of the solution decreases, approaching that of water. During charging, however, the lead sulphate is changed to peroxide of lead at the positive terminal and to spongy lead at the negative terminal, sulphuric acid is formed in both reactions, and the density of the solution increases. One may determine the extent of the battery's charge, therefore, by reading the density of the electrolyte. This is the test made at a

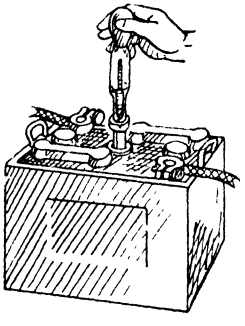


Fig. 200. Testing a Storage Battery.

service station when you have your automobile battery "checked." Distilled water is added if the level of the electrolyte is low (Fig. 200).

During charging, electrical energy is sent into the cell and there converted into chemical energy. Later this chemical energy is released as electrical energy. Thus, the storage battery does not store electrical energy; it stores chemical energy.

Generation by the Action of Light—Photoelectric Effect

Let a piece of zinc, with a surface which has been *freshly* sandpapered, be mounted on a gold leaf electroscope and be given a negative charge. The zinc surface, then, has an extra supply of electrons and the gold leaves are spread apart. Now let the light from an open arc fall upon the zinc. The gold leaves slowly collapse; in a few seconds the electroscope is completely discharged, and electrons must have left the zinc surface. Let the electroscope again be charged negatively, but this time place a pane of window glass between the electroscope and the arc. The leaves fall very slowly. Electrons are ejected, but at a much reduced rate. Thus, it seems that electrons are ejected from zinc by the action of the invisible ultraviolet rays, which do not readily pass through ordinary glass. Protons are not ejected by the action of light, because when the electroscope is charged positively and light falls upon the zinc surface, the leaves do not collapse and no discharge takes place. The ejection of electrons by the action of light is called the *photoelectric effect.*

A photoelectric cell, as manufactured commercially, has much the same appearance as an electric light globe or radio tube. On the inside of the evacuated glass bulb, a light-sensitive surface (sodium, potassium, cesium, or specially contaminated surfaces of these pure metals) is deposited in such a manner as to cover all the inside surface except a windowlike opening through which light may enter. A wire ring, or a grid of wires, made of nonsensitive material,

is centrally located in the tube, as shown in Fig. 201. Under the action of light, electrons are ejected from the sensitive surface, and the space near by becomes literally an atmosphere of electrons. When a battery is placed between the sensitive surface and the nonsensitive ring, with the positive terminal connected to the ring, electrons are swept away from the sensitive surface to the positive ring as soon as they are ejected by the light. Under suitable conditions, the current of electrons is directly proportional to the intensity of the light. Thus, a variation in light intensity may be converted into an identical electric current variation. This principle is used in talking pictures, television, and in sending photographs by telegraph or radio.

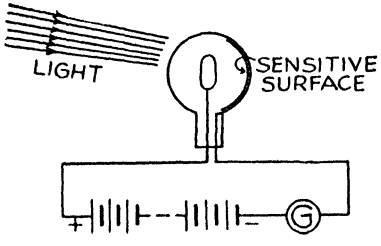


Fig. 201. A Photoelectric Cell.

Generation by the Action of Heat—Thermoelectricity

Let the ends of an iron wire be connected to two copper wires (as illustrated in Fig. 202), and let the circuit then be completed through a sensitive galvanometer. When one copper-iron junction is heated and the other is kept at a low temperature, a current flows through the galvanometer. This is a *thermoelectric* current. Such electric currents are always very small, but thermocouples built upon this principle are widely used for high temperature measurements.

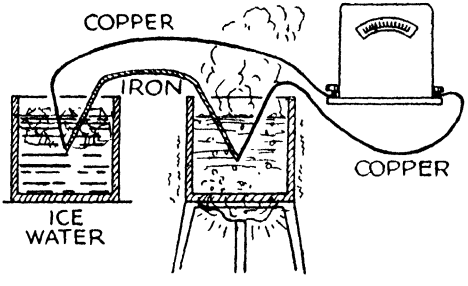


Fig. 202. Thermoelectric Effect.

By making a very small couple of bismuth and platinum and enclosing it in a vacuum, it is possible to compare the heat radiated from stars. Nicholson and Petit, of Mount

Wilson Observatory, make instruments so sensitive that by the use of the 100-inch Hooker telescope the heat from a candle 120 miles away may be detected.

Thermoelectricity is explained when it is assumed that heat energy is converted into electrical energy by a sort of diffusion of electrons from one metal to the other across the junction.

Generation by Cutting Magnetic Lines of Force

Let a battery be connected to a large electromagnet, and let the ends of a long copper wire be fastened to the terminals of a galvanometer. Then if one quickly forces the wire down through the magnetic field between the poles of the strong magnet, the galvanometer shows a current reading so long as the wire is in *motion* (Fig. 203). If one now cuts the wire up across the field, the needle moves in the opposite direction, indicating that the current has reversed. When one moves the wire slowly through the field, the deflection is not as great as when one moves it rapidly. Again, if one moves it along and not across the field, no deflection results.

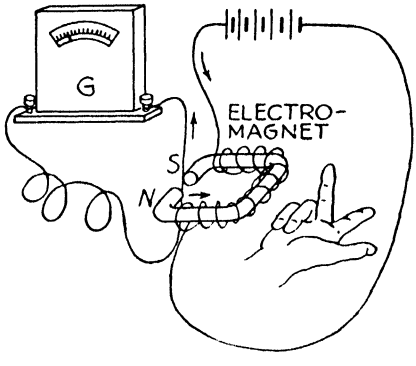


Fig. 203. Electromagnetic Induction.

Thus, we conclude that if a conductor moves across a magnetic field, electricity is generated, the size of the current depending on the strength of the field cut and the speed of the motion through the field.

Next let the long wire be wrapped a few times around the electromagnet. When the current flowing in the electromagnet is suddenly stopped, the needle of the galvanometer moves. As the current is again started and the magnetic field is being once more established, the needle moves in the opposite direction. This must mean that the effect obtained from a wire at rest in a changing magnetic field is

equivalent to that obtained from a wire cutting through a steady field.

Michael Faraday (1791–1867) discovered this very important phenomenon, known as *electromagnetic induction*, in 1831. And today practically all electricity used in this age of machinery is generated by the forcing of conductors through a magnetic field or by the forcing of a magnetic field through conductors.

The generator rule. In electromagnetic induction, the direction of motion of the conductor, the direction of flow of the current, and the direction of the magnetic field are related in a simple manner. The student may remember and apply the relationship by making use of the right-hand rule: Place the thumb and first and second fingers of the right hand at right angles to each other (Fig. 203). Point the first finger in the direction of the magnetic field (from the north to the south pole) and the thumb in the direction of motion of the conductor; then the second finger will point in the direction of the flow of current.

Source of the energy transferred to electricity in the process of electromagnetic induction. We have already shown that a current flowing in a wire produces a magnetic effect. Therefore, just as soon as a current is established in a wire as it moves through a magnetic field, this generated current produces a field of its own. The two fields add up, and the net result is a force which opposes the motion of the wire. This means that work must be done to force a closed wire through a magnetic field. The energy expended appears loaded upon the electrons of the electric current. In everyday practice, the mechanical energy of falling water or expanding steam is used to force wires through magnetic fields, and the mechanical energy expended is changed to electrical energy.

The generator. In the practical utilization of electromagnetic induction, wires are wound on an armature (an iron cylinder), and this is rotated in a magnetic field. The direction of the current generated reverses as the wires

cease cutting through the field in one direction and begin cutting in the opposite direction. Thus, a current which changes its direction periodically (an alternating current) is drawn from such a generator. An alternating current is diagrammatically represented in Fig. 204. The electric current commonly used in the household makes 60 alternations per second. Details of generator construction cannot be included in this survey. Students may easily find literature giving the details desired.

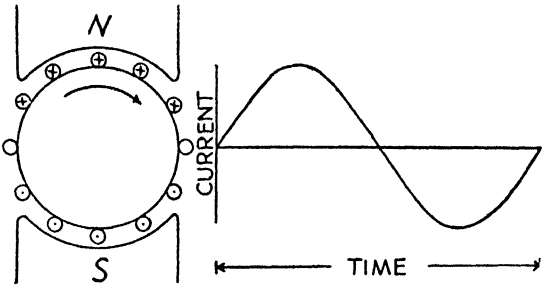


Fig. 204. Illustrating an Alternating Electric Current.

The telephone transmitter. The electrodynamic transmitter or microphone, the kind used in radio broadcasting stations, where high speech quality is required, is con-

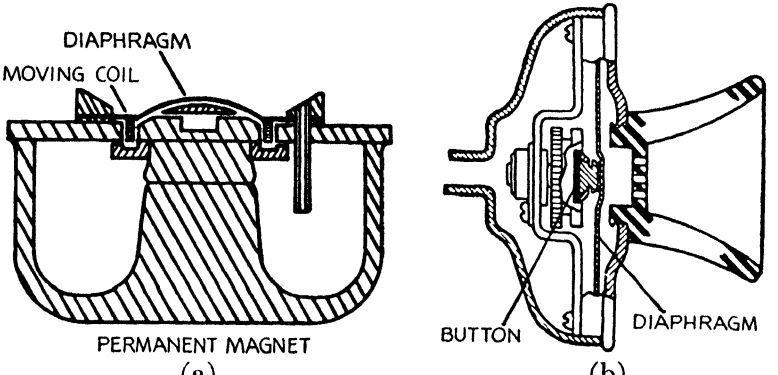


Fig. 205. (*a*) A Moving-Coil Transmitter. (*After Wente and Thuras.*) (*b*) A Carbon Microphone. (*Courtesy of Bell Telephone Laboratories.*)

structed on the principle of the electrical generator. The moving diaphragm causes a tiny coil of wire to move back and forth in a strong magnetic field, and an electric current, with all the vibration characteristics of the sound waves, is generated. Such a high-quality moving-coil transmitter,

designed by Wente and Thuras, of Bell Telephone Laboratories, is illustrated in Fig. 205.

The microphone used in the common telephone depends for its action upon tiny pieces of carbon loosely packed in a small box located just behind the transmitter diaphragm. The vibrating diaphragm changes the packing of the particles, and the electric current flowing through the assemblage varies as the packing changes. The necessary pulsating currents needed to carry the characteristics of sound are thus established.

Changing Electrical Energy into Mechanical Energy

The electric motor. Let a wire be placed in a magnetic field as illustrated in Fig. 206. When the circuit is closed, the *wire moves* in a direction perpendicular to the magnetic field. When the current is reversed, the motion takes place in the opposite direction. It has been found that the action, when carefully tested, may be summarized in the left-hand rule: Place the thumb, first and second fingers of the left hand at right angles to each other. Point the first finger in the direction of the magnetic field and the second finger in the direction of the current; then the thumb points in the direction of the motion of the wire. All electric motors are built on this principle.

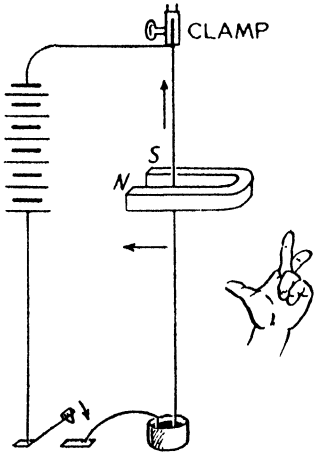


Fig. 206. Motor Effect.

In a generator, electricity is generated by the use of a mechanical means to force wires through a magnetic field; in a motor, an electric current is sent through a wire in a magnetic field, and mechanical forces are produced. With these principles once established, man soon learned the art of transferring mechanical energy into electrical energy and then back again, and our whole scheme of energy utilization was extended and perfected.

The telephone receiver. The ordinary telephone receiver is composed of a diaphragm located near but not touching a permanent magnet. The field of the magnet is modified, when the receiver is in use, by a variable electric current which flows through a coil of wires surrounding the magnet. If the pulsating current, set up by sound waves impinging on a transmitter or microphone, is sent through the wire coils of the receiver, the magnetic field is modified; forces on the diaphragm are periodically changed, so that it is caused to vibrate, and sound waves are again established. The essential parts of the common telephone receiver may be observed when the cap is unscrewed and the metal diaphragm removed.

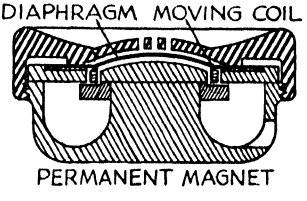


Fig. 207. High-Quality Moving-Coil Head Receiver. (*After Wente and Thuras. Courtesy of Bell Telephone Laboratories.*)

A high-quality receiver which makes use of the principle of the motor, that a wire in which current is flowing when placed in a magnetic field receives a mechanical force, is illustrated in Fig. 207. Such a receiver, when constructed on a larger scale, may be used as the active element in a loudspeaker.

Electrical Measurements

What electrical quantities should one try to measure? Certainly one should have a measure of the electric current. It is also very important to know how much energy one should expect to receive when a certain current is utilized.

The measure of the electric current—The ampere. The simplest current one can imagine is that produced at a point in a conductor by one electron passing per second. This size of current is found to be altogether too small for practical purposes, in fact, about a million million million times too small. The practical unit of current, the *ampere*, is standardized as the current which will deposit 0.00111800 grams of silver out of a solution of

silver nitrate in one second. When such a current flows in a wire, 6.29 million million million electrons pass per second. This number of electrons, when lumped together, constitutes the charge called the *coulomb*. Hence, the ampere is a current of one coulomb per second.

The *ammeter* is an instrument used to measure electric current. It is essentially a low-resistance galvanometer, and is always placed so that the current to be measured goes through it (Fig. 208).

The measure of the electrical potential—The volt. When electricity is generated, energy is stored on the electrons. Can we measure this energy? What we wish is usually the difference in energy content between different points in the electric circuit. For example, we wish the difference in energy between the terminals of a battery or across the filament of an incandescent lamp. If the current measured in amperes is known, then the most useful quantity would be the energy difference per coulomb between the two points. We need a new unit to measure this potential energy difference per coulomb, or simply the *difference of potential*. The *volt* has been selected. The difference of potential is one volt when the energy drop per coulomb is one joule. Thus, when the voltage (difference of potential) of a house-lighting line is 110 volts, the energy on a coulomb is 110 joules more on one side of the circuit than on the other. In other words, this amount of energy is taken from each coulomb which passes through a lamp, a sweeper, a stove, or a pair of curling irons, if operating on this circuit.

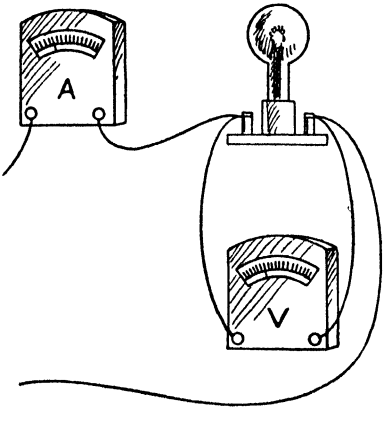


Fig. 208. Measurement of Amperes and Volts.

The voltmeter is used to measure difference of potential or voltage. It is essentially a high-resistance galvanometer.

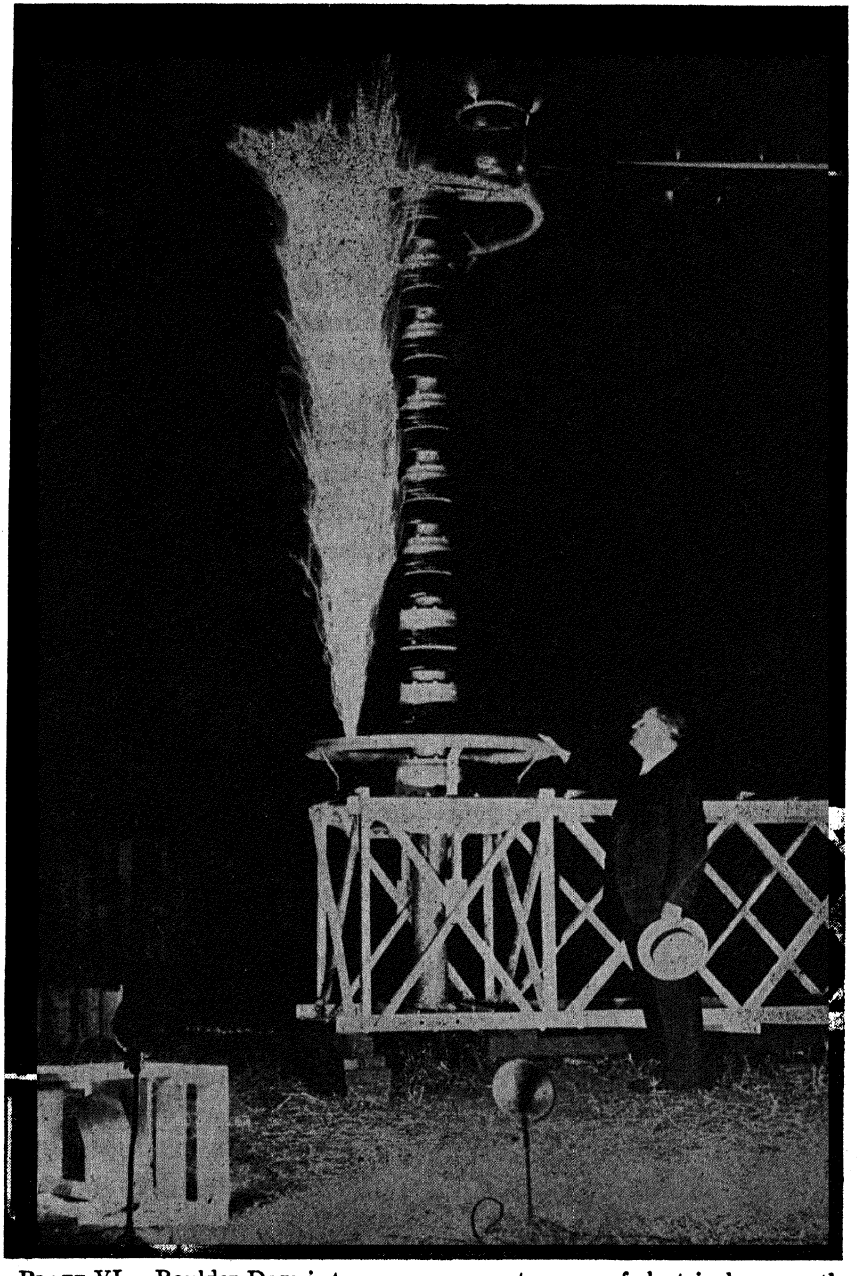


PLATE VI. Boulder Dam is to serve as a great source of electrical power, the energy of the escaping water being transferred to electricity. The generated power will be transmitted at high voltage to great distances—as far as Los Angeles, California. In order to turn the electricity on and off the power lines, specially designed switches will be used. Here you see a photograph of an electrical discharge produced when 830,000 volts were applied across a switch designed to operate normally on 287,000 volts. It required this very high potential to produce the arc-over, thus establishing the fact that the switch is built to operate with safety on the much lower potential. Such switches are to be used at the Boulder Dam Switching Station. (*Photograph through the courtesy of Prof. Joseph S. Carroll, Stanford University, California.*)

Its needle deflection is proportional to the current flowing through it, but the size of this current depends upon the difference of potential across the high resistance. Hence, the deflection, if properly designated, measures the voltage directly. A voltmeter is always placed across the points between which the voltage is desired (Fig. 208).

The rate at which electrical energy is used—The watt. Suppose that a current of one ampere is flowing through a lighting fixture on a 110-volt lighting circuit. Each second, a coulomb goes through an energy drop of 110 joules. This rate of doing work is 110 watts, since a joule per second is a *watt*. With two amperes flowing, the energy consumed per second is doubled, and the power used is 220 watts. In general, we may write,

$$\text{Watts} = \text{Volts} \times \text{Amperes}.$$

A kilowatt is 1,000 watts. If energy at this rate is fed into the appliances of a home for *one hour*, the amount of energy used is one kilowatt-hour. When one settles his electric bill, he pays for the kilowatt-hours used.

Electrical resistance—The ohm. When a current flows in a metal, the electrons are assumed to move through the open spaces between the atoms. The opposition encountered by the current is called *electrical resistance*. A unit of resistance is defined as follows: When a current of one *ampere* flows in a wire, the resistance between two points is one *ohm* if the difference of potential between these points is one *volt*.

Ohm's law. George Simon Ohm (1789–1854), after careful experimentation, formulated this law which bears his name:

$$\text{Amperes} = \frac{\text{Volts}}{\text{Ohms}}.$$

This means, for example, that on a 110-volt circuit a resistance of 110 ohms will allow a current of 1 ampere to pass, a resistance of 55 ohms will allow 2 amperes to pass,

and a resistance of 220 ohms will allow ½ ampere to pass. On a 220-volt circuit, the current passing through each resistance would be doubled. On a constant-potential line (the type used in the home), the current is controlled by introducing resistances such as iron or chromel wire, carbon plates, a salt solution, or electric lamps. Such devices are called *rheostats.*

Copper wire is the conductor generally used to carry electricity from one place to another. Although their resistance is low, small wires with heavy currents may produce sufficient heat to set a building on fire. Conductors of sufficient size to carry the current load should be installed. The following data give the safe carrying capacity of a few sizes of wire:

Brown and Sharp gauge.....	16	14	12	10	8	6	4
Amperes.....	6	12	17	25	33	45	63

Electrical Heating and Lighting

The electric circuit. A complete circuit of conducting material is required if an electric current is expected. If sufficient voltage is present to force electricity through the difficult path of a very small electric light filament, and if by accident a large copper wire is placed between the terminals of the lamp, the electricity is given too easy a path; a so-called *short circuit* is established, a very large current is drawn from the power line, wires are overheated, and a house may be set on fire unless protected by fuses. A fuse plug is made of a metal which melts and breaks the circuit before damage is done (Fig. 209).

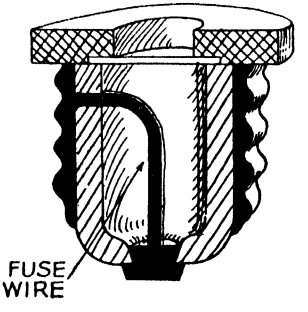


Fig. 209. A Fuse.

Here is a set of simple rules: *In making electrical connections, be sure to make a complete circuit; keep out short circuits; and as a precaution, protect the line with a fuse. Never*

replace a fuse until the cause of its burning out is discovered and corrected.

Heating appliances. Heating appliances, such as stoves, flatirons, toasters, and so forth, are rated in terms of the power consumed. For example, a 6.5-pound flatiron uses 525 watts; a toaster, 400 watts; an electric range oven, 2,500 watts. The heating elements of an electric range are made of chromel wire, an alloy consisting of 80 parts nickel and 20 parts chromium. A nickel-iron-chrome alloy consisting of 63 per cent nickel, 25 per cent iron, and 12 per cent chromium is often used in flatirons, toasters, and similar devices. These metals may be heated to

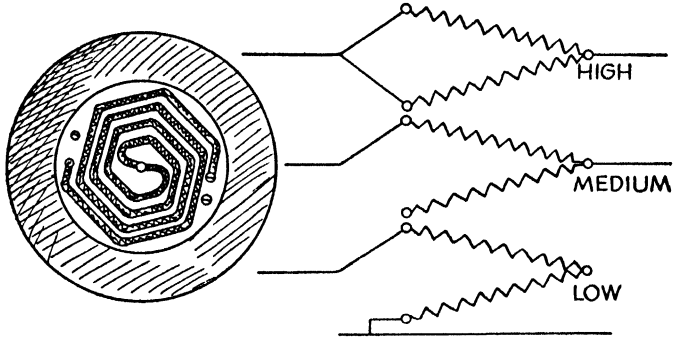


Fig. 210. Three "Heats" in an Electric Range.

redness without oxidation. In time, however, the heating elements may need to be replaced. In electric stoves each plate has three "heats": high, medium, and low. At high, the current has a double path formed by the two coils making up the plate (parallel connection); at medium, only one of the coils is used; at low, the current is made to pass through the coils in series (Fig. 210).

Electric lamps. Lamps are rated according to the power consumed, as, 40-watt, 100-watt, 500-watt. A 100-watt lamp costs 2.5 times as much to operate as a 40-watt lamp. The carbon filament lamp was invented by Thomas A. Edison (1847–1931). His first commercial product had an efficiency of 0.088 candle power per watt. This first lamp has been followed by the metallized carbon filament

lamp, the tantalum lamp, the tungsten lamp, and the gas-filled lamp. The modern 100-watt gas-filled tungsten lamp has an efficiency of 1.06 candle power per watt, an efficiency over ten times that of the first carbon lamp produced by Edison.

The efficiency of modern lamps increases with size. The larger lamps can be raised to higher temperatures, and the radiation in the visible part of the spectrum increases with the temperature. This is clearly shown by the following data.

Type of Lamp	*Watts*	*Candle Power per Watt*	*Temperature in Kelvin Degrees*
Vacuum, inside frost....	15	.75	2470
" " "	25	.78	2505
" " "	40	.80	2535
Gas-filled, inside frost...	50	.81	2670
" " " ...	60	.90	2695
" " " ...	100	1.06	2755
Gas-filled, clear.....	200	1.29	2865
" "	500	1.52	2965
Stereopticon.....	1,000	1.92	3185
Movie.....	10,000	2.47	3350

These data show clearly that by using two 50-watt lamps, only $\frac{8}{10}$ as much illumination is obtained for the same cost as if one 100-watt lamp were used. However, the large units bring the danger of glare unless properly shaded.

We have shown how great groups of electrons may be detected, and how they may be used as carriers of energy. In the next chapter we shall outline the experiments which have led to the conclusion that matter itself is built out of electrical corpuscles.

Questions and Problems

1. Why is it good practice to remove the batteries from a flashlight before storing it?

2. What is the benefit of having a storage battery checked at a service station?

3. Describe a practical use of the photoelectric cell.

4. The Generator Rule: Place the thumb and the first and second fingers of the .... at right angles to each other. Point

the first finger in the direction of the .... and the thumb in the direction of the .... ; then the second finger will point in the direction of the ....

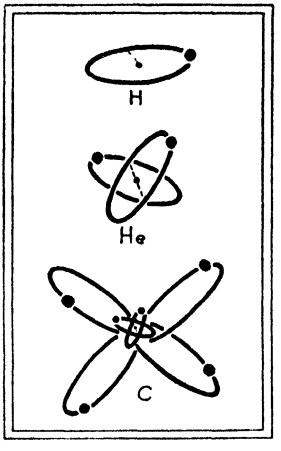
5. The Motor Rule: Place the thumb and the first and second fingers of the .... at right angles to each other. Point the first finger in the direction of the .... and the second in the direction of the ..... ; then the thumb will point in the direction of the motion ....

6. Explain how the thermoelectric effect of a current may be used to measure the temperature of furnaces.

7. If electric power costs ten cents per kilowatt-hour, what will it cost to burn five 100-watt lamps four hours per day for 30 days?

Suggested Readings

(1) Black, N. H., and Davis, H. N., *New Practical Physics*, The Macmillan Company, New York, 1929, Chaps. XVII–XXIV.
(2) Farnum, D. T., *et al.*, *Profitable Science in Industry*, The Macmillan Company, New York, 1925, Chaps. II and III.
(3) Gibson, C. R., *Our Good Slave, Electricity*, J. B. Lippincott Company, Philadelphia, 1915, Chaps. IV–XV.
(4) Harris, F. S., and Butt, N. I., *Scientific Research and Human Welfare*, The Macmillan Company, New York, 1924, Part II.



CHAPTER XVII

Inside the Atom

During the past thirty years, we have learned to "look" inside the atom and discover its structure. This is not done by the use of the microscope. The eye cannot penetrate into the minute mechanism; but man has learned to see the invisible by interpreting the messages which come out of these interesting structures. By perfecting new instruments of measurement, by developing new laboratory technique, and by working out careful inferences, the scientist not only has discovered the building materials of atoms, but has made a rather good guess as to atomic architectural design. It is impossible in this short chapter to follow through all the experiments which have been performed, all the hypotheses which have been built up, torn down, rejected, and accepted during the past thirty years as modern physics has been in the making, but we shall describe a few of the experiments out of which the modern picture of atomic structure has evolved.

Electric Discharge in Partial Vacua

In an earlier chapter, we discussed the disruptive spark discharge formed at atmospheric pressure. We explained the passage of electricity by supposing that ions of air actually carry the charges across the gap.

When electrodes are placed inside a glass tube, and the pressure is reduced to about one-half to one-third of an atmosphere, a streamer-like discharge fills the tube. When the pressure is reduced to about one-thousandth of an

atmosphere (0.76 mm. Hg), the so-called *Faraday dark space* appears *near* the cathode (the terminal charged with electrons). But *at* the cathode a narrow, grayish region, called the *negative glow*, is observed. With further reduction of pressure, a second dark space, the *Crookes dark space*, appears between the negative glow and the cathode; the Faraday dark space and the negative glow move toward the anode; and a new glow, known as the *cathode glow*, appears at the cathode. With a pressure of one-millionth of an atmosphere or less (0.00076 mm. Hg), the Crookes dark space reaches from the cathode to the anode. An X-ray tube has been produced.

Cathode Rays

Since the glow due to the residual gas has disappeared, what carries the electricity across the tube? Positive ions may still play a role in the process, but probably a stream of electrons are shot from cathode to anode. Something invisible is given out at the cathode which causes a brilliant fluorescence in the walls of the tube opposite the cathode, produces a red-hot heating of a metal placed in its path, and casts sharp shadows of the metal obstacle.

When a mica slit is placed in front of the cathode and a zinc sulphide screen is located at the opposite end of the tube, a fluorescent image of the slit is seen. The invisible stream passes through the slit and strikes the screen. When a magnet is brought near the tube, the image of the slit moves exactly as if an electric current consisting of a stream of electrons moved from cathode to anode. When the poles of the magnet are reversed, the direction of the shift is also reversed. In 1895 Sir J. J. Thomson (1856–) proved that this stream, called

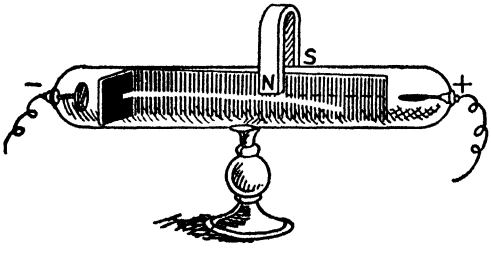


Fig. 211. Bending Cathode Rays with a Magnetic Field.

cathode rays, is also deflected by an electric charge exactly as if it were made up of moving negative charges. Today cathode rays are universally recognized as a stream of electrons shot out from the cathode in directions normal to the surface (Fig. 211).

Using the deflections produced by a magnetic field and by an electric charge, Thomson was able to measure the speed of the cathode ray electrons and the ratio of the electron charge to its mass. Velocities above 100,000 miles per second were measured, and the mass of the electron was found to be $\frac{1}{1840}$ of the mass of the hydrogen atom.

Measuring the Charge on the Electron

In the now famous oil drop experiment, apparatus is arranged as shown in Fig. 212. A small oil drop floats in the air between the condenser plates; the air is ionized by X-rays, and when the field is applied, the oil drop rises against gravity or falls with gravity with a greater than usual speed, indicating that it has received a charge from the air ions. By the proper reversal of the electric field between the plates, a drop may be kept for hours in the field of view of a telescope. Often the speed of rise or fall suddenly changes—an indication that a charge has been added or subtracted. Especially does this take place when the X-rays are on. A careful study of oil-drop speeds reveals that the charges captured by the drop are always a small whole number times a definite charge. It is easy to determine whether one, two, three, or more charges are caught by the drop or lost by it. By measuring the mass of the drop, the difference of potential across the plates, the velocity of fall under the action of gravity only, and the velocity of rise against gravity under the action of the field, the elementary electrical charge, the charge on the electron, can be calculated. Robert A. Millikan used this method

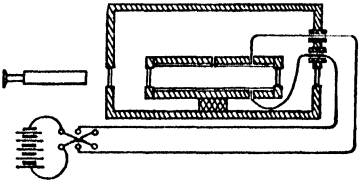


Fig. 212. Oil Drop Apparatus.

to give us a most accurate value of the charge on the electron.

Positive Rays

A discharge tube contains, in addition to the free electrons shot out from the cathode, positively charged atoms or molecules of the residual gas. If a hole is cut through the cathode, these rapidly moving positive ions or positive rays will stream through it and produce a fluorescence in the residual gas behind the cathode. If the tube is designed so that a photographic plate may be hit by these rays, then the deflection produced on a stream of them by the action of magnetic and electric fields may be photographed. When such a stream of positive rays, consisting of ions of nitrogen, oxygen, chlorine, and argon, is deflected, a series of lines is found on the photographic plate. Each gas produces a line of its own, the extent of the deflection depending on the type of apparatus used and on the atomic mass of the gas.

Aston discovered that all ions of nitrogen fall on one line and all ions of oxygen on another, but that the ions of chlorine fall on two distinct spots, as if this gas were made up of a mixture of two elements of atomic weights 35 and 37. These newly discovered components of chlorine are called its *isotopes*. They cannot be distinguished chemically; hence, when chlorine is generated in the laboratory, a mixture is obtained which cannot by ordinary chemical means be detected as such. This mixture of atoms with the same chemical properties but with atomic weights represented by the whole numbers 35 and 37 gives rise to 35.457, the average atomic weight of chlorine. It is now known that nitrogen has two isotopes, and that oxygen has three.

In fact, most elements consist of several isotopes, the total number of different isotopes being upwards of 250, among the 92 different chemical elements.

X-Rays

In 1895 William Conrad Roentgen discovered that whenever cathode rays strike the walls of the tube or any object placed in the tube, such as a platinum target, they give rise to another type of invisible radiation which is now known as *X-rays* or *Roentgen rays*. These rays penetrate many substances which are opaque to light; they have the ability to ionize air; they produce fluorescence in various types of crystals; they set electrons free in organic matter, curing certain diseases, but under long-continued exposure produce burns which are difficult to heal; and they affect a photographic plate, making possible X-ray pictures (Fig. 213).



Fig. 213. X-ray Picture.

Coolidge tube. In the early X-ray bulbs, the electrons were released from the cathode by the bombardment of the positive ions of the residual gas. The intensity of the cathode stream, and consequently the intensity of the X-rays, was difficult to regulate because of the uncertainty of the positive ions. However, in 1913 W. D. Coolidge invented an X-ray tube which uses a hot filament as the source of electrons. This tube is very highly evacuated, and the target is usually made of molybdenum. Since the hot filament furnishes the carriers, low as well as high voltages may be used. Thus a control of the energy

with which the electrons strike a target is made possible; the so-called "soft" and "hard" X-rays may be generated at will (Fig. 214).

Coolidge has constructed a tube of special design in which the electrons move through a potential of approximately a million volts. Instead of having the corpuscles hit a target and produce X-rays, he has them strike a nickel window five ten-thousandths of an inch thick. The window is supported by a grid so that it does not collapse under the air pressure. Large quantities of cathode rays pass through the window and to a considerable distance into the air. This stream of rapidly moving electrons causes a crystal of calcite to shine with a red glow; it causes castor oil to change from a liquid to a solid, and it causes clear crystals of sugar to become white.

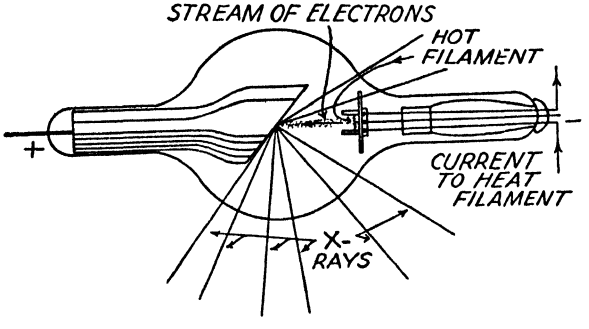


Fig. 214. Coolidge X-ray Tube.

The nature of X-rays. X-rays are not deflected by a magnet or by an electric charge. They do not, therefore, consist of a stream of electric particles. Some scientists suspected that they might be light waves of very short wave length, much shorter than the waves of ultraviolet light, but for many years there was no proof for this guess. The problem was difficult because the X-rays could not be reflected from mirrors or refracted by prisms. But Max von Laue and Sir William Bragg, as early as 1912, suggested that a crystal with atoms arranged in planes might be used as a sort of diffraction grating. Their ideas were soon shown to be correct. By using crystals, such as quartz and rock salt, they were not only able to measure the wave length of X-rays and establish the fact that these rays belong to the same family as light waves, but

they were able also to determine the nature of the crystal structure.

Crystal structure. Common table salt (sodium chloride) forms cubical crystals. Can we say that its atoms of sodium and chlorine form a cubical lattice as shown in Fig. 215? One might guess that this is so, but how may we find evidence in favor of the picture? We have already stated that Laue and Bragg were successful in measuring the wave length of X-rays by using a crystal as a diffraction grating for these very short wave lengths. Before this measurement could be made, the nature of the crystal had to be determined much as follows:

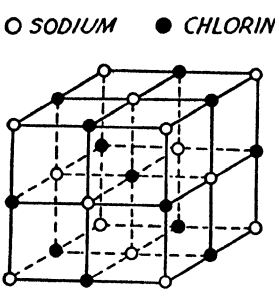


Fig. 215. Salt Crystal.

X-rays of a definite frequency are allowed to fall upon a crystal face. There is no observable reflection until a proper angle between the face and the beam is reached; then a sudden flash of reflection is observed. Another face is presented; the angle of maximum reflection is found to be the same. Finally, the third face is presented with similar results. This means that the spacings are cubical in form, and that the atoms are arranged as shown in the figure—one atom at each corner of a tiny cube. If d is the distance between corners, there will be $1/d^3$ such little cubes in one cubic centimeter of crystal. Each little cube contains an atom, and two of them a molecule; consequently, the number of molecules of salt in a cubic centimeter is $1/2d^3$. But the weight of a salt molecule is 58.0 times the weight of a hydrogen atom, or $58.0 \times 1.662 \times 10^{-24}$ grams. The weight of one cubic centimeter of rock salt is 2.17 grams; hence, the number of molecules per cubic centimeter is $2.17/(58.0 \times 1.662 \times 10^{-24})$ grams. This quantity is also equal to $1/2d^3$; therefore, we can solve and find the distance between corners, the so-called *grating space*, and get

$$d = 2.81 \times 10^{-8} \text{ cm.} \qquad (17.1)$$

By means of this distance and the angle necessary for maximum reflection, the wave lengths of certain X-rays may be determined, and then the nature of other crystals may be inferred. This process opens up the broad field of crystal analysis. Models built in accordance with this type of analysis are made, and if available will supply many hours of fascinating study.

In the salt crystal, the sodium atom gives up one electron to the chlorine atom, and the attractive forces between these ions hold the crystal together. In the diamond, the electrons are shared by the atoms, and the forces are very strong, as is evidenced by the hardness of this precious gem.

General X-radiation. Two kinds of X-radiation are known, the so-called *general* X-radiation and the characteristic X-radiation. The general X-radiation has a short wave length limit, that is, an upper limit of vibration frequency which is determined by the voltage applied to the tube (Frequency = 2.43×10^{14} per volt). When an electron is stopped instantaneously by the target, it converts its kinetic energy into X-radiation. But what determines the frequency of vibration of the X-rays produced?

This brings us to the *quantum theory* first promulgated by Max Planck. According to this theory, energy is put up in bundles of sizes depending on the product of the frequency of vibration and a universal constant, called *Planck's constant.* Applying the quantum theory to the production of X-rays, Planck and Einstein have shown that the energy given up by the electron is equal to the frequency of the X-rays produced times Planck's constant. If the energy of the impinging electron is changed into X-radiation not all at one time, but in installments, then the frequency of the vibration at each installment will be lower than that produced when the whole amount of energy is given up at once. This gives rise to a rather wide band of frequencies called *general X-radiation.* The highest frequency can never be greater than that produced when all the energy of an electron is transferred to radiation at one time.

Hence, "hard" or penetrating rays may be produced by the aid of high voltages across the tube (a large amount of energy on an electron when it strikes); "soft" rays may be produced by the use of low voltages.

Characteristic X-radiation. If the voltage across the tube is sufficiently high, the characteristic X-radiation appears as a series of definite vibrations superimposed upon the general radiation. The frequencies in the series are determined by the kind of metal used as a target. To stimulate the characteristic X-radiation, impinging electrons must have energy above a certain amount. But the frequencies are not controlled by this energy, as in the case of the general radiation. Some mechanism within the atoms of the target seems to be the controlling factor. X-ray messages are thus able to give us information concerning the structure of atoms.

The characteristic X-radiation may be classified into a number of series, as the K, L, M, N, O, . . . , series. Each series has a number of frequencies designated as follows: K-alpha, K-beta, K-gamma, and so forth.

Moseley, who left his researches to lose his life in the Dardanelles during the World War, discovered that the chemical elements can be arranged in a natural series according to the frequency of their characteristic X-rays. The change of frequency of the characteristic K-alpha radiation, for example, occurs with remarkable regularity from element to element. With such exactness do the steps take place that every missing element in the series is immediately revealed by too large a jump. The sequence of the elements in this natural series is almost exactly the same as the sequence obtained by arranging the elements in the order of increasing atomic weight. The steps in atomic weight are not regular. But the displacement toward a higher frequency, as one ascends the series, occurs with remarkable precision. These results justify the assigning of a number, known as the *atomic* number, to each element according to its position in the series. Hydrogen

is given the atomic number 1; uranium, the atomic number 92.

Fig. 216 shows how the frequencies of the K-series increase with the atomic number, and how a gap is clearly

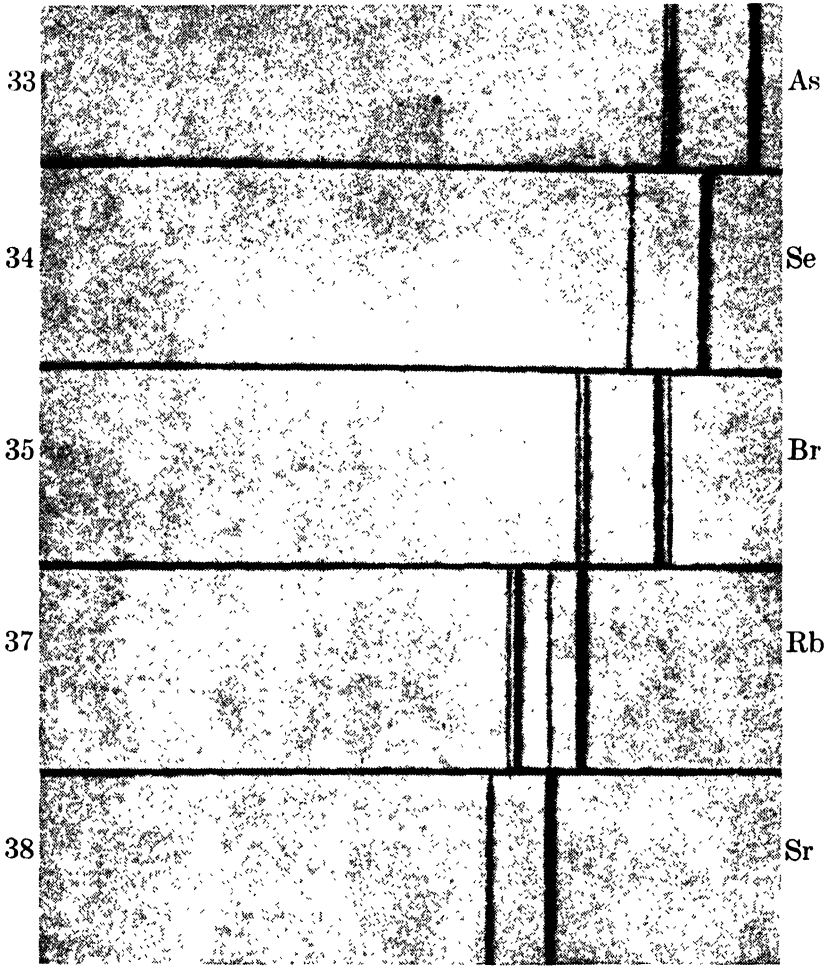


Fig. 216. K-Series. (*From Millikan's "Electrons (+ and −), etc."; courtesy of the University of Chicago Press.*)

suggested. The element missing is the rare gas krypton, whose X-ray spectrum is known to fall into place between bromine and rubidium. Notice the exact similarity of the spectra. The wave length of each line is proportional to the distance from the edge at the left.

Wave lengths of X-rays. All wave lengths which lie between 0.000000001 and 0.0000001 centimeters may be classed as X-rays. The shortest wave length of visible light is approximately 0.000039 centimeters. The wave lengths of the K-alpha lines (Fig. 216) are listed below. (The wave lengths are expressed in *angstroms*. It takes 100,000,000 angstroms to make one centimeter.)

WAVE LENGTHS OF K-ALPHA LINES

Element	*Atomic No.*	*Wave Length of K-alpha Line*
Sodium (Na)	11	11.884
.....	..	.....
Arsenic (As)	33	1.170
Selenium (Se)	34	1.104
Bromine (Br)	35	1.035
Rubidium (Rb)	37	0.926
Strontium (Sr)	38	0.871
.....	..	.....
Tungsten (W)	74	0.209

Radioactivity

In 1896 Henri Becquerel wrapped a photographic plate in black paper, placed a coin on top of the paper, and suspended uranium above the coin. When he developed the plate, he discovered the same shadow picture which X-rays would have produced. Madame Marie Sklodowska Curie (1867–1934) discovered that pitchblende, the crude ore from which uranium is obtained, discharged an electroscope about four times as rapidly as pure uranium discharged it. She looked for some new substance, and after a long and difficult search she was able to separate from several tons of pitchblende a few hundredths of a gram of the new substance she called *radium*.

In 1899 Ernest Rutherford (now Lord Rutherford) showed that the Becquerel rays from uranium and radium are made up of three types of radiations, which he named *alpha*, *beta*, and *gamma rays*. The beta rays are found to be identical with cathode rays, and consist of rapidly moving electrons with speeds of from 60,000 to 180,000 miles per second. Each particle of the alpha rays—an alpha

particle—is found to have a positive charge and a mass about four times the mass of the hydrogen atom and 7,400 times the mass of the electron, and a velocity of about 20,000 miles per second. The alpha particle turns out to be a positive ion of helium. The gamma rays are simply "hard" X-rays.

When an alpha particle drives through the air with its terrific speed and heavy mass, it knocks electrons out of practically all the molecules it encounters. Thus, it leaves a trail of ionized air, which may be made visible as follows: Water condenses on dust particles or ions. If a space is dust- and ion-free, the dew point may be reached without the formation of a fog. When an alpha particle is shot into such a space, a fog track forms along its path because of the air ions that it produces. Such alpha tracks were first demonstrated by C. T. R. Wilson (Fig. 217). A study of these tracks reveals that the alpha particle may penetrate through many thousands of atoms without its direction being appreciably affected. Occasionally, however, its course changes in an abrupt manner. A head-on collision is made with the nucleus of an atom.

In 1903 Sir William Crookes devised an instrument called the *spinthariscope*. When alpha particles, furnished by a tiny speck of radium, strike a zinc sulphide screen, flashes of visible light are emitted. It is possible to see these scintillations by looking through the lens of the spinthariscope. The screen has the appearance of being fiercely bombarded by an incessant rain of projectiles.

If alpha particles are allowed to pass through a thin metal foil and fall upon a zinc sulphide screen, their scattering caused by coming in close proximity to the metal atoms can be measured. This scattering is always small, whether measurements are made in air or in a thin metal foil. Here is evidence that the nuclei of the atoms are very, very small as compared to the diameters of the atoms as determined, say, from the amount of space an atom occupies when solidly packed with others in a crystal. This has led

Rutherford to assume that the atom is made up of a positively charged nucleus around which electrons revolve like

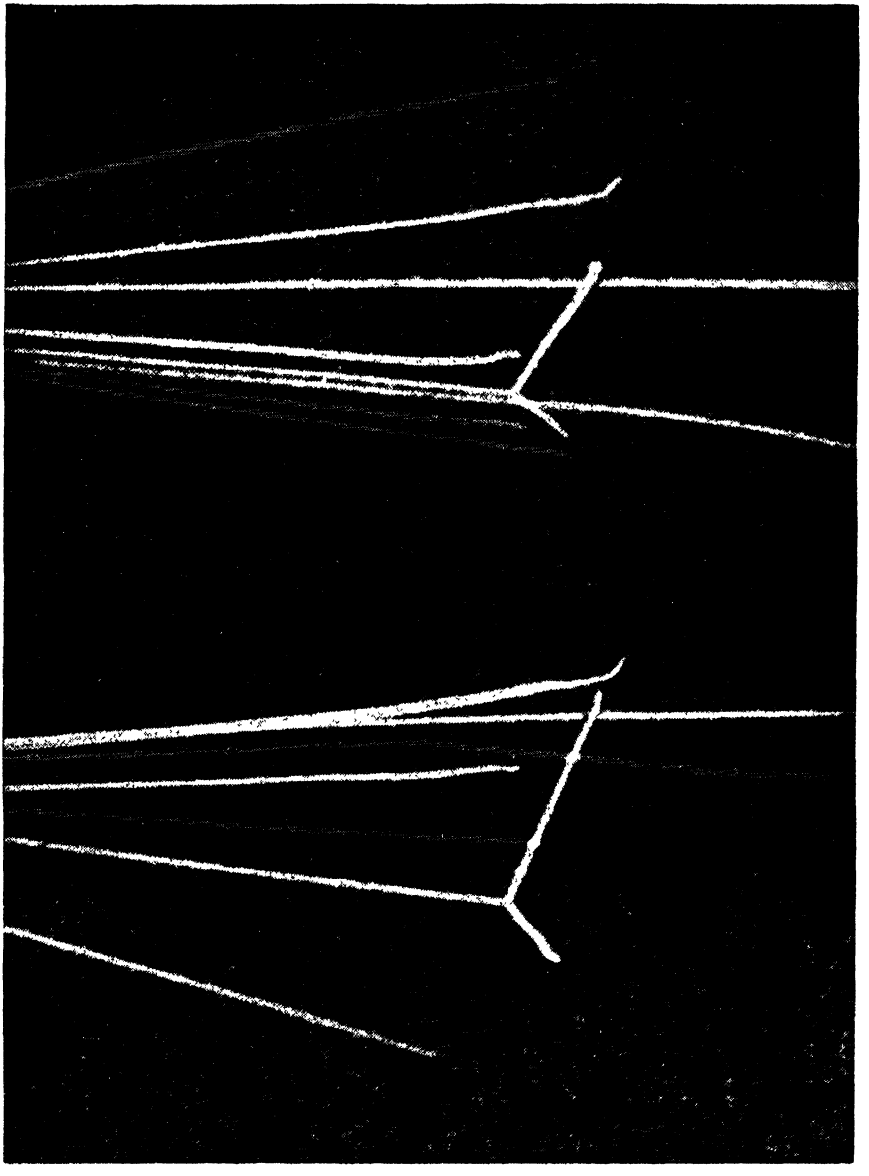


Fig. 217. Collision of Alpha Particles with Oxygen Atoms. (*From Millikan's "Electrons (+ and −), etc."; courtesy of the University of Chicago Press.*)

planets in a solar system. As an alpha particle passes through such an atom, it usually simply brushes aside the

electrons without experiencing a deflection. But sometimes it comes very close to the nucleus; it is then deflected and in certain cases receives a terrific impact. In this way the sharp changes in direction of the alpha tracks are explained.

It is thought that the alpha particles and the electrons making up, respectively, the alpha and beta rays of radium and other radioactive substances, are shot out of the nuclei of the radioactive atoms, the final end-products being the isotopes of lead (atomic weights 206, 207, and 208).

Altogether, there are several dozen radioactive substances known. These disintegrate spontaneously at average rates which man has so far been unable to influence in any way. This spontaneous disintegration seems to have been going on at unaltered rates through long geological ages, each radioactive substance having its own rate of decay. For example, radium disintegrates at such a rate that it is one-half gone in 2,000 years. Other elements are half gone in a small fraction of a second and would not now be available for observation if it were not for the fact that they are the "offspring" or decomposition products of much more long-lived parents.

Artificial disintegration of atoms. In 1919 Rutherford opened up a whole new field of investigation by showing that it is possible to break up the nuclei of nitrogen by bombarding nitrogen with alpha particles. A breaking up of the nucleus means, of course, a real transmutation of the element. However, the nuclei are so small that very few alpha particles make effective hits; hence, physicists are far from being able to produce even microscopic quantities of transmuted elements. It was soon found that many other elements could be disintegrated in this way. This field of investigation has been vastly widened in the last few years by the development of means of producing strong beams of positive ions which have been given energy in a potential drop of upward of a million volts. This has made possible the study of a large number of nuclear

reactions in which the ordinary elements are bombarded with *protons* (the nuclei of hydrogen atoms) and also with *deuterons*, as the nuclei of heavy hydrogen atoms are called. Out of this work has come recently the extraordinary discovery that some nuclear changes do not require such enormous voltages. Thus, a proton striking lithium will combine with it to produce two atoms of helium, and this will happen to some extent at voltages no higher than those used in an ordinary dental X-ray apparatus.

The structure of the nucleus. In the process of the artificial disintegration of atoms, the atomic nuclei are known to yield, under one circumstance or another, one or more of the following particles: alpha particles, deuterons, protons, neutrons, electrons, and positrons. Even so, it is not known exactly how these basic building stones are fitted into the structure making up the atomic nucleus. So far, we have explained electrical phenomena in terms of negatively charged electrons and positively charged protons. Until the discovery, in 1932, of the *neutron* by Chadwick, of the Cavendish Laboratory, and of the *positron* by C. D. Anderson, of the California Institute of Technology, the electron and proton were considered to be the primordial entities out of which all things were built. But the neutron turns out to be a particle with *no electrical charge* and with a mass very nearly the same as that of a proton (the nucleus of a hydrogen atom). The positron is found to have a positive charge equal and opposite to that of the electron and a mass thought to be the same as that of the electron. Following the hypothesis that nuclei are composed of protons and neutrons, we conclude that an *alpha particle* is composed of two protons and two neutrons and that a *deuteron* is composed of one proton and one neutron. Thus we have reduced the list of what seemed to be six basic ingredients to four. Is the proton simply a neutron with a closely attached positron? Is the neutron a proton with a closely attached electron? If either of these questions can be answered in the affirmative, we may reduce the

fundamental entities to three: the *electron*, the *positron*, and either the *proton* or the *neutron*.

Atomic Structure

Radioactive transformations (both natural and artificial), the scattering of alpha particles, Moseley's discovery of the characteristic X-rays (with the need for a mechanism which will give rise to frequencies dependent upon the so-called atomic number), the discovery of isotopes of the elements with fractional atomic weights, and the discovery of neutrons are facts which may be explained and correlated by the so-called Rutherford-Bohr atom—the Rutherford atom redesigned by Bohr to meet the needs of the quantum theory. (This atomic model is not adequate to interpret all present-day facts, but it still serves as an excellent picture for the beginning student.)

Making use of the experimental facts presented in the above pages and accepting the Rutherford-Bohr atomic model, we build the picture of the microcosm much as follows: There are just 92 elements. These elements have a positively charged nucleus of minute size surrounded at a considerable distance by the number of electrons requisite to maintain the structure electrically neutral. The nucleus is composed of protons and neutrons, and the atomic number is equal to the number of protons in the nucleus. The atom of hydrogen, the lightest element, consists of one proton as nucleus with one electron at a *considerable distance from it*—a miniature solar system with a single planet. The uranium atom, the element with greatest atomic weight, contains a nucleus, made up of 92 protons and 146 neutrons, and, at varying distances from the nucleus, as planets in a solar system, 92 electrons. Fractional atomic weights (such as 35.5 for chlorine) are explained as being caused by a mixture of isotopes, atoms having identical chemical properties owing to the presence of the same number of free charges on the nucleus but

having different atomic weights because of the extra neutrons which add mass to the nucleus but do not add free charge to it.

Theory of Magnetism

H. A. Rowland (1848–1901) showed in a very important experiment that a rotating electric charge produces a magnetic effect similar to a current of electricity. It follows, therefore, that a rotating electron produces a magnetic field, and that the atom is equivalent to a magnetic particle, a so-called *magneton*. All substances are found to be affected by a magnetic field: some are pushed out of the strongest part of the field and are called *diamagnetic* (bismuth is a splendid example); others are attracted into the strongest part of the field and are called *paramagnetic*. Some paramagnetic substances, such as iron, are very magnetic and are called *ferromagnetic*.

The natural reaction between an outside magnetic field and magnetons is of two types. The first reaction is one in which the rotation of the electrons in their orbits is speeded up or slowed down by the outside magnetic field; this reaction explains diamagnetism. The effect is a property of all atoms, but is so small as to be overshadowed by paramagnetism or ferromagnetism when either of the latter is present. The second reaction orients the atoms, molecules, or groups of molecules so that they are all aligned in a definite manner. This reaction explains paramagnetism. Such an orientation occurs only with those substances whose atoms show an exterior magnetic field. For example, silver in the solid form is diamagnetic, but the single atoms are paramagnetic. In the solid state, the atoms seem to be combined in such a manner as to have their external magnetic fields cancel out. Ferromagnetism may be explained by assuming interactions of atomic fields to form powerful magnetic groups. These groups are probably much smaller than those making up individual crystals.

Electromagnetic Waves

James Clerk Maxwell (1831–1879) showed by a profound mathematical analysis that the phenomena discovered by Oersted and Faraday—the production of a magnetic field by an electric current, and the production of an electric current by a varying magnetic field—should give rise to electromagnetic waves having a velocity equal to the velocity of light. This was the beginning of a great unifying process that has caused visible light, radiant heat, radio waves, and X-rays to be recognized as members of a great family of waves produced by moving charges of electricity.

Bohr assumed that the electrons which revolve about the nucleus of the atom do not radiate energy except as they change from one so-called *energy level* to another, the frequency of the radiation depending upon the energy radiated in accordance with the quantum theory. Thus, the high-frequency X-rays are produced in the heavier atoms by an electron's falling close to the nucleus and giving up, as it does so, a large amount of energy. When an electron falls toward a hydrogen nucleus, the force of attraction is due to one proton. When it falls toward a uranium atom, the attraction is due to 92 protons. This means that the size of energy bundle which the uranium atom can radiate is greater than that which the hydrogen atom can radiate. In terms of frequencies, it means that uranium may be the source of penetrating X-rays, but hydrogen may not produce frequencies higher than those of ultraviolet light.

Let the circles about the nucleus of the hydrogen atom illustrated in Fig. 218 represent the energy levels of the planetary electrons. If the electron jumps from orbit 2 to orbit 1, the K-alpha radiation is given out; if it jumps from orbit 3 to orbit 1, the K-beta radiation is emitted; if it jumps from orbit 4 to orbit 1, the K-gamma radiation is radiated. These K-radiations are all in the ultraviolet range of frequencies and make up the so-called Lyman series. If an electron jumps from orbit 3 to orbit 2,

orbit 4 to orbit 2, and so on, the L-alpha, L-beta, and continuing radiations are emitted, and these radiations give rise to the visible Balmer series. Again, if an electron jumps from orbit 4 to orbit 3, orbit 5 to orbit 3, and so on, the M-alpha, M-beta, and continuing radiations are emitted, and they give rise to the infra-red Paschen series. The great triumph of the Rutherford-Bohr atom is that it

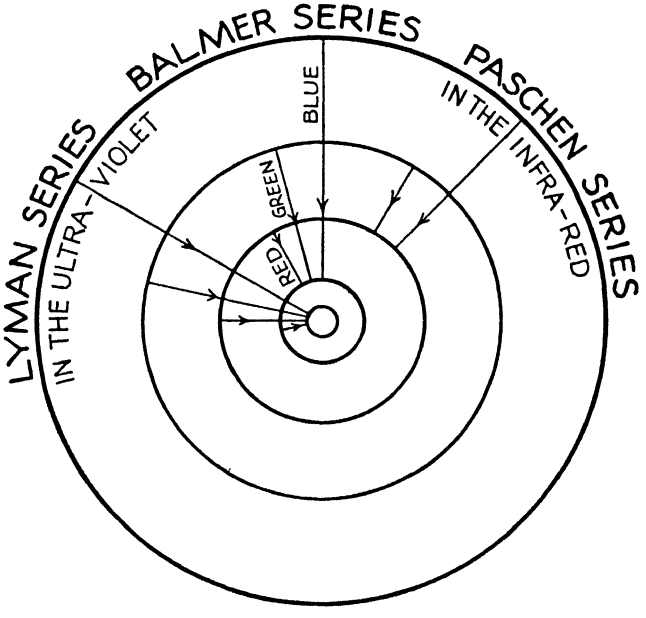


Fig. 218. The Hydrogen Atom.

fits the facts of spectrum analysis not only qualitatively but quantitatively.

The shift between energy levels near the heavier atoms means a greater energy radiation; hence, the K, L, M, and continuing radiations increase in frequency with the atomic number until the heavier atoms emit these radiations as X-rays. Roughly speaking, then, X-rays are produced by electrons shifting in orbits near the heavy atoms; ultra-violet and visible light waves are produced by electrons shifting in the outermost orbits of the heavier atoms or in the inner orbits of the lighter atoms; infra-red waves are

produced by vibrating atoms or molecules which are electrically charged.

Representative electromagnetic waves may be listed in the order of increasing frequency as follows: alternating current, radio, Hertzian waves, infra-red (heat radiations), visible rays, ultraviolet, "soft" X-rays, "hard" X-rays, gamma rays, cosmic rays. Beyond the visible red light have been found waves with wave lengths of three-tenths of a millimeter. Electric waves, which ordinarily have lengths between 50 and 300 meters, have been produced by Nichols and Tear with a wave length as short as six-tenths millimeter; hence, the gap between the so-called wireless waves and light waves has been practically bridged. On the other end of the visible spectrum beyond the ultraviolet, a complete series of decreasing wave lengths and, therefore, increasing frequencies has been found. That is, the ultraviolet passes without a gap into the "soft" X-rays, the "soft" X-rays into "hard" X-rays, and the "hard" X-rays into gamma rays.

When electromagnetic waves are absorbed and re-radiated, the frequency of the re-radiated wave is ordinarily not greater than that of the original, and, owing to the lack of a complete energy transfer, it may be less. This is in accordance with the quantum theory and explains why X-rays and ultraviolet light may produce a visible fluorescence and why the heat waves from the sun which pass through the glass of a hotbed reappear as very much longer heat waves which may not pass out through the glass.

Cosmic Rays

Einstein has suggested that mass may be converted into energy; that the mass lost, measured in grams, times the square of the velocity of light, measured in centimeters per second, is equal to the energy gained in ergs. Now the atomic weight of hydrogen is 1.0078, and it is composed of one proton and one electron; the atomic weight of helium is 4.002, and it is composed of four protons and four elec-

trons (or two protons, two neutrons, and two electrons); the atomic weight of oxygen is 16.000, and it is composed of sixteen protons and sixteen electrons (or eight protons, eight neutrons, and eight electrons). When helium and oxygen are built up out of the primordial elements, protons and electrons (or protons, electrons, and neutrons), what happens to the extra mass (0.0292 for helium and 0.1248 for oxygen)? It is radiated as cosmic rays, according to an hypothesis due to Millikan.

Millikan and his students have measured the penetrating power of these rays and have found that they can pass through 200 feet of water, or the equivalent of 18 feet of lead, before being completely absorbed. The most penetrating X-ray produced in hospitals cannot go through one-half an inch of lead. These new rays seem to originate out in space, and, according to a hypothesis proposed by Millikan, have their origin in the creation of one or more of the common elements out of hydrogen. At least, the frequencies of these radiations, as determined by their penetrating power, indicate that the bundles of energy radiated could come from the decrease of mass, as protons and electrons are packed into the nuclei of atoms. Thus we may guess that elements are created in the laboratories of the stars.

However, the complete story of the cosmic rays is far from being known. A large part of the radiation as it is observed consists of charged particles of extremely high energy instead of electromagnetic waves of high frequency. This has been revealed in the last two or three years through the interpretation of an elaborate set of intensity measurements made at different places on the earth. An adequate interpretation of these observations requires that the rays be deflected in the earth's magnetic field, and this leads us to charged particles rather than to electromagnetic waves. No doubt some particles are secondary in nature, being generated as outside primary rays are absorbed in the air or in other materials.

Corpuscles and Waves

The Compton effect. In 1923, A. H. Compton found that when an X-ray beam of single frequency strikes a piece of material such as carbon, X-rays are scattered in two portions, the "unmodified" part having a frequency of the incident ray and the "modified" part a slightly lower frequency. No explanation based on classical wave theory can be suggested, but a combination of the quantum theory and the principle of the conservation of energy and momentum furishes a simple explanation, provided the *quantum of X-rays* is considered as being able to collide with a free electron much as one ball collides with another. Thus, the quantum of X-rays (*photon*) strikes the electron a glancing blow; the electron receives energy and recoils, and the quantum moves off at an angle as a "modified" X-ray quantum. This explanation accounts quantitatively for the lower frequency (smaller energy content per quantum) of the "modified" portion of the scattered X-rays. This type of experiment evokes the corpuscle rather than the wave aspect of radiant energy.

The diffraction of electrons. In 1927 C. J. Davisson and L. H. Germer performed an experiment in which a stream of electrons striking a single crystal of nickel was diffracted as if the electrons were waves with a wave length of the order of ordinary X-rays. Thus it is clear that under a certain laboratory procedure, the electron, which ordinarily acts like a particle, plays the role of a wave.

Uncertainty principle. Heisenberg in 1927 proposed the uncertainty principle, which states that it is impossible to determine at the same time the exact position and velocity of an electron. Either property may be determined accurately, but the measurement involved in getting one property spoils the accurate measurement of the other. We have been wrong, according to this principle, in the simple assumption that we can measure anything as accurately as we please. This conclusion follows not because of our

failure to be omnipotent, but because of the nature of the universe itself.

It has been our attempt, through the pages of this text, to emphasize that *our scientific concepts emerge from patterns inferentially woven out of data obtained from apparatus.* We have seen that certain types of experiments evoke the particle aspect of the electron, while others evoke its wave aspect; the same is true of radiant energy. With keen insight, Bohr has pointed out that the uncertainty principle leads to the conclusion that it is impossible for an experiment to evoke the *particle and wave aspect* of an electron or quantum of energy *at the same time.* Thus, we never shall find the "particle-wave" electron, but only the "particle-or-wave" electron. We determine the "cloak" the electron concept wears by the apparatus we choose to use.

Suggested Readings

(1) Blackwood, O. H., *et al.*, *Outline of Atomic Physics*, John Wiley and Sons, Inc., New York, 1933.
(2) Darwin, Charles, *New Conceptions of Matter*, The Macmillan Company, New York, 1931.
(3) Harrow, B., *From Newton to Einstein*, D. Van Nostrand Company, Inc., New York, 1920.
(4) Jaffe, B., *Crucibles*, Simon and Schuster, New York, 1930, Chaps. XII–XVI.
(5) Kaye, G. W., *The Practical Application of X-rays*, E. P. Dutton and Company, Inc., New York, 1923.
(6) Langdon-Davies, John, *Man and His Universe*, Harper and Brothers, New York, 1930, Chap. VII.
(7) Millikan, R. A., *Electrons (+ and −), Protons, Neutrons, and Cosmic Rays*, University of Chicago Press, Chicago, 1935.
(8) Planck, Max, *Where Is Science Going?* W. W. Norton and Company, Inc., New York, 1932.
(9) Pupin, M. I., *The New Reformation*, Charles Scribner's Sons, New York, 1927.
(10) Rasetti, F., *Nuclear Physics*, Prentice-Hall, Inc., New York, 1936.
(11) Sullivan, J. W. N., *The Limitations of Science*, The Viking Press, Inc., New York, 1933.

APPENDIX

Solving Problems and Answering Questions

In solving the problems and answering the questions at the close of each chapter, the student may find the following suggestions helpful.

Problem solving. In solving problems

(*a*) Determine the exact relationship between the elements of the problem, and, if possible, express the relationship in terms of an algebraic formula.

(*b*) Substitute the numerical values in the equation and solve for the unknown.

(*c*) Example: In Problem 3, Chapter II, the relationship needed is the law of the lever,

$$WL_w = FL_f$$

Substituting the numerical data, we get

$$10W = 100 \times 8$$

$$W = \frac{100 \times 8}{10} = 80$$

Thus, the weight of the child is 80 pounds.

Completing sentences. Important statements which the student should thoroughly understand and master are introduced with certain important words missing. By carefully reviewing the material of the text, the student should be able to determine and supply these words.

Procedure commands. The questions make use of the commands: *List, Outline, Compare, Contrast, Describe how,* and *Explain why.* As used in this text, these words have the following meaning:

(*a*) *List.* A simple naming of the elements, one after the other.

(*b*) *Outline.* A listing in which coördinate and subordinate elements are properly arranged.

(*c*) *Compare.* A separate listing of similarities and dissimilarities.

(*d*) *Contrast.* A listing of dissimilarities only.

(*e*) *Describe how.* A listing of the steps of the procedure.

(*f*) *Explain why.* A listing of the steps of the procedure and under each step a listing of the reasons why.

Index

A

Aberration:
chromatic, 289
spherical, 286
Absolute temperature scale, 92, 94
Absolute zero, 92
Absorption:
color the result of, 264
of sound, 173
Acceleration, 27
Achromatic lens, 289
Acoustics of auditoriums, 233
Adhesion, 113
Air columns, vibration of, 155
Air resistance, 46, 66
Alpha particle, 354, 357
Alpha tracks, 354
Alternating current, 333
Ammeter, 336
Ampère, André, 316
Ampere, the, 335
Amplification:
by forced vibration, 206
by resonance, 205
Amplitude of vibration, 141
Anastigmatic lens, 289
Anderson, C. D., 357
Anderson, O. A., 75
Aneroid barometer, 74
Antinode, 148
Aristotle, 28, 240
Armature, 332
Articulation:
testing, 230
word list, 231
Astigmatism, 289
Aston, F. W., 346
Atomic number, 351
Atomic structure, 358
Atoms, 69
artificial disintegration of, 356
evidence for, 69
Audiometer, 187
Audition, limits of, 179
Auditoriums:
acoustics of, 232
good design of, 234
poor design of, 234
size and shape of, 233
Auditory sensation area, 180
Ault, James P., 320

B

Balance:
beam, 60
conditions for, 21
spring, 11
Balloon, stratosphere, 75
Balmer series, 361
Banking a road, 35
Barometer:
aneroid, 74
mercury, 73
Bartholomew, W., 230
Basilar membrane, 193, 195
Battery, lead storage, 328
Beats (sound), 199
Becquerel, H., 353
Bell, A. G., 181
Beta rays, 353
Blue sky, 249
Bohr, N., 241, 358, 365
Boiling, description of, 118
Boiling point:
defined, 120
effect of pressure on, 120
of solutions, 124
Boulder Dam:
photograph, 107
switching station, 337
Bowing a string, 210
Boyle, R., 76
Boyle's law, 74
Bragg, Sir William, 348
Brass instruments, 215
Brightness:
of a color, 276
of a surface, 251
British thermal unit, 96

C

Caloric, 85
Calorie, defined, 96
Camera:
  depth of focus, 291, 293
  fixed focus, 294
  operation, 290
  speed, 291
Candle power, 243
Canning fruit, 121
Capillary action, 114
  law of, 115
Carbon-button microphone, 333
Cathode rays, 344
Cavendish, H., 45
Cells:
  photoelectric, 330
  voltaic, 325
    action, 326
    charge on terminals, 311
Celsius, A., 89
Center of gravity, 17
  of human body, 19
  position, 18
Centigrade scale, 89
Centimeter, 10
Centrifugal force, 34
Centripetal force, 34
Charge on electron, 345
Charles, J. A. C., 91
Charles's law, 92
Chimney, 131
Chladni, E. F. F., 151
Chladni's figures, 152
Chromatic aberration, 289
Circular motion, 27
Clarinet, 212, 214
Cochlea, 193
Cohesion, 81
Colds, 129
Collins, Jimmy, 47
Color:
  absorption, 281
  blindness, 273
  characteristics of, 276
  chart, 280
  circle, 281
  groups, 282
  intensity, 279
  pyramid, 277
  saturation, 278
  value, 272, 279
Colored lights, 267
  mixing of, 270
    law of, 272
  three fundamental, 271
Colors:
  by absorption and reflection, 264
  complementary, 268
  in art, 279
  in nature, 260
  of soap bubbles, 263
  of thin films, 261
Color vision, 273
  Ladd-Franklin theory of, 274
Compass, magnetic, 319
Complementary colors, 268
Complex vibrations, 153
Component of force, 20
Compton, A. H., 241, 364
Compton effect, 364
Concurrent forces, 20
Condensation, 118, 169
Condenser (electric), 308
Conduction (heat), 130
Conductors (electric), 308
Conservation:
  of energy, 57
  of matter, 70
Consonance, 199
Consonants, 227
Contact electromotive force, 323
Convection, 131
Convex lens, defects of, 286
Coolidge, W. D., 347
Corpuscles, 241
Cosmic rays, 362
Coulomb, the, 336
Crandall, I. B., 223
Crookes, Sir William, 354
Crystal:
  grating space, 349
  snow, 83
  structure, 349
Curie, Marie S., 353

D

Dalton, John, 69
Damped oscillations, 142
Davisson, C. J., 241, 364
Davy, Sir Humphrey, 86
Deafness, 186
  caused by masking, 187
  of school children, 187

Decibel, the, 181
Declination, 319
Deductive system of discovery, 3
Definite proportions, 70
Depth of focus, 291, 293
Descartes, 8, 41
Deuterons, 357
Dew point, 128
Diamagnetism, 359
Diaphragm, vibrating, 216
Diatonic scales, 201
Diffraction:
  grating, 264
  of light, 245
Diffuse reflection, 247
Diffusion:
  in liquids, 82
  of gases, 72
  through a porous cup, 72, 79
Dispersion, 266, 289
Dissonance, 199
Distortion of image, 288
Dry cell, 328
Du Fay, C. F., 307
Dyne, the, 44

E

Ear:
  operation, 194
  structure, 192
Edison, Thomas A., 340
Efficiency of machines, 67
Einstein, A., 241, 350, 362
Electric:
  battery, 311
  cell:
    action, 326
    defects, 327
  chemical effect, 313
  circuit, 339
  current, alternating, 333
  discharge in partial vacua, 343
  heating, 313, 339
  heating appliances, 340
  lamps, 341
  measurements, 335
  motor, 334
  power, 338
  resistance, 338
  spark, 312
Electricity:
  generation of, by:
    chemical action, 325
    friction, 322
    heat, 330
    induction, 323
    light, 329
    motion in a magnetic field, 331
  magnetic effect, 315
  two kinds, 305
Electrification by:
  friction, 322
  induction, 323
Electrodes, 313
Electrolysis of water, 314
Electrolyte, 313
Electromagnet, 317
Electromagnetic induction, 33[illegible]
Electromagnetic waves, 360
Electron, 307
  charge on, 345
  diffraction, 364
  particle and wave aspect, 365
Electrophorus, 324
Electroplating, 315
Electroscope, 307
Energy, 2
  bodily requirement, 102
  conservation of, 57
  electric to mechanical, 334
  heat to mechanical, 99
  in foods, 101
  kinetic, 54
  latent, 109
  level in atoms, 360
  mechanical to electric, 332
  nature of, 50
  on light waves, 240
  of oscillations, 140
  potential, 57
  solar, 103
  speech, 222
  through plant life, 104
  use of natural, 50
Engines:
  gasoline, 100
  steam, 100
Equilibrium:
  requirements for, 21
  types, 21
Evaporation, 115
Experiment, controlled, 14

Eye, 295
defects, 296
resolving power, 297
Eyepiece, 301, 302

F

Facts, 4
and hypotheses, 6
Fahrenheit, G. D., 89
Fahrenheit scale, 89
Falling bodies, 27
effect of friction on, 46
Faraday, M., 332
Far-sightedness, 296
Ferromagnetism, 359
Films, 290
Fletcher, H., 180, 194, 225, 231
Fluorescence, 243
Flute, 213
Focal length, 285
Fog, 173
Foot-candle, 250
meter, 253
Foot-pound, 52
Foot, the, 10
Force:
and acceleration, 31
and motion, 25
centrifugal, 34
centripetal, 34
precise definition, 43
simple definition, 10
Forces:
components of, 20
concurrent, 20
in equilibrium, 20
parallel, 19
parallelogram of, 20
Formulas, method of numbering, 12
Franklin, B., 312, 315
Freezing point of water:
effect of pressure on, 110
effect of salt on, 111
Frequency of vibration, 142
of vibrating strings, 165
Fresnel, A. J., 241, 245
Friction, 65
reducing, 67
rolling, 66
sliding, 65
Fundamental (vibration), 152
Fuses, 339

G

Galileo, 4, 28, 41, 87
Galvani, A., 327
Galvanometer, 325
Gamma rays, 354
Gasoline engine, 99
Gas pressure, 73
cause of, 76
Generator, 332
rule, 332
Germer, L. H., 364
Gilbert, W., 306
Glare, 252
Glottis, 220
Gram, the, 41
Grating, diffraction, 264
Gravitational attraction, 44
Newton's law, 45
Gravity, center of, 17
Greek theater, 235

H

Hammer, operation of, 53
Harmonics, 153
Hearing:
impaired, 186
limits, 185
Heat:
and work, 97
conductivity, 134
insulation, 133
measurement, 95
mechanical equivalent, 99
nature, 85
of vaporization, 123
work from, 99
Heisenberg, W., 241, 364
Helmholtz, H. von, 194
Heyl, Paul R., 46
Home heating, 130
Hooke, Robert, 13
Hooke's law, 148
Horn, 216
Horsepower, 53
Hot air furnace, 132
Hot water furnace, 133
Hot water tank, 132
Hue, 277, 279
Human body, 1
Human machine, 100
Humidity, 127

Huygens, C., 241
Hydrogen atom, 361
Hygrometer, 129, 130
Hypothesis, 4, 6

I

Ice, 109
  heat of fusion, 109
Ice cream freezer, 111
Ideal gas, 94
Illumination, 249
  foot-candle requirement, 252
Image:
  brightness, 290
  of convex lens, 286
  real, 285
  size of, 285
  virtual, 284
Incandescence, 242
Inclined plane, 63
Induced current, 331
Inertia, 39
Insulators, 308
Intensity (sound), 171
  level, 180
Interference, 163
  fringes, 262
Interferometer, Michelson, 262
Ions, 326
Isotopes, 71, 346
  separation by mechanical means, 36

J

Joule, J. P., 98
Joule, the, 52

K

K-alpha lines, 353
Kelvin, Lord, 92
Kepler, J., 33
Kilogram, 41
Kinetic energy, 54
  calculation, 55
  equation for, 56
Knudsen, V. O., 236
Koenig, R., 178
Kundt's apparatus, 156

L

Ladd-Franklin, C., 274
Lagrange, J. L., 30
Latent heat, 109
Laue, Max von, 348
Leaning Tower of Pisa, 29
Left-hand rule (motor), 334
Length, units of, 9
Lens, convex, 284
Lever:
  arm, 15
  classes of, 16
  law of, 14, 15, 60
  of the foot, 16
  of the human body, 17
  simple, 59
Leyden jar, 310
Light:
  diffraction, 245
  interference, 262
  reflection, 246
    diffuse, 247
    law of, 247
    regular, 247
    selective, 266
  refraction, 256
  selective absorption, 265
  source, 242
  valve, 300
  velocity, 244
  wave length, 262
  waves, 241
Limits of hearing, 179
Liquids, 81
Longitudinal vibrations, 155
Longitudinal waves, 159
Loop (vibration), 148
Loudness, 178, 195
  level, 180
  scale, 183
Loud-speaker, 218, 335
Lubrication, 67
Lucretius, Titus, 69
Luminescence, 242
Luminous animals, 243
Luminous vapors, 242
Lyman series, 361

M

Machines, 59
  efficiency, 67

Machines (*cont.*):
human, 100
law of, 59
Magnetic compass, 319
Magnetic field, 318
Magnetism:
of earth, 320
theory, 359
Magneton, 359
Magnets, 317
Magnifying glass, 300
Magnifying power:
microscope, 301
projection lantern, 297
simple lens, 300
telescope, 302
Major diatonic scale, 201
Mass, 39
and weight, 40
units of, 40
Maxwell, J. C., 360
Measurements, 9
Mechanical advantage, 62
Mechanical equivalent of heat, 99
Mechanical vibrations, 138
Meter, the, 10
Michelson, A. A., 244
Microphone, 333
Microscope, 301
Miller, D. C., 175, 192, 209, 214
Millikan, R. A., 241, 307, 345, 363
Minor diatonic scale, 203
Mirror, plane, 284
Mixing colored lights, 267
Molecules:
attraction of, 108
bombardment of, 77
magnitudes, 80
motion, 71
speed, 79
Moment, 15
Momentum, 41
Moseley, H. G. J., 351
Motion:
accelerated, 30
damped, 141
quantity of, 42
simple harmonic, 138
Motor, 334
Mouthpiece, 212, 216
Multiple proportions, 70
Musical instruments, 205
frequency range, 219
Musical intervals, 201
Musical scale, 201
Musical tone, 175
Music halls and studios, 237
Musschenbroek, Pieter van, 308

N

Natural laws, 7
Near-sightedness, 296
Neutrons, 71, 307, 357
Newton, Sir Isaac, 8, 31, 42, 241, 266
Newton's laws of motion:
first, 32
second, 43
third, 47
Newton's law of gravitation, 45
Nichols and *Tear*, 362
Nicholson and *Petit*, 330
Node (vibration), 148
Noise, 175, 188, 235
insulating against, 236
Nucleus of atom, 357
Numerical aperture, 286, 292

O

Objective lens, 301, 302
Oboe, 214
Octave, 177
Oersted, H. C., 315
Ohm, G. S., 338
Ohm's law, 338
Ohm, the, 338
Oil drop experiment, 345
Optimal reverberation time, 237
Osmosis, 125
Overtone, 152
structure, 190

P

Paget, Sir Richard, 225
Parallelogram of forces, 20
Paramagnetism, 359
Pascal, B., 73
Paschen series, 361
Passy, Paul E., 229
Period of vibration, 142
Permanent magnet, 317
Perpetual motion, 56
Phase, 143
Phonetic alphabet, 229

Phonograph, 217
Phosphorescence, 243
Photoelectric cell, 299, 330
Photoelectric effect, 329
Photon, wave and particle aspect of, 364
Piano, 209
Pigments:
  mixing of, 280
  primary, secondary, complementary, 281
Pipe organ, 212
Pitch, 176
  change with loudness, 195
  of a complex tone, 196
  variation along the basilar membrane, 195
Planck, Max, 241, 350
Planck's constant, 350
Plato, 240
Porous cup, 72
Positive rays, 346
Positron, 357
Posture, 22
  correct, 23
Potential, difference of, 336
Potential energy, 57
  examples, 58
Pound, the, 13
Power, 52
Principal focus, 285
Problem solving, 367
Projectiles:
  maximum range, 38
  motion, 37
Projection lantern, 297
Proton, 307, 357
Public address system, 219
Pulleys, 62
Putting the shot, 39
Pythagoras, 240

Q

Quality of sound, 188
Quantum theory, 350

R

Radiation, 130
  from hydrogen atom, 361
Radio, 217
Radioactivity, 353
  artificial, 356
Radium, 353
Rainbow, 260
Rarefaction, 169
Rays:
  alpha, 353
  beta, 353
  cathode, 344
  cosmic, 362
  gamma, 353
  positive, 346
  Roentgen, 347
  X-, 347
Reeds, 151
Reflection:
  light, 247
  sound, 173, 233
  waves, 161
Refraction:
  light, 256
  sound, 173
Refrigerator, 110
Regelation, 110
Resolving power, 246
  eye, 297
  microscope, 301
  telescope, 302
Resonance, 144
Respiration calorimeter, 101
Reverberation, 236
Rheostats, 339
Right-hand rule (generator), 316
Rise of liquids in tubes, 114
Roentgen, W. C., 347
Rolling friction, 66
Rowland, H. A., 99, 359
Rumford (Count), 86
Russell, G. O., 221
Rutherford (Lord), 358

S

Sabine, W. C., 232
Saturated vapor, 116
  pressure, 117
Saturation of color, 278
Scale:
  diatonic, 201
  equal-tempered, 203
Science:
  achievements, 1
  concepts, 7, 365

Science (*cont.*):
co-operative, 8
facts, 4
growth of scientific law, 5
instruments, 6
method, 4, 11
newborn theories of, 8
Secondary foci, 285
Second, the, 25
Shadows, 255
Simple harmonic motion, 138
conditions for, 139
Singing voice, 230
Siren, 176
Size, 9
Skidding, 35
Sky, colors of, 249
Sliding friction, 65
Snow crystals, 83
Snow, W. B., 219
Solids, 81
Sonometer, 189
Sound:
absorption, 173
amplifier, 205
direction of travel, 172
focusing, 233, 234
generator, 205
intensity, 171, 178, 234
level, 180
motion pictures, 298
reflection, 173, 233
refraction, 173
spectrum, 190
subjective and physical, 190
track, 299
transmission, 173
Spark, 312
Specific heat, 96
table, 97
Spectrum:
bronze, 263
solar, 267
Speech:
frequency and intensity range, 228
power, 222
sounds, 223, 226
Speed:
of light, 244
of sound:
effect of temperature, 170
in air, 170
table, 171
Spherical aberration, 286
Spinthariscope, 354
Spring balance, 11
Stability, 21
increasing, 22
Standing waves:
in air columns, 166
patterns, 168
Static machine, 325
Steam:
engine, 99
heating, 133
Steinberg, J. C., 224, 228, 231
Stevens, A. W., 75
Stevens, S. S., 195
Stiffness of a spring, 140
Stop watch, 26
Storage battery, 328
Stove, 131
Stratosphere balloon, 75
Stringed instruments, 207
Stroemer, M., 89
Subjective tones, 196
Sun, source of energy, 103
Surface:
gravity, 46
tension, 112
Sympathetic vibrations, 144

T

Tables:
articulation word lists, 231
current capacity of copper wire, 339
efficiency of electric lamps, 341
efficiency of heat engines, 100
energy of food, 101
heat conductivity, 134
illumination requirements, 252
light reflected from surfaces, 251
man's energy requirements, 102
musical intervals, 201
phonetic alphabet, 229
photographic films, color ranges and sensitivities of, 290
specific heat, 97
speed of sound, 171
standing wave patterns, 168
temperature color scale, 242
vowel sounds, 225
wave length of light, 262
Talking motion picture, 299

Telephone, 216
  receiver, 335
  transmitter, 333
Telescope, 302
  mirror for 200-inch, 303
Temperature, 87
  absolute zero, 92
  coldest produced, 94
  Kelvin and centigrade, 94
  what it measures, 91
Tempered scale, 203
Terminal velocity, 47
Testing airplanes, 36
Thales of Miletus, 305
Theory of color vision, 274
Thermocouple, 330
Thermoelectric effect, 330
Thermometer:
  air, 87
  calibration, 88
  centigrade, 89
  Fahrenheit, 89
  gas, 91, 95
  Kelvin scale, 92
  mercury, 88
  scales compared, 90
  standard, 95
Thermostats, 90
Thompson, B. (*Rumford*), 86
Thomson, J. J., 344
Thomson, W. (*Kelvin*), 92
Thuras, A. L., 334
Time, unit of, 25
Tone quality, 188, 198
Torricelli, E., 73
Tower of Pisa, 29
Transverse vibrations:
  in plates, 151
  in rods and bars, 147
  in strings, 152
Traveling deformation, 154
Truth, search for, 7
Tuning fork, 151

U

Uncertainty principle, 364
Uniformly accelerated motion, 30
Uniform motion, 26, 32
Uranium, 353

V

Value (color), 279
Vapor pressure, 117
Velocity, 26
  terminal, 47
Ventilation, 126
Vibrating strings, 152
  law of, 208
Vibration:
  amplitude, 141
  complex, 153
  forced, 147
  in air columns, 155
  longitudinal, in rods, 154
  mechanical, 138
  patterns, 149
  period and frequency, 142
  sympathetic, 144
  transverse:
    in plates, 151
    in rods and bars, 147
    in strings, 152
Vibrato, 230
da Vinci, Leonardo, 28
Violin, 209
Virtual image, 284
Vision, 273, 295
Voice, 219
Volta, A., 327
Voltmeter, 326, 336
Volt, the, 336
Vowel sounds, 223, 225

W

Washboard road, 146
Water:
  pressure, 108
  supply, 106
Watt, James, 52
Watt, the, 338
Wave length, 159
Wave form, 191, 224
  of musical instruments, 192
Waves, 158
  demonstration, 158
  in free air, 168
  interference, 162
  light, 241
  longitudinal, 159
  on ropes, 160
  reflection, 161
  sound, 171
  standing, 154, 163
  transverse, 159
Weighing machine, 60

Weight, 9, 11
Wente, E. C., 334
Wheatstone, C., 226
Wilson, C. T. R., 354
Wind instruments, 211
Windlass, 61
Woolf, Stanley, and *Sette*, 230
Work:
  and heat, 97
  defined, 52
  measurement, 51
  unit, 52

X

X-rays:
  characteristic radiation, 351
  Coolidge tube, 347
X-rays (*cont.*):
  general radiation, 350
  hard and soft, 351
  nature, 348
  photograph, 347
  wave length, 353
Xylophone, 150

Y

Yard:
  American, 10
  imperial, 9
Young, Thomas, 241

Z

Zero, absolute, 92