# QUANTUM MECHANICS

VOLUME I

## OLD QUANTUM THEORY

.

## SERIES IN PHYSICS

General editors:

J. DE BOER Professor of Physics, University of Amsterdam

H. BRINKMAN Professor of Physics, University of Groningen

H. B. G. CASIMIR Director of the Philips Research Laboratories, Eindhoven

Some books of related interest in this Series are :

H. A. KRAMERS Quantum Mechanics

H. J. LIPKIN Beta Decay for Pedestrians

J. MCCONNELL Quantum Particle Dynamics

> A. MESSIAH Quantum Mechanics

P. ROMAN Theory of Elementary Particles

J. L. SYNGE Relativity: The Special Theory

J. L. SYNGE Relativity: The General Theory

A. H. WAPSTRA, G. J. NIJGH AND R. VAN LIESHOUT Nuclear Spectroscopy Tables

D. A. BROMLEY AND E. W. VOGT (editors) Proceedings of the International Conference on Nuclear Structure (Kingston, 1960)

> P. M. ENDT AND M. DEMEUR (editors) Nuclear Reactions, Volume 1

H. J. LIPKIN (editor) Proceedings of the Rehovoth Conference on Nuclear Structure

> K. SIEGBAHN (editor) Beta- and Gamma-Ray Spectroscopy

J. G. WILSON AND S. A. WOUTHUYSEN (editors) Progress in Elementary Particle and Cosmic Ray Physics, Volumes I-V

# QUANTUM MECHANICS

# VOLUME I

# OLD QUANTUM THEORY

ВY

# SIN-ITIRO TOMONAGA

The Tokyo University of Education, Japan

TRANSLATED FROM THE JAPANESE BY KOSHIBA



1962

NORTH-HOLLAND PUBLISHING COMPANY, AMSTERDAM

INTERSCIENCE PUBLISHERS INC., NEW YORK

No part of this book may be reproduced

in any form by print, photoprint, microfilm or any other means without written permission from the publisher

#### THIS BOOK WAS ORIGINALLY PUBLISHED

#### IN JAPANESE BY THE MISUZU PUBLISHING COMPANY, TOKYO

PUBLISHERS:

#### NORTH-HOLLAND PUBLISHING COMPANY, AMSTERDAM

INTERSCIENCE PUBLISHERS INC., NEW YORK

PRINTED IN THE NETHERLANDS

## PREFACE

THE intention of the author in writing this book is to serve the reader with a textbook by which he can learn quantum theory from the beginning. Younger students trying to appreciate quantum concepts often find some inconveniences, since there are few books in which the infancy of the theory, the course of how the quantum concepts have been discovered and generally recognized, is described in the light of the modern point of view.

For example, in most of the textbooks on quantum mechanics, only few pages are devoted to the description of Planck's theory of heat radiation or Bohr's theory of atomic structure, topics where the quantum concepts had originated and had revealed their universal nature. It may be said that today there is no need to start from Planck's discovery and to learn Bohr's old theory in order to master quantum mechanics. But the theory of heat radiation is in itself an important and interesting branch of physics and Bohr's theory still remains a useful way of describing intra-atomic events. Thus the author expects that this book may be appreciated not only by the beginners, but also by those who are already familiar with modern quantum mechanics.

The author has tried to show in this book how the present quantum mechanics has been built up instead of trying merely to introduce it in its established form. The task of theoretical physicists can be grossly divided into two categories. One category is to apply an established theory to those problems open to theoretical interpretation and to clarify the nature of the phenomena. The other is to make efforts to discover a new theory when the older one has been proven to fail. In the latter task which is no less important than the former task, examples which show how such a work had been performed in the past will be very helpful in guiding scientists by stimulating their imagination, to the discovery of a new theory. Experiences in building up quantum mechanics are highly instructive in this respect, since there the ways of solving the riddles set by nature were most abundantly shown by physicists of various types in various ways of thinking.

This book is thus written in the form of tracing the historical development of quantum theory. But the intention of the author is not to write a book of science history. The reader will find everywhere many anachronistic descriptions and distortions from the historical point of view. The author arbitrarily rearranged the materials so as to fit the purpose of elucidating as clearly as possible the intricate ways of thinking and of raising questions of many geniuses.

Since the emphasis has been laid on catching the physical points in dealing with various problems, mathematical rigor, generalities and elegantness were considerably sacrificed. This sacrifice was also unavoidable in order to make the description understandable in an intuitive way. The reader who demands those things will have no trouble in finding other appropriate textbooks.

Though this book is intended to be comprehensible even to the undergraduate students, it is assumed that the reader is reasonably familiar with some mathematics, analytical dynamics, and electromagnetic theory. Appendices will supply the beginners with what is lacking in the text.

The author owes a great deal to many other books and papers, but as the book is written mainly for the beginners, no list of references quoted in the text is given. It is hoped that the description is inclusive enough to be understood without referring to the original publications.

This book will be completed with three volumes. In this first volume, the earlier development of quantum theory, from Planck to Heisenberg through Einstein, Bohr and others, is treated. The second volume will deal with the development due to De Broglie, Schrödinger and others. It will describe how Heisenberg's matrix theory and the wave theory of De Broglie were unified into a comprehensive theory, which led to the settlement of the wave-particle dilemma. The third volume will be devoted to the presentation of the theory in a more systematic way rather than in the form of the historical development.

This book was written by Tomonaga first in Japanese and then translated into English by Koshiba. The manuscripts were read by Professors N. Fukuda and H. Umezawa to whom the author is much indebted for valuable comments. The English manuscripts were carefully checked by Professor J. B. French and Dr. M. Rosen to whom the author wishes to express his sincere acknowledgements. Thanks are

#### PREFACE

also due to the Misuzu Publishing Company, Tokyo, the publishers of the Japanese edition of this book, who so willingly, without any condition, gave the author the permission for publishing its English edition; and particularly to Mr. M. Matsui of the Company for his help in preparing figures and diagrams and in collection photographs anew for the English edition.

Sin-itiro Tomonaga

Tokyo, Japan March, 1962.

PREFACE																													v
---------	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	---

# CHAPTER 1

# THE DISCOVERY OF THE ENERGY QUANTUM

1.1	The beginning of the problem
1.2	The theory of specific heat
	(i) Ideal gas of monatomic molecules
	(ii) Ideal gas of diatomic molecules
	(iii) Crystalline substance
	(iv) Comparison with experiments
1.3	The specific heat of "vacuum"
1.4	The Rayleigh-Jeans formula
1.5	Wien's displacement law
	(i) The adiabatic invariant
	(ii) The pressure of black body radiation
	(iii) The Stefan-Boltzmann law
	(iv) The temperature change during an adiabetic volume change 29
	(v) Wien's displacement law
1.6	Wien's formula
1.7	Planck's formula
1.8	The quantum of energy
1.9	The quantum theory of specific heat

# CHAPTER 2

# THE CORPUSCULAR NATURE OF LIGHT

2.10	The hypothesis of the light quantum	43
2.11	The fluctuation of the energy in a hollow cavity	45
	(i) The fluctuation with the Rayleigh-Jeans formula	49
	(ii) The fluctuation with Wien's formula	50
	(iii) The fluctuation with Planck's formula	55
2.12	The photoelectric effect	57

-

2.13	The Compton effect.	64
	(i) The momentum of a light quantum	64
	(ii) The quantum theory of X-ray scattering	66
	(iii) Comparison with experiments	69
	(iv) The recoil electron $\ldots$	69
	(v) The Compton-Simon experiment	71
2.14	The photon gas	72
2.15	The corpuscular nature and the wave nature	75

# CHAPTER 3

# THE OLD QUANTUM MECHANICS

3.16	The structure of the atom	78
	(i) Zeeman effect	78
	(ii) The Thomson model and the Nagaoka model of the atom $\ldots$ .	82
	(iii) Scattering experiments with $\alpha$ -particles	83
	(iv) Rutherford's calculation	84
	(v) Rutherford's conclusion	89
	(vi) The difficulty in Rutherford's atom model	90
3.17	The spectrum of atoms	94
3.18	Bohr's theory	97
	(i) The fundamental idea	97
	(ii) The energy levels of hydrogen	100
	(iii) The size of the atom and its magnetic moment	105
3.19	The quantum condition	107
	(i) Ehrenfest's adiabatic hypothesis	107
	(ii) The generalized adiabatic invariant	110
	(iii) The adiabatic invariant for the system of one dimensional periodic	
	motion	110
	(iv) The quantum condition for a one-dimensional periodic system .	118
	(v) Generalization to multi-periodic systems.	119
	(vi) Example	122
3.20	The stationary states of the hydrogen atom	124
3.21	The azimuthal quantization	130
3.22	Experimental verifications of the stationary states	135
	(i) The Franck-Hertz experiment	136
	(ii) The Stern-Gerlach experiment	138
3.23	Bohr's correspondence principle	139
	(i) The mathematical foundations of the correspondence principle .	142
	(ii) The transition probability	147
£	(iii) The case of the harmonic oscillator	150
	(iv) The selection rule $\ldots$	151
	(v) The correspondence principle as a guiding principle	159
3.24	The quantization of rotational and translational motion	162
	(i) The quantization of rotational motion	163
	(ii) The quantization of translational motion	164

Т

## Chapter 4

# THE SHELL STRUCTURE OF THE ATOM

4.25	The theory of optical spectra	167
4.26	The X-ray spectrum	173
4.27	The shell structure of the atom and the periodicity	180
4.28	The multiplet structure of spectral terms and the inner quantum	
	number	189
4.29	The spin of the electron and the intrinsic magnetic moment	192
4.30	Pauli's principle	200

## CHAPTER 5

#### THE BIRTH OF MATRIX MECHANICS

5.31	The clue for solving the difficulties	203
5.32	Heisenberg's discovery	204
5.33	Matrix mechanics	224
5.34	The canonical equations of motion and Bohr's frequency relation	235
	(i) Mathematical preliminaries	236
	(ii) The law of energy conservation, and Bohr's relationship	241
5.35	The eigenvalue problem	244
	(i) The inverse of the law of energy conservation	245
	(ii) The unitary transformation	248
	(iii) The eigenvalue problem	252

#### APPENDIX I

Boltzmann's principle	2
(i) The trajectory of a dynamical system	2
(ii) The probability density	4
(iii) Liouville's theorem	6
(iv) Calculation of the probability density $\ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots 26$	9
(v) A system immersed in a thermal reservoir	4
(vi) Introduction of the temperature $\ldots \ldots \ldots \ldots \ldots \ldots \ldots 27$	7
APPENDIX II	
A simple model of a crystal	0
APPENDIX III	
Transverse oscillations of a continuous string	3
APPENDIX IV	
Dynamical treatment of electromagnetic oscillations	ô
APPENDIX V	
Pressure due to transverse vibrations of a string	0

#### APPENDIX VI

Adiabatic invariants of an oscillating system	294
APPENDIX VII	
Fluctuation of the energy in the oscillation of a string	298
APPENDIX VIII	
Another proof of $T = dJ/dW$ .	305
EXERCISES	307
INDEX	311



I. Electron tracks produced by X-ray in the cloud chamber. The longer ones are due to the photoelectric effect and the shorter, comma-like, ones are due to the Compton effect.



II. Tracks in the Compton-Simon experiment. In the upper photograph, the track is traced by a white curved line.



III. Photographs of fraction and interference of light by slits.(a) Images produced by one slit. (b) Images produced by two slits.



IV. Zeeman effect for the Zn 4680 Å line.



V. The spectrum of hydrogen.



VI. The spectrum of sodium. Upper side shows principal series and the lower 1st and 2nd subordinate series.



VII. The Stern-Gerlach experiment. The left shows the image of slit. in the presence of a magnetic field and the right without field.



VIII. The shift of K-lines.

#### CHAPTER 1

#### THE DISCOVERY OF THE ENERGY QUANTUM

#### § 1. The Beginning of the Problem

It was not too long ago that quantum mechanics was born into the family of physics. We may say that it began with Planck's discovery of what is called the energy quantum at the very beginning, 1900, of this century.

Until the end of the last century, physics had dealt mainly with natural phenomena which we could see, hear, or touch; just recall such branches of physics as mechanics, optics, acoustics, electrodynamics, thermodynamics, and so on. Toward the end of the century, the discoveries of strange phenomena, such as cathode rays, X-rays and radioactivity led the concern of physics from the world of our direct perception into the structure of the micro-world of molecules and atoms. The question now is: "What are the laws of physics that govern this micro-world?" In order to answer this question, experimental physicists engaged enthusiastically in studies of the newly discovered strange phenomena, while a group of theoretical physicists were attacking the problem of "black body radiation" in which they endeavoured to explain thermal properties of radiation in terms of the atomistic theory of heat.

Already in the last century, the study of the micro-world, molecule and atom, had been extensively carried out by Maxwell, Boltzmann and others. These people assumed that the micro-world of molecules and atoms is different only in its size from the macro-world and that the same physical laws govern both worlds. On this basis they tried to complete mathematically the atomistic theory which had long been programmed by the ancient Greek atomists. Their efforts were rewarded with remarkable successes.

The laws which govern the world of perception are those of Newtonian mechanics and Maxwell's electrodynamics. The theoreticians of those days considered molecules and atoms as objects which move according to the laws of Newtonian mechanics and which emit or absorb light in accordance with Maxwell's electrodynamics. Any directly perceptible object consists of a great number of molecules and atoms which obey these laws; and it is the gross effect of their individual behaviors that we directly perceive as the characteristics of this object.

In order to handle this gross effect of a great many individual motions, we have to apply a statistical method. What we see is not the individual motion of atoms or the light emitted by them, but instead, a kind of statistical average over a great number of molecules and atoms. For example, a vast number of gas molecules flying around in the container hit against the wall incessantly, each collision giving an impulse to it, and its average with respect to time is what we observe as the gas pressure. In the same way, what we measure as the temperature of a gas is the average kinetic energy of the gas molecules per degree of freedom.

This atomistic theory which presumes the validity of the laws of mechanics and electrodynamics in the micro-world together with the statistical considerations had considerable success in explaining various characteristics of matter. It turned out, however, that this theory did not work in the problem of black body radiation. Consequently people began to doubt the direct extrapolation of Newtonian mechanics or of Maxwell's electrodynamics into the micro-world as the working physical laws. In order to solve the problem of black body radiation, we had to construct a new picture of the physical world based on physical laws different from those of the above two theories. This new theory is called quantum theory, and is the subject of our volumes, while the Newtonian mechanics and Maxwell's electrodynamics are now to be called classical theories.

The problem of "black body radiation" is in short as follows: What color of light will an object emit when it is heated to a certain temperature? The light radiated from a heated body shows a red color first and becomes whiter as the temperature rises. The problem now is how to explain this phenomenon atomistically. The problem can be stated in a more physically precise form by the following idealization which will make the problem easier to handle. Consider a hollow cavity surrounded by walls at absolute temperature T. In thermal equilibrium, what spectrum of light will exist in the hollow cavity? The color we see when we look into a small opening of a flaming furnace depends on the spectrum. This spectrum was proved, first by Kirchhoff, to depend only on the temperature but on neither the substance, the shape nor the size of the wall.

There are various methods of treating this problem. It will be easier to understand if we arrive at the solution by generalizing the theory of specific heat rather than by the direct exposition of the method by which Planck first solved the problem. Therefore we begin with the theory of specific heat.

#### § 2. The Theory of Specific Heat

One of the important results of the statistical mechanics developed by Boltzmann is that we may interpret the thermal quantities of an object as some average of the dynamical quantities of molecular motions; the heat, or the temperature, for instance, is considered to be due to the random and complicated motion of a great many molecules which constitute the object. According to this idea, the more heat in an object, the more energy of random motion of its constituent molecules, and a higher temperature means a higher average kinetic energy per degree of freedom of the molecules. In this sense, heat is not a substance like phlogiston was considered in olden days, but instead a form of energy.

Let us suppose that an object is immersed in a thermal reservoir kept at a certain temperature; that both are in thermal equilibrium in the state in which the object is (at the same temperature as that of the reservoir) and that there is no exchange of thermal energy between them. Although we do not apparently observe the exchange of energy in this state, a certain amount of energy is from time to time received from or transfered to the reservoir through the boundary, when we look at the atomic motions of the object in detail. The apparent balance is due to the invisibility to us of this minute and random energy exchange. The actual energy of a molecule fluctuates with time around its average value. Boltzmann's statistical mechanics gives the following conclusions concerning this fluctuation.

When we look at an object atomistically, it is an extremely complicated dynamical system with an enormous number of degrees of freedom. The state of this dynamical system is specified by giving the coordinates  $q_1, q_2, q_3, \ldots, q_f$ , and their corresponding momenta  $p_1, p_2, p_3, \ldots, p_f$ , where f is the number of degrees of freedom. These coordinates and momenta vary with time in a very complicated manner, but satisfy the equations of motion which include also the coordinates and momenta of the reservoir. The energy of the system, as a given function of these 2*f* coordinates and momenta, also varies with time. It will have constant value only when the system is isolated from the surroundings; in our present case, however, it varies since the system is exchanging energy with the reservoir. The time variations of these coordinates, momenta, and energy are so complicated that it is not always possible to determine these quantities theoretically in each particular case. The statistical discussion, however, of these variations is in general possible. Namely we can discuss the probability distribution for coordinates and momenta, in the sense that they take certain values very frequently and certain other values very rarely in the course of the time variation of the dynamical system.

Let the energy of the system as a function of q's and p's be denoted by  $E(q_1, q_2, \ldots, q_f, p_1, p_2, \ldots, p_f)$ . It is then shown by statistical mechanics that the probability for the first coordinate, the second, ... to take some values between  $q_1$  and  $q_1 + dq_1$ ,  $q_2$  and  $q_2 + dq_2$ , ..., and, at the same time, for the first momentum, the second, ... to take some values between  $p_1$  and  $p_1 + dp_1$ ,  $p_2$  and  $p_2 + dp_2$ , ..., respectively, is given by

$$P(q_1q_2 \dots q_f \not p_1 \not p_2 \dots \not p_f) dq_1 dq_2 \dots dq_f dp_1 dp_2 \dots dp_f$$

$$= A \exp\left\{-\frac{1}{kT} E(q_1q_2 \dots q_f \not p_1 \not p_2 \dots \not p_f)\right\}$$
(2.1)
$$dq_1 dq_2 \dots dq_f dp_1 dp_2 \dots dp_f$$

where A is the factor which normalizes the total probability to unity. Thus A is given by

$$A = \frac{1}{\int \int \int \dots \int \exp\left\{-\frac{1}{kT} E(q_1 q_2 \dots p_f)\right\} dq_1 dq_2 \dots dp_f} .$$
 (2.2)

In Eqs. (2.1) and (2.2), T is the absolute temperature of the reservoir which is, in the state of thermal equilibrium, equal to the absolute temperature of the object itself. The constant k is determined by giving the unit of measurement of temperature and energy. In units of degree Kelvin and erg for temperature and energy, respectively, it becomes

$$k = 1.38 \times 10^{-16} \text{ erg/}^{\circ} \text{K}$$
.

#### CH. I, § 2] THE THEORY OF SPECIFIC HEAT

This constant k is called the Boltzmann constant. Eq. (2.1) is wellknown as Boltzmann's principle, a proof of which is given in Appendix I.

The following important conclusion can be drawn from Boltzmann's principle; if the total kinetic energy of the system is expressed as a sum of the kinetic energies in each degree of freedom and if, furthermore, the kinetic energy in each degree of freedom is proportional to the square of the momentum corresponding to that degree of freedom, then the average value of the kinetic energy per degree of freedom in the reservoir at temperature T is equal to  $\frac{1}{2}kT$ . Since according to this law the total kinetic energy is divided equally among the degrees of freedom without preference to any particular degree of freedom, it is called the law of equipartition of energy. This law also gives the atomistic interpretation of the temperature.

It is not difficult to derive this law from Boltzmann's principle. In our case, the kinetic energy of the system is in the form

$$E_{\rm kin} = \alpha_1 p_1^2 + \alpha_2 p_2^2 + \ldots + \alpha_s p_s^2 + \ldots + \alpha_f p_f^2 \qquad (2.3)$$

where the general term  $\alpha_s p_s^2$  in the kinetic energy corresponds to the s-th degree of freedom and  $\alpha_s$  is a constant; or more generally,  $\alpha_s$  can be an arbitrary but positive definite function of  $q_1, q_2, \ldots, q_f$ . Since the probability for q's and p's to take certain specified values is given by Eq. (2.1), the average value of the kinetic energy  $\alpha_s p_s^2$  corresponding to the s-th degree of freedom, which we denote by the expression  $\langle \alpha_s p_s^2 \rangle$ , becomes

$$\langle \alpha_s p_s^2 \rangle = A \int \int \dots \int \alpha_s p_s^2 \exp \left\{ -\frac{1}{kT} E(q_1 \dots q_f p_1 \dots p_f) \right\}$$
  
 $\mathrm{d}q_1 \dots \mathrm{d}q_f \mathrm{d}p_1 \dots \mathrm{d}p_f$ 

$$= A \iint \dots \int \alpha_s p_s^2 \exp \left\{-\frac{1}{kT} \left(\sum_{s=1}^f \alpha_s p_s^2 + V\right)\right\} dq_1 \dots dq_f dp_1 \dots dp_f$$

where V is the potential energy which depends only on the q's.

5

In order to obtain this average value, we first notice that the integral

$$\int \alpha_s p_s^2 \exp\left\{-\frac{1}{kT} \alpha_s p_s^2\right\} \mathrm{d}p_s$$

is readily reduced to

$$\frac{kT}{2}\int \exp\left\{-\frac{1}{kT}\,\alpha_s p_s^2\right\} \mathrm{d}p_s\,.$$

Therefore,

$$\langle \alpha_s p_s^2 \rangle = A \frac{kT}{2} \int \int \dots \int \exp\left\{-\frac{1}{kT} \left(\sum_{s=1}^f \alpha_s p_s^2 + V\right)\right\} dq_1 \dots dq_f dp_1 \dots dp_f.$$

With the use of Eq. (2.2), we arrive at the result

$$\langle \alpha_s p_s^2 \rangle = \frac{1}{2}kT$$
. (2.4)

This formula, Eq. (2.4), holds for any value of s from 1 to f and hence we have proved the law of energy equipartition. With the help of this law, we can readily calculate the specific heat of various substances. We start with the simplest case.

#### (i) IDEAL GAS OF MONATOMIC MOLECULES

We consider a unit mole of a monatomic ideal gas confined in a thermal reservoir of absolute temperature T. By a monatomic gas we mean a gas, like helium or neon, the molecule of which consists of a single atom. In such cases, we may consider the molecules as point masses. Furthermore, when such a gas is sufficiently diluted, we may neglect the potential energy between molecules. Hence we have only to consider as the energy of the gas the kinetic energy, which, as the readers are well aquainted with, has the form

$$E = E_{kin} = \sum_{n=1}^{N} \frac{1}{2m} \left( p_{nx}^2 + p_{ny}^2 + p_{nz}^2 \right)$$
(2.5)

where *m* is the mass of the molecule and  $p_{nx}$ ,  $p_{ny}$  and  $p_{nz}$  are the *x*-, *y*- and *z*-components, respectively, of the momentum of the *n*-th molecule. *N* is the total number of molecules per unit mole and its value is

$$N = 6.02 \times 10^{23}$$
.

6

# CH. I, § 2] THE THEORY OF SPECIFIC HEAT

Eq. (2.5) has the same form as Eq. (2.3). Therefore, the law of energy equipartition holds in this case, and the average value of the kinetic energy associated with each degree of freedom is

$$\frac{1}{2m}\langle p_{nx}^2\rangle = \frac{1}{2m}\langle p_{ny}^2\rangle = \frac{1}{2m}\langle p_{nz}^2\rangle = \frac{1}{2}kT.$$

The energy per gas molecule is, accordingly,

$$\langle E_n \rangle = \frac{1}{2m} \left( \langle p_{nx}^2 \rangle + \langle p_{ny}^2 \rangle + \langle p_{nz}^2 \rangle \right)$$
$$= \frac{3}{2} kT . \qquad (2.6)$$

The factor 3 on the right hand side of the equation corresponds to the fact that the number of degrees of freedom of a molecule, considered in this case as a point mass, is three.

We obtain the total energy, U, of a unit mole of the gas by multiplying Eq. (2.6) by N:

$$U = \frac{3}{2} NkT . \tag{2.7}$$

With the use of the conventional quantity R, defined by R = Nk, Eq. (2.7) can be written also as

$$U = \frac{3}{2} RT$$
 (2.8)

If we use calories instead of ergs as the energy unit, we have

$$R = 1.98 \text{ cal/}^{\circ} \text{K}$$
.

Since Eq. (2.7) is the thermal energy of a unit mole of the gas at temperature T, the necessary heat, C, to raise the temperature by one degree is

$$C = \frac{3}{2} R$$
  
= 2.97 cal/°K. (2.9)

This is the quantity we have been seeking, namely the specific heat of a unit mole of a monatomic gas.

#### (ii) IDEAL GAS OF DIATOMIC MOLECULES

The molecules of certain gases, such as oxygen, hydrogen, or nitrogen, etc., have two atoms which are connected in the shape of a



dumbell. The gas molecule of this kind has a moment of inertia for rotations around an axis perpendicular to the dumbell axis. It, therefore, has two more degrees of freedom for rotations, in addition to the three translational ones. The state of motion of a dumbell is specified by giving the five coordinates;  $x, y, z, \theta$  and  $\varphi$ , as well as their corresponding momenta  $p_x, p_y, p_z, p_{\theta}$ , and  $p_{\theta}$ ; x, y and z for specifying the position of the center of mass, the zenith angle  $\theta$  and azimuth

angle  $\varphi$  for specifying the direction of the dumbell

Fig. 1. Coordinates for the dumbbell shaped molecule.

axis as is shown in Fig. 1.

Denoting the moment of inertia by I, the energy is

$$E = \frac{1}{2m} \left( p_x^2 + p_y^2 + p_z^2 \right) + \frac{1}{2I} \left( p_\theta^2 + \frac{p_\phi^2}{\sin^2 \theta} \right).$$
(2.10)

Since it again has the form of Eq. (2.3), the law of energy equipartition allows us to write down the expression for the average energy per molecule as

$$\langle E \rangle = \frac{1}{2m} \left( \langle p_x^2 \rangle + \langle p_y^2 \rangle + \langle p_z^2 \rangle \right) + \frac{1}{2I} \left( \langle p_\theta^2 \rangle + \langle \frac{p_\phi^2}{\sin^2 \theta} \rangle \right)$$
$$= \frac{5}{2} kT ,$$

whence we obtain

$$U = \frac{5}{2}RT \tag{2.11}$$

and

$$C = \frac{5}{2}R = 4.95 \text{ cal/}^{\circ}\text{K}$$
 (2.12)

for the total energy and the specific heat per mole, respectively, of the gas.

The factor 5 reflects the fact that there are five degrees of freedom in the case of a diatomic molecule.

#### (iii) CRYSTALLINE SUBSTANCE

The specific heat of a crystal can also be obtained in a simple way according to the law of energy equipartition. We note that most of the solid substances consist of randomly oriented minute crystals. We consider a unit mole of a crystalline substance immersed in a thermal reservoir. Atoms of the crystal are located in a regular fashion, making what we call a space lattice. At the absolute temperature zero, these atoms are at rest at their own lattice points. At non-vanishing temperatures, however, these atoms oscillate randomly around their equilibrium positions. The oscillations of this kind, when their amplitudes are not too large, can be treated in a general way.

In case the amplitude is small, any arbitrary oscillation of the space lattice, no matter how random and irregular it may be, can be described as a superposition of a certain number of proper oscillations which vary sinusoidally with time. Thus, even a thermal motion can be expressed as a superposition of the fundamental and all their higher modes, with a proper choice of the amplitudes and phases of these "elastic" proper oscillations. The total energy of the thermal oscillation is given as a sum of the energies of all the superposed proper oscillations.

We now introduce what is in mechanics called the normal coordinate  $q_s$  and the corresponding momentum  $p_s$ , i.e. the coordinate and the momentum describing the s-th proper oscillation. In terms of this normal coordinate, the energy of the proper oscillation takes the simple form corresponding to a harmonic oscillator, namely

$$E_s = \alpha_s p_s^2 + \beta_s q_s^2 \tag{2.13}$$

where the first term is the kinetic energy while the second term is the potential energy of the oscillation. Some simple examples of this general statement are given in Appendix II and III.

From Eq. (2.13), we see that the law of energy equipartition holds in the dynamical system now under consideration. We hence have  $\langle \alpha_s p_s^2 \rangle = \frac{1}{2}kT$ . Our oscillator, however, differs from that of the molecules of an ideal gas in that its total energy contains the potential energy besides the kinetic energy. Thanks to the fact that in the case of a harmonic oscillator the potential energy is also of the form (2.3), its mean value  $\langle \beta_s q_s^2 \rangle$  equals  $\frac{1}{2}kT$ , giving rise to

$$\langle E_s \rangle = \langle \alpha_s p_s^2 \rangle + \langle \beta_s q_s^2 \rangle$$
  
=  $kT$ . (2.14)

Since we now have the average energy of a proper oscillation, we readily get the energy per unit mole of the crystal by summing up the energies of all the proper oscillations. Denoting the total number of the proper oscillations by f, it is then given by

$$\langle E \rangle = fkT$$
.

Now, our crystal contains a total of N atoms. In such a system as we are considering, the number of proper oscillations is equal to the number of degrees of freedom as will be shown in Appendix II and III. Hence,

$$f=3N$$
.

Therefore, the thermal energy per mole of the crystal is

$$U = 3NkT$$
  
= 3RT. (2.15)

From Eq. (2.15), the specific heat per mole of solid substances becomes

$$C = 3R$$
  
= 5.85 cal/°K . (2.16)

(iv) COMPARISON WITH EXPERIMENTS

We now compare the theoretical results derived in the preceding sections with the experimental values. In Table 1 one finds the theoretical values as well as the experimental ones. According to this table, one can see that, over the temperature range of the measurements, He and A have values close to  $\frac{3}{2}R = 2.97$  and that N<sub>2</sub> and O<sub>2</sub> have values approximately  $\frac{5}{2}R = 4.95$ . It has long been known experimentally as Dulong-Petit's law that the specific heat of solid substances has a value close to 3R = 5.85.

These experimental results all show that the law of energy equipartition holds rather well. However, when one extends the range of

Substance	Molecular heat (cal/deg.)	Temperature of measurements (°C)	Theoretical molecular heat (cal/deg.)
Monatomic gas $\left(egin{array}{c} He \ A \end{array} ight)$	(3.008 (2.949) (3.07 (2.86)	291 93 288 93	$\frac{1}{2}R=2.97$
Diatomic gas $\begin{pmatrix} N_2 \\ O_2 \end{pmatrix}$	( 4.984 ( 4.733 ( 4.985 ( 4.41	293 92 293 92	$rac{1}{2}R = 4.95$
Solid { Al Pb	5.83 5.43	a mean of 15 ~ 100	3R = 5.85

TABLE 1

temperature in measuring the specific heat, one finds that the experimental value becomes very different from the theoretical one. Fig. 2 shows the experimental results on the specific heat of a hydrogen gas



Fig. 2. The relation between the molar specific heat of hydrogen (experimental value) and the temperature. The abscissa shows the temperature (°K), the ordinate the molar specific heat (cal/R).

at various temperatures. From the specific heat-temperature curve shown in the figure one sees that at comparatively high temperatures the experimental value is close to  $\frac{5}{2}R$  but at lower temperatures it deviates substantially from this theoretical value. It is to be noted that at lower temperatures the experimental value is close to  $\frac{3}{2}R$ rather than to  $\frac{5}{2}R$ ; thus a H<sub>2</sub>-molecule at lower temperatures behaves like a monatomic molecule. Does the rotational degree of freedom have, at low temperatures, nothing to do with the specific heat?

The same situation arises also in the case of the specific heat of solid substances. In Fig. 3 is given the specific heat-temperature curve of lead. Dulong-Petit's law is observed here to hold at high temperatures



Fig. 3. The relationship between the molar specific heat of lead (experimental value) and the temperature. The abscissa shows the temperature ( $^{\circ}$ K), the ordinate the molar specific heat (cal/R).

but the specific heat differs greatly from it at low temperatures. This implies that at low temperatures not every degree of freedom receives enough energy to satisfy the energy equipartition law.

# § 3. The Specific Heat of "Vacuum"

We have seen that at low temperatures not all the degrees of freedom of a dynamical system receive the amount of energy allotted by the energy equipartition law. This fact is still more clearly observed in the case of "Vacuum". In the preceding sections we immersed various substances in the thermal reservoir. This time we put "Vacuum" in it; that is, we consider the hollow cavity itself surrounded by walls of temperature T. Although this hollow cavity is literally "hollow" and free of material, it contains the energy of random electromagnetic oscillations since it contains light inside. That is to say, it has some thermal energy and we can accordingly speak of the specific heat of "Vacuum" itself.

#### CH. I, § 3] THE SPECIFIC HEAT OF "VACUUM"

Unlike the elastic oscillations of a material object the electromagnetic oscillation cannot be handled by mechanics. However, Maxwell's laws which govern the electromagnetic phenomena can be expressed in a form similar to that of the mechanical laws which govern the elastic phenomena. When the laws are of similar mathematical structure then the phenomena themselves should also have some similarity. Thus, we can expect the same situations to occur in the specific heat of "Vacuum" as in that of material substances treated in the preceding sections.

More specifically, it goes as follows: any arbitrary oscillation of an electromagnetic field can be expressed as a superposition of its proper oscillations. Mathematically speaking, a state of the electromagnetic field is described in terms of the normal coordinates  $q_1, q_2, q_3, \ldots$  and their corresponding momenta  $p_1, p_2, p_3, \ldots$ . The energy of a state is given as a sum of terms of the form Eq. (2.13). Due to this fact, which will be explained in more detail in Appendix III and IV, it becomes possible to use the law of energy equipartition and we have an amount kT of energy in each degree of freedom. Hence, the total energy contained in the hollow cavity amounts to

$$E = fkT \tag{3.1}$$

where *f* is the number of the proper oscillations in this hollow cavity.

Unfortunately we here run into a difficulty. In the case of a crystal, the number f is related to the number, N, of atoms in it by the equation f = 3N and N as well as f is finite. The number of proper oscillations in our present case, however, is infinite since the electromagnetic oscillation is not an oscillation of a dynamical system consisting of a finite number of atoms, but instead an oscillation of a continuum, the "ether". In fact, as will be discussed in the next section, it is possible for an electromagnetic wave of any short wave length to exist in the hollow cavity while in a crystal an oscillation of wave length shorter than the inter-atomic distance can not occur. Consequently, the number f of Eq. (3.1) is infinite and hence E becomes so, too. This means that the specific heat of "Vacuum" is infinite.

The fact that E is infinite implies that the hollow cavity will absorb heat from the wall endlessly, before thermal equilibrium is reached. The heat energy thus absorbed is to be endlessly fed into proper oscillations of shorter and shorter wave lengths.

13

Nevertheless, the actual energy of a hollow cavity is finite. According to experiment, the energy per unit volume of a hollow cavity is related to its temperature by Stefan's law which says, denoting this energy density by U,

$$U = \sigma T^4 . \tag{3.2}$$

The constant  $\sigma$  is known from Stefan's measurement and is  $\sigma = 7.64 \times 10^{-15} \text{ erg/(cm}^3 \,^{\circ}\text{K}^4)$ . One now sees that the amount of energy distributed to each degree of freedom of electromagnetic oscillations is not in accord with the law of energy equipartition and is much smaller than is given by the law. Due to this very fact it was possible to derive the specific heat of a gas or a solid in the preceding sections and obtain fair agreement with experiments even with the neglect of the electromagnetic oscillations. At the temperatures where the experiments of the preceding sections were performed, the electromagnetic oscillations have scarcely any heat energy allotted to them. We have already seen that, at still lower temperatures even the mechanical oscillation receives energy in an amount smaller than  $\frac{1}{2}kT$ .

### § 4. The Rayleigh-Jeans Formula

Although it was found that the law of energy equipartition did not hold in the theory of specific heat, the law cannot be totally false since it gives specific heats which are in approximate agreement with experiments at not too low temperatures as one can see in Table 1. The specific heat-temperature curve of Fig. 2 suggests that in the case of the molecule the rotational degrees of freedom start to violate the equipartition law with decreasing temperature and the due share of energy is not received by them. The translational motions of the molecule seem to be still in accord with the equipartition law since the specific heat takes the value  $\frac{3}{2}R$  at low temperatures.

From the above considerations, one can now conjecture that, even in the case of a hollow cavity, certain degrees of freedom are, at the given temperature, still satisfying the energy equipartition law. The problem now is to find which degrees of freedom satisfy the law and which do not.

In order to decide this, one may measure separately the energies distributed to each proper oscillations instead of the total energy of the hollow cavity. This can be achieved by measuring the spectral intensity

#### CH. I, § 4] THE RAYLEIGH-JEANS FORMULA

distribution of the light coming from the hollow cavity through a spectrometer embedded in a small opening of the wall. By measuring spectroscopically the intensity of light of frequencies between  $\nu$  and  $\nu + d\nu$ , one experimentally knows the amount of energy distributed to the proper oscillations of frequencies between  $\nu$  and  $\nu + d\nu$ .

Assuming the law of energy equipartition, which we now know is against the facts, we can calculate this spectrum theoretically. For this purpose it is enough to know the number of proper oscillations having frequencies between v and v + dv, since the intensity of light in this frequency range can be obtained by simply multiplying it by kT.

Let this number of degrees of freedom be denoted by  $Z(\nu) d\nu$ . This statement may require some comments concerning the size of  $d\nu$ . This frequency range has to be small enough for the variation therein of the intensity to be negligible and, at the same time, it has to be large enough to contain a sufficient number of proper oscillations in the frequency range; but it must also be small enough, so that the number of degrees of freedom is proportional to  $d\nu$ . It can be shown that there exists such a  $d\nu$  for the radiation in a hollow cavity of practical size.

For the sake of simplicity, we consider a hollow cavity which is a cube of linear dimension L. The determination of Z(v) will be understood easily if we proceed as follows. We first treat the one-dimensional problem before proceeding to the three-dimensional one. Furthermore, we put aside for a while the problem of electromagnetic oscillations which is our final aim and take, as a model for it, the elastic oscillation of a string which is more familiar to us.

The question now is how many proper oscillations with frequencies between  $\nu$  and  $\nu + d\nu$  exist for a string of length L which is fixed at both ends.

The proper oscillations possible for a string of length L are, as are shown in Fig. 4, known to be those of wave length 2L, 2L/2, 2L/3, ... 2L/s, ... The proper oscillation of wave length 2L/s has (s - 1) nodes. If we denote the velocity of elastic waves on this string by c, the frequencies of these oscillations are given by

$$v_s = s \frac{c}{2L}, \quad s = 1, 2, \dots$$
 (4.1)

Thus the frequency of each proper oscillation is given by multiplying the fundamental frequency c/2L by 1, 2, 3, .... We see then that the

15

frequencies of the successive proper oscillations of the string have the same spacing,  $\Delta = c/2L$ . Therefore, the number of proper oscillations contained in the frequency range between v and v + dv is

$$Z(\mathbf{v}) d\mathbf{v} = \frac{d\mathbf{v}}{\Delta}$$
$$= \frac{2L}{c} d\mathbf{v} . \qquad (4.2)$$

A more detailed discussion on the subject is given in Appendix III.

In the above case of one-dimensional oscillations, it was possible to arrange the proper oscillations in order by giving each of them an



Fig. 4. The normal vibrations of a string.

Fig. 5. The normal vibrations of a square plane.

integral number s which has the physical meaning that (s-1) is the number of nodes of the oscillation.

In the three-dimensional case, however, the node of oscillation is not a point but is a plane which requires a set of three integers,  $s_x$ ,  $s_y$  and  $s_z$ , to specify it.

Since it is too complicated to show in the three-dimensional case the pattern of nodes for a given set of  $s_x$ ,  $s_y$  and  $s_z$ , we give the schematic representation of the two-dimensional oscillations. In Fig. 5 are shown the nodal patterns of oscillations of a square drum. They all have the common features of lattice patterns which can be arranged in order by giving each of them a set of two integers  $s_x$  and  $s_y$ . The three-dimensional case can be exemplified by adding one more dimension.

сн. 1, § 4]

#### THE RAYLEIGH-JEANS FORMULA

In this way three-dimensional oscillations can be specified by sets of three positive integers,  $s_x$ ,  $s_y$  and  $s_z$  and their frequencies are given by

$$v_{s_x s_y s_z} = \sqrt{s_x^2 + s_y^2 + s_z^2} \cdot \frac{c}{2L} \,. \tag{4.3}$$

Hence it is convenient to use a three-dimensional space to display them. We denote the three orthogonal coordinate axes of this space by x-, y- and z-axes and consider the points in the space having the coordinates  $x = (c/2L) s_x$ ,  $y = (c/2L) s_y$  and  $z = (c/2L) s_z$ . Since  $s_x$ ,  $s_y$ and  $s_z$  are integers, this point coincides with one of the space lattice points. The cell volume  $\Delta^3$  is  $(c/2L)^3$ . The distance from the origin to this point is

$$\sqrt{x^2 + y^2 + z^2} = \sqrt{s_x^2 + s_y^2 + s_z^2} \cdot \frac{c}{2L}.$$

By comparing this expression with Eq. (4.3), we see that this distance gives just the frequency of this particular oscillation.

When the frequency is expressed in this way, we can easily obtain the desired  $Z(\nu)$  geometrically. In order to find the number of oscillations with frequencies between  $\nu$  and  $\nu + d\nu$ , we have only to count the number of lattice points in the shell spanned by the two spheres of radius  $\nu$  and  $\nu + d\nu$  in the above x, y and z space. Since  $s_x$ ,  $s_y$  and  $s_z$ are all positive numbers, only the first octant is to be considered. The volume of the shell in the first octant being  $\frac{4}{5}\pi\nu^2 d\nu$ , we get  $Z(\nu) d\nu$ by dividing this volume by the volume of the lattice unit cell:

$$Z(\mathbf{v}) \, \mathrm{d}\mathbf{v} = \frac{\pi v^2 \, \mathrm{d}\mathbf{v}}{2\Delta^3} = \frac{4\pi L^3}{c^3} \, \mathbf{v}^2 \, \mathrm{d}\mathbf{v} \,. \tag{4.4}$$

In order to apply this result to the case of black body radiation, we have to take account of the following fact. Light, being a transverse wave, has the property of polarization and to each nodal pattern there correspond two kinds of oscillations of different polarization as explained in Appendix IV. In other words, there are two kinds of oscillations for each set of  $s_x$ ,  $s_y$  and  $s_z$  and, accordingly, Z(v) dv in this case has twice the value given by Eq. (4.4). Namely, the number of proper oscillations of light in the frequency range from  $\nu$  to  $\nu + d\nu$  is

$$Z(\mathbf{v}) \, \mathrm{d}\mathbf{v} = \frac{8\pi L^3}{c^3} \, \mathbf{v}^2 \, \mathrm{d}\mathbf{v} \,. \tag{4.5}$$

The spectrum of light in a hollow cavity can now readily be obtained from this expression multiplying Z(v) dv by kT. The following remarks are, however, necessary. Namely, in determining the number of proper oscillations, we considered a string fixed at both ends or a drum membrane fixed all around its circumference; i.e., the oscillators which have no exchange of energy with the external system. In the case of the electromagnetic oscillations, the above Z(v), therefore, applies when the walls of the hollow cavity are made of perfectly reflecting mirrors. Since this implies that there is no exchange of energy between the hollow cavity and the thermal reservoir, any initially given oscillation of the hollow cavity remains unchanged and thermal equilibrium will never be realized between the hollow cavity and the reservoir. We, hence, consider that there is a small window in the wall to make the thermal equilibrium attainable. If the window is small enough, it does not disturb appreciably the frequency or the nodal pattern of the proper oscillations and we can apply the same  $Z(\nu)$ obtained above for the case under consideration. Through this small opening in the wall, the energy is exchanged between the hollow cavity and the reservoir to bring about eventually a state of thermal equilibrium. Since according to the equipartition law, each proper oscillation receives the same amount kT of energy in thermal equilibrium, the amount of energy contained in the light of frequency between  $\nu$  and  $\nu + d\nu$  is

$$E(\mathbf{v}) d\mathbf{v} = Z(\mathbf{v}) kT d\mathbf{v}$$
$$= \frac{8\pi^3 L^3}{c^3} kT \mathbf{v}^2 d\mathbf{v} . \qquad (4.6)$$

This gives the spectral distribution of the energy of the whole hollow cavity and is to be directly related to the frequency distribution of the light intensity observed by means of a spectrometer.

Dividing Eq. (4.6) by the volume of the hollow cavity we get the spectral distribution of the energy per unit volume of the hollow cavity, which is

$$U(\nu) \, \mathrm{d}\nu = \frac{8\pi \, kT}{c^3} \, \nu^2 \, \mathrm{d}\nu \,. \tag{4.7}$$

In the above derivation, we considered specifically a cubic hollow cavity surrounded by mirror walls with a small window. Nevertheless, Eq. (4.6), or (4.7), is applicable more generally and is called the Rayleigh-Jeans formula.

This formula has been derived from the equipartition law and hence gives a infinite specific heat as was explained in the preceding section; it is clear then that the formula is incompatible with experiment. In fact, the integration of Eq. (4.6) over v, which gives the total energy of the hollow cavity, diverges to infinity.

Now then, what kind of spectrum is observed experimentally? In Fig. 6 are shown the observed spectrum of the black body radiation. As can be seen in the figure, the spectrum observed at a certain temperature has a maximum at a certain frequency beyond which the intensity of the radiation decreases with increasing frequency.

Since the Rayleigh-Jeans formula gives an intensity which is proportional to  $\nu^2$ and is thus increasing monotonically it obviously contradicts the experimental results. However, when we investigate the region of very small *v* in more detail, we find that the observed spectrum is in good agreement there with the Rayleigh-Jeans formula. For example, at temperatures around 2000 °K this for- Fig. 6. The spectrum of black mula agrees quite well with experiment in the frequency range around  $10^{13}$ /sec, i.e., in the far infra-red region.



body radiation. The abscissa shows the frequency, the ordinate intensity of radiation.

The frequency region of good agreement extends to still higher frequencies for higher temperatures.

According to the Rayleigh-Jeans formula, the spectral distribution of light is independent of the temperature T, while the total intensity is proportional to it. This implies that the color of the light from the hollow cavity does not depend on the temperature. In reality, however, not only does the total intensity of light increase more rapidly with increasing temperature, proportional to  $T^4$ , obeying Stefan's law, but also its spectrum changes with the temperature. Namely, at low temperatures the light of maximum intensity resides in a rather longwave-length region and with increasing temperature it shifts to short wave lengths. In fact, it is a daily common experience that an object emits infrared light at low temperatures and with increasing temperature red light, then white light, and finally it begins to radiate ultraviolet light. The equipartition law can not explain this daily common phenomenon.

It has become clear through the foregoing discussions that the energy equipartition law does not hold in general but it still is correct for oscillations of lower frequencies and that the frequency region where this law is applicable widens with increasing temperature. In other words, with decreasing temperature the degrees of freedom corresponding to large frequencies start violating the equipartition law and "die"; i.e., the energy is not duly distributed to these oscillations.

In calculating the specific heat of a gas or of solid matter we considered translational, rotational and oscillational degrees of freedom only and obtained rather good agreements with experiment. It is clear, however, that an atom, since it contains many electrons, has not only translational, rotational and oscillational degrees of freedom but also some internal degrees of freedom; but these internal degrees of freedom are "dead" at room temperatures owing to their high frequencies. When the temperature is raised a great deal the specific heat becomes larger since these internal degrees of freedom then begin to become excited. The fact that the specific heat of hydrogen gas approaches with decreasing temperature that of a monatonic gas can be interpreted in a similar way, as due to the "dying" of the rotational degrees of freedom at low temperatures

# § 5. Wien's Displacement Law

The fact that the energy equipartition law is incompatible with observation implies that we have no method for finding theoretically the spectrum of the black body radiation. It was Wien's theory which introduced some hope into this sad situation. Wien has shown ingeniously that, even though one cannot derive the spectrum theoretically, once the spectrum is known by some means at a certain
## сн. I, § 5]

temperature, the spectrum at any other temperature can be derived from it.

In this chapter, however, we deviate from the original method of Wien in arriving at his conclusion and follow a reasoning which seems more intimately related to the development of quantum theory. Before we start this main subject, a few preliminary considerations are needed.

## (i) THE ADIABATIC INVARIANT

The general definition of the adiabatic invariant will be given in a later section and we at present satisfy ourselves with the following definition which is sufficient for the special case to be considered here.

We consider a system oscillating sinusoidally. An example is a simple pendulum or a stretched string. Another example is the radiation in a hollow cavity provided that it is completely isolated and does not exchange any thermal energy with the reservoir. We then slowly change the shape of this dynamical system which is in the state of a certain proper oscillation. This can be done, for example, by gradually shortening the string of a pendulum, by changing the length or tension of the stretched string, or by pushing in one of the walls of the hollow cavity with something like a piston. It can be shown that, if the deformation is made very slowly, the dynamical system is still in state of a proper oscillation after the deformation.

Naturally the frequency v of a proper oscillation varies with the deformation of the dynamical system. The amplitude, and hence the energy

E. likewise varies. In other words, a certain amount of work is done in the process of deformation. The important thing to be noted here is that the rate of change of v is equal to that of E so that the ratio E/vremains constant during the slow deformation. This quantity  $E/\nu$  is called the adiabatic invariant.

When, at the beginning, the system is not in a state of proper  $_{\rm Fig. 7. The vibration of a pendulum and$ oscillation, it is described by a superposition of various proper



its adiabatic deformation.

oscillations and  $E/\nu$  is invariant for each proper oscillation during the deformation process.

We now demonstrate the above fact for the case of a simple pendulum. In Fig. 7 is shown a pendulum consisting of a weight and a string which is supported by a pulley. We pull this string slowly while the pendulum is in oscillation; then we can see experimentally that the amplitude as well as the frequency of the oscillation increases with the shortening of the hanging string.

We first calculate the amount of work to be done in pulling the string. This gives the change of the energy of this oscillating pendulum. Let the angle of the pendulum from the vertical, the length of the string and the mass of the weight be denoted by  $\theta$ , L and m, respectively. We have for the equation of motion

$$mL\ddot{\theta} = -mg\sin\theta$$

where g is the acceleration of gravitation.

On the other hand, the energy of the pendulum is

$$E = \frac{1}{2} m L^2 \dot{\theta}^2 + mgL \left(1 - \cos \theta\right)$$

where the first term is the kinetic energy and the second is the potential energy. We next find the tension of the string T. This can be obtained by adding the component of gravitational force in the direction of the string,  $mg \cos \theta$ , and the centrifugal force,  $mL\dot{\theta}^2$ , thus

$$T = mg\cos\theta + mL\theta^2.$$

If we restrict ourselves to the case of small amplitude, we can expand sin  $\theta$  and cos  $\theta$  in terms of  $\theta$  and neglect higher order terms. The above three relations then become

$$mL\ddot{\theta} = -mg\theta \tag{5.1}$$

$$E = \frac{1}{2} m L^2 \dot{\theta}^2 + \frac{1}{2} m g L \theta^2$$
 (5.2)

$$T = mg (1 - \frac{1}{2} \theta^2) + mL\dot{\theta}^2.$$
 (5.3)

Eq. (5.1) can readily be solved to give

$$\theta = a\cos\left(2\pi\nu t + \delta\right) \tag{5.4}$$

# сн. I, § 5]

where a and  $\delta$  are the constants of integration and signify the amplitude and phase, respectively, of the oscillation. The frequency v is given by

$$\nu = \frac{1}{2\pi} \sqrt{\frac{g}{L}} \,. \tag{5.5}$$

Substitution of Eq. (5.4) into Eq. (5.3) yields the tension as a function of time,

$$T = mg - mga^{2} \left\{ \frac{1}{2} \cos^{2} (2\pi \nu t + \delta) - \sin^{2} (2\pi \nu t + \delta) \right\}.$$
 (5.6)

Due to the trigonometrical terms in the curly bracket the tension has, besides the gravitational force mg, a component which varies with time.

Suppose L is varied by pulling the string and the variation be denoted by  $\delta L$ . Then the work  $\delta W$  done on the pendulum is given by the product of  $(-\delta L)$  and the average tension  $\langle T \rangle$  during the pulling, provided that the pulling is performed slowly enough for the pendulum to have time to make a considerable number of oscillations. Thus,

$$\delta W = -\langle T \rangle \, \delta L \tag{5.7}$$

where  $\langle T \rangle$  is the time average of Eq. (5.6) and is given by

$$\langle T \rangle = mg + \frac{1}{4} mga^2 . \tag{5.8}$$

From the work done on the pendulum, we can now calculate the change in the oscillational energy of the pendulum. Of this work the amount  $mg \ \delta L$  is spent in lifting the pendulum as a whole, i.e., in increasing the gravitational potential energy. The net increase of the oscillational energy is then obtained by subtracting this amount from  $\delta W$  of Eq. (5.7) and is

$$\delta E = -\frac{1}{4} mga^2 \,\delta L \,. \tag{5.9}$$

Observing that substitution of Eq. (5.4) into Eq. (5.2) yields

$$E = \frac{1}{2} mga^2 L \tag{5.10}$$

and that the change of frequency,  $\delta \nu$ , due to the change of L by  $\delta L$  is from Eq. (5.5)

$$\delta \mathbf{v} = -\frac{\mathbf{v}}{2} \frac{\delta L}{L} \,,$$

we can conclude from Eq. (5.9) that

$$\frac{\delta \nu}{\nu} = \frac{\delta E}{E} \,. \tag{5.11}$$

Eq. (5.11) now readily gives by integration

$$\frac{E}{\nu} = \text{const.} \tag{5.12}$$

This is the result we wanted to derive. The proof of the law of adiabatic invariance for the oscillating system in general is given in Appendix VI.

## (ii) THE PRESSURE OF RADIATION

We consider the radiation filling a hollow cavity. In whatever oscillatory state this radiation may be, it exerts some pressure constantly on the walls of the hollow cavity. We will show that this pressure is closely related to the energy density of the radiation inside.

It may sound somewhat strange that radiation exerts pressure on the wall, but this is due to the same reason which made the average tension of a string,  $\langle T \rangle$ , different from the mere gravitational tension mg. There it was  $\langle T \rangle --mg$ , and hence  $\frac{1}{4}mga^2$  (according to Eq. (5.8)), that we might call the "pressure" of the oscillation. For a more detailed discussion of pressure in the case of an oscillating string, the reader should refer to Appendix V.

In order to obtain the relation between the pressure and the energy density of the black body radiation, we calculate the amount of work to be done to compress the hollow cavity, assuming naturally that the walls are plastic. For the sake of simplicity, we assume that the hollow cavity is always kept in the form of a cube.

The calculation of the work is essentially the same as that of the work in pulling the string of a pendulum as explained above. The work done,  $\delta W$ , gives in this case directly the sum of the energy increments,  $\delta E_{\star}$ , in each proper oscillation, i.e.,

$$\delta W = \sum_{s} \delta E_{s} \tag{5.13}$$

where s stands for three integers  $s_x$ ,  $s_y$  and  $s_z$ .

сн. 1, § 5]

On the other hand, since  $E_s/\nu_s$  is an adiabatic invariant,  $\delta E_s$  is to satisfy a relation similar to Eq. (5.11)

$$\frac{\delta E_s}{E_s} = \frac{\delta \nu_s}{\nu_s}.$$
(5.14)

We already have the relationship between the frequency  $\nu_s$  and the dimension L of the cube, namely, according to Eq. (4.3)

$$v_s = s \frac{c}{2L}$$

where s is to be understood to mean  $\sqrt{s_x^2 + s_y^2 + s_z^2}$ . Consequently we have

$$\frac{\delta v_s}{v_s} = -\frac{\delta L}{L} \tag{5.15}$$

and furthermore, by combining this with Eq. (5.14), we arrive at

$$\frac{\delta E_s}{E_s} = -\frac{\delta L}{L} \,. \tag{5.16}$$

Denoting the volume of the cubic space by V we rewrite Eq. (5.16) as

$$\frac{\delta E_s}{E_s} = -\frac{1}{3} \frac{\delta V}{V} \tag{5.17}$$

since  $\delta V = 3L^2 \delta L$ . This gives the energy increment of the s-th proper oscillation when we change the volume of the hollow space adiabatically by  $\delta V$ .

Substitution of Eq. (5.17) into Eq. (5.13) immediately gives us the amount of work  $\delta W$  necessary to change the volume of the hollow cavity by  $\delta V$ . Thus

$$\delta W = -\frac{1}{3} \frac{\sum E_s}{V} \delta V \,. \tag{5.18}$$

Since  $\sum_{s} E_{s}$  is the total energy contained in the hollow cavity,  $\sum_{s} E_{s}/V$  is the energy per unit volume of the hollow cavity, i.e., the energy density. In terms of this energy density, U, Eq. (5.18) can be written also as

$$\frac{\delta W}{\delta V} = -\frac{1}{3} U \,. \tag{5.19}$$

25

This equation (5.19) implies that the average pressure, P, of the black body radiation on the surrounding six walls is given by

$$P = \frac{1}{3} U$$
. (5.20)

This result has been verified experimentally by means of a very precise torsion balance.

#### (iii) THE STEFAN-BOLTZMANN LAW

It was stated in a previous section that the energy density of black body radiation in thermal equilibrium with a reservoir of temperature T is proportional to  $T^4$ .

This law originally was an empirical one discovered by Stefan. Boltzmann later succeeded in deriving it theoretically. This can be done by combining the mechanical result, Eq. (5.19), with thermodynamical considerations which determine the temperature of the system. With this aim in mind we consider a thermodynamic cycle as shown in Fig. 8, which probably reminds the reader of the famous Carnot cycle.

Let us imagine a hollow cavity of perfectly reflecting walls with a small window. We immerse this into a heat reservoir of temperature  $T_1$ ; the hollow cavity will reach the state of thermal equilibrium, the state 1 say, which is specified by the temperature  $T_1$ , the volume  $V_1$  and the energy density  $U_1$ . We then isolate the hollow cavity from the reservoir, by closing the window, and compress it adiabatically until its volume becomes  $V_2$ . We call this state the state 2. As the result of the adiabatic compression the energy density and the temperature become higher in the state 2. Let these be denoted by  $U_2$  and  $T_2$  respectively.

According to the consideration we carried through in calculating the radiation pressure, the amount of work to be done in the process of changing the volume adiabatically is given by Eq. (5.19). This work results in a change of the energy in the hollow cavity. The change of the energy density, during an adiabatic volume change, therefore, must satisfy the equation

$$\frac{\delta(UV)}{\delta V} = -\frac{1}{3} U$$

which can readily be integrated to give

$$U = \frac{A}{V^{\sharp}}.$$
 (5.21)

The integration constant A is unchanged during the adiabatic process. In words, the energy density is inversely proportional to the 4 power of the volume, and in our process

$$U_1 V_1^4 = U_2 V_2^4 = A . (5.22)$$

We now immerse the hollow cavity in a thermal reservoir of temperature  $T_2$  and, by opening the window, allow the volume to expand isothermally to  $V_3$ . Since it has been proved by Kirchhoff that the

energy density of a hollow cavity is a function of temperature only, the energy density of this state, the state 3, is still  $U_2$ . In the process of this expansion, the hollow cavity has to absorb some heat from the reservoir since the internal energy of the state 2 is  $V_2U_2$  while that of the state 3 is  $V_3U_2$ , and  $V_3 > V_2$ . Fur- Fig. 8. The UV diagram for the thermore it performs work in expanding the volume from  $V_2$  to  $V_3$  with the pres-



hollow cavity.

sure  $\frac{1}{3}U_2$ . Thus the amount of heat absorbed,  $Q_{2\rightarrow 3}$ , during this process is

$$Q_{2 \to 3} = \frac{4}{3} U_2 (V_3 - V_2) . \tag{5.23}$$

The hollow cavity is now isolated again from the reservoir and is expanded adiabatically until its temperature sinks to the original value  $T_1$ . This state is called the state 4, the volume of which is denoted by  $V_{\mathbf{a}}$ . Since again the energy density of black body radiation is a function of temperature only, the energy density in the state 4 is equal to  $U_1$ , and, according to a consideration similar to that used in arriving at Eq. (5.22), we obtain

$$U_2 V_3^{\ddagger} = U_1 V_4^{\ddagger} = A'. \tag{5.24}$$

We now proceed to the last step of the cycle, namely we immerse the hollow cavity in the thermal reservoir of temperature  $T_1$  and compress it isothermally until its volume becomes  $V_1$ ; i.e. until it comes back to the original state 1. During this process,  $4 \rightarrow 1$ , it liberates some heat, or absorbs some negative heat, the amount, by the same argument used in deriving Eq. (5.23), being given by

1

$$Q_{4\to 1} = \frac{4}{3} U_1 (V_1 - V_4) . \tag{5.25}$$

In the meantime, Carnot's principle, or, in the modern terminology, the second law of thermodynamics, requires the following relation to be satisfied, in order that we should not have a perpetual motion of the second kind,

$$\frac{Q_{2-3}}{T_2} + \frac{Q_{4-1}}{T_1} = 0 \; .$$

By means of Eq. (5.23) and Eq. (5.25), this becomes

$$\frac{U_2}{T_2}(V_3 - V_2) + \frac{U_1}{T_1}(V_1 - V_4) = 0.$$
 (5.26)

Since we have to find a relation between energy density and temperature we eliminate the V's from Eq. (5.26) by making use of Eqs. (5.23) and (5.24). Then the result is

$$\frac{U_2^{\frac{1}{2}}}{T_2} (A'^{\frac{1}{2}} - A^{\frac{1}{2}}) + \frac{U_1^{\frac{1}{2}}}{T_1} (A^{\frac{1}{2}} - A'^{\frac{1}{2}}) = 0$$
$$\frac{U_1^{\frac{1}{2}}}{T_1} = \frac{U_2^{\frac{1}{2}}}{T_2}.$$

or

Since the last relationship should hold for arbitrary states 1 and 2, it requires that

$$\frac{U^4}{T} = \text{const.}, \qquad (5.27)$$

which is just Stefan's law.

Stefan's law is thus derived by requiring that the second law of thermodynamics be satisfied. The proportionality constant in Eq. (5.27), the so-called Stefan's constant, however, remains undetermined by such a thermodynamical requirement only.

СН. I, § 5]

# (iv) THE TEMPERATURE CHANGE DURING AN ADIABATIC VOLUME CHANGE

From the above law,  $U(T) \propto T^4$ , we can derive a formula for the temperature change of the hollow cavity during an adiabatic compression or expansion. The energy density of a hollow cavity changes during the adiabatic deformation according to Eq. (5.21), i.e.

$$U \propto V^{-\frac{4}{3}}$$
.

On the other hand, it has to satisfy  $U \propto T^4$ . We thus obtain

$$T \propto V^{-\frac{1}{3}}.$$
 (5.28)

which says that the temperature of a hollow cavity changes proportionally to the inverse of the cube root of its volume in the process of adiabatic compression or expansion.

## (v) WIEN'S DISPLACEMENT LAW

We are now ready to start on the main subject of this chapter. We consider a hollow cavity in thermal equilibrium at temperature T. When we compress or expand this hollow cavity adiabatically, the temperature changes according to Eq. (5.28), and in this way we can construct a black body radiation of any temperature from the radiation at temperature T.

According to the theorem of the adiabatic invariant,  $E_s$ , the energy in the s-th proper oscillation of the hollow cavity, changes in such a manner that  $E_s/v_s$  remains invariant for each proper oscillation.

Meanwhile it was derived above that the temperature varies proportionally to the inverse of the cube root of the volume in adiabatic processes. Furthermore, as can be seen from Eq. (4.3), the variation of  $v_s$  is also inversely proportional to the cube root of the volume,  $V^4$ . Hence we can conclude that in adiabatic processes the frequency variation of each proper oscillation is proportional to the temperature; i.e., we have

$$\frac{v_s}{T} = \text{invariant}$$
 (5.29)

for each proper oscillation.

If we assume that the energy distribution of a hollow cavity in thermal equilibrium at temperature T is given by  $E_s = f(v_s)$ , we have

in general to expect that this function  $f(v_s)$  will vary with temperature. As a matter of fact if the energy equipartition law holds f(v) is simply kT and is obviously temperature dependent. However, from the above consideration that led us to the adiabatic invariance of  $v_s/T$  as well as  $E_s/v_s$ , it follows that if we write the energy distribution as

$$\frac{E_s}{v_s} = F\left(\frac{v_s}{T}\right)$$

the functional form of this function F should be independent of the temperature, since neither  $E_s/\nu_s$  nor  $\nu_s/T$  is temperature dependent. Hence the energy distributed to the s-th proper oscillation may be written in general as

$$E_s = F\left(\frac{\nu_s}{T}\right) \cdot \nu_s \tag{5.30}$$

provided that Carnot's principle holds, whereas the energy equipartition law need not be satisfied here.

If the energy distributed to each proper oscillation is given by Eq. (5.30), we can obtain the spectral intensity distribution of hollow cavity radiation by multiplying this by  $Z(\nu) d\nu/V$ , as we did in deriving the Rayleigh-Jeans formula, namely

$$U(\mathbf{v}) \,\mathrm{d}\mathbf{v} = \frac{8\pi}{c^3} F\left(\frac{\mathbf{v}}{T}\right) \cdot \mathbf{v}^3 \,\mathrm{d}\mathbf{v} \,. \tag{5.31}$$

The Rayleigh-Jeans formula is contained in Eq. (5.31) as a special case with

$$F(x)=\frac{k}{x}.$$

Through the above consideration, it has become clear to what extent we can determine the spectrum of a black body radiation without having recourse to the energy equipartition law, though the specific form of the function F(x) can not be determined by such considerations. However, if Eq. (5.31) holds, the determination of the spectrum at a certain temperature leads to the determination of the functional form of F(x) and hence the spectrum can be predicted, for any other temperature. We have now obtained a formula which tells us how the

#### WIEN'S FORMULA

spectrum of the black body radiation changes with varying temperature and especially how the intensity maximum of the spectrum shifts. This Eq. (5.31) is called Wien's displacement law and has been completely verified by experiment. In deriving Eq. (5.31), we have used Stefan's law as well as the adiabatic invariance of  $E_s/v_s$  of an oscillating system. Stefan's law is quite well established experimentally, even if one may not feel quite certain about its theoretical derivation. The agreement with experiment of Eq. (5.31) thus seems to suggest that the adiabatic invariants might in some form or other survive in the future theory even if Newtonian mechanics should prove not to be satisfactory. The energy equipartition law, on the other hand, cannot survive at all.

# § 6. Wien's Formula

It was a great achievement to find that the energy spectrum of the hollow cavity radiation has a form given by Eq. (5.31). The next problem of determining the functional form of F(x), however, still remained unsolved by considerations only of the thermodynamic laws or of the adiabatic invariants.

As to the functional form of F(x), we can say at present only that it has to tend asymptotically to zero stronger than  $1/x^5$  as x approaches infinity, since otherwise the total energy, which is obtained by integrating Eq. (5.31) with respect to  $\nu$  from zero to infinity, would diverge to infinity.

A number of people proposed various forms for F(x), some as empirical formulae and others from tentative theories, without reaching a satisfactory one. From his theoretical considerations, Wien assumed the following form for F(x):

$$F(x) = k\beta e^{-\beta x} \tag{6.1}$$

and hence

$$U(\mathbf{v}) \,\mathrm{d}\mathbf{v} = \frac{8\pi k\beta}{c^3} \,\mathrm{e}^{-\rho_{\mathbf{v}/T}} \,\mathbf{v}^3 \,\mathrm{d}\mathbf{v} \tag{6.2}$$

where k is Boltzmann's constant.

It was found by experimentalists that with proper adjustment of the constant  $\beta$  in the formula it can well reproduce the experimental results in the range

$$\frac{\nu}{T} \gtrsim 10^{11} \, \mathrm{sec}^{-1} \, \cdot \, {}^{\circ}\mathrm{K}^{-1}$$
.

This formula, however, fails to agree with experiment outside the above frequency range. In fact, Eq. (6.2) varies as  $\nu^3$  for small  $\nu/T$ , while the Rayleigh-Jeans formula gives a good agreement there with the experiment implying that  $U(\nu)$  has to tend to zero as  $\nu^2$ .

Not only could Wien's formula not reproduce the experimental results, but also his theory was not well received, since it deviated entirely from Maxwell's theory. His theory was indeed a modification of the classical theory but it was considered too arbitrary, involving unnatural assumptions. However, it could not be denied that his formula conveyed a part of the truth since it agreed quite well with experiments in the range  $\nu/T \gtrsim 10^{11} \sec^{-1} \cdot {}^{\circ}\mathrm{K}^{-1}$ , as a matter of fact better than any other formula proposed up to that time.

The correct theoretical meaning of Wien's formula will become clear as we proceed and we will come back to this point in § 14.

## § 7. Planck's Formula

It has thus been found that the Rayleigh-Jeans formula gives good agreement with experiment for small  $\nu/T$  while for large  $\nu/T$  Wien's formula takes its place. Planck then considered an interpolation formula between the above two.

According to him, F(x) takes the form

$$F(x) = \frac{k\beta}{e^{\beta x} - 1}.$$
 (7.1)

F(x) as given by Eq. (7.1) does indeed take the form  $k\beta e^{-\beta x}$  for large x and the form k/x for small x, corresponding to Wien's and Rayleigh-Jeans' formula, respectively.

Assuming F(x) to be of the form of Eq. (7.1), the spectrum of the black body radiation becomes

$$U(\nu) \, \mathrm{d}\nu = \frac{8\pi k\beta}{c^3} \frac{1}{\mathrm{e}^{\beta\nu/T} - 1} \, \nu^3 \, \mathrm{d}\nu \,. \tag{7.2}$$

This is Planck's well-known formula. Usually we write

$$h = k\beta . \tag{7.3}$$

This constant h has the dimension of action and its value has been determined by comparison of Eq. (7.2) with experiment to give

$$h = 6.62 \times 10^{-27} \,\mathrm{erg} \cdot \mathrm{sec}$$
 (7.4)

With this constant h, Eq. (7.2) can also be written as

$$U(\nu) \, \mathrm{d}\nu = \frac{8\pi h}{c^3} \frac{1}{\mathrm{e}^{h\nu/kT} - 1} \, \nu^3 \, \mathrm{d}\nu \,. \tag{7.5}$$

The constant h is called Planck's constant or the quantum of action.

This formula has been verified experimentally by a number of people and found to reproduce the experimental data very well in the entire temperature and frequency range observed, while the other formulae did not. This can be seen in Fig. 9.

The integration of Eq. (7.5) with respect to  $\nu$  from zero to infinity gives the total energy density of the black body radiation, which is

$$U = \frac{8\pi h}{c^3} \int_0^\infty \frac{\nu^3 \,\mathrm{d}\nu}{\mathrm{e}^{h\nu/kT} - 1}$$
$$= \frac{8\pi k^4}{c^3 h^3} T^4 \int_0^\infty \frac{x^3 \,\mathrm{d}x}{\mathrm{e}^x - 1}.$$
(7.6)

This is nothing but Stefan's formula and Stefan's constant is thus given in this theory by

$$\sigma = \frac{8\pi}{c^3 h^3} k^4 \int_0^\infty \frac{x^3 dx}{e^x - 1}$$
(7.7)

the value of which is in good agreement with experiment.

Eq. (7.5) also gives the distribution law which is to replace the energy equipartition law. Namely,

the distribution law is obtained by dividing this formula Eq. (7.5) by  $Z(\nu) d\nu/L^3$  of Eq. (4.5). The result is that the amount of energy



Fig. 9. The relationship between the intensity of black body radiation and the frequency. The abscissa shows the frequency, the ordinate the intensity, the curves Planck's formula, the small circles measured values. actually distributed to a degree of freedom of frequency v is

$$\langle E_{\nu} \rangle = kTP\left(\frac{h\nu}{kT}\right),$$
 (7.8)

whereas according to the energy equipartition law the same amount of energy kT was to be distributed to each degree of freedom independently of the frequency. The function P(x) is defined by

$$P(x) = \frac{x}{\mathrm{e}^{x} - 1} \tag{7.9}$$

and has values close to, or smaller than, unity for small, or large x, respectively. Accordingly, for such degrees of freedom which have small hv/kT the equipartition law holds approximately, while for other degrees of freedom the amount of energy given is not as much as the equipartition law requires.

# § 8. The Quantum of Energy

Planck has thus found a formula which is in excellent agreement with experiment. If he had been contented with this, his work might have been considered merely as a fortuitous discovery. He worked hard to find the fundamental meaning of this formula and finally, after several weeks of struggling, arrived at the concept of the energy quantum.

His idea goes as follows. No material substance is a continuum and every one of them consists of atoms which cannot be divided any further. Energy might also not be a continuum capable of being divided without end but might consist of a kind of atom. Planck called this elementary quantity of energy "the energy quantum".

When the energy consists of quanta, the energy equipartition law, Eq. (2.14), does not hold and the law to replace it is derived as follows. The energy of an oscillating system is, as in Eq. (2.13),

$$E = \alpha p^2 + \beta q^2 \,. \tag{8.1}$$

The average energy is according to Boltzmann's principle

$$\langle E \rangle = A \int \int (\alpha p^2 + \beta q^2) e^{-E/kT} dp dq$$
 (8.2)

сн. 1, §8]

where A is given by

$$\frac{1}{A} = \iint e^{-E/kT} dp dq . \qquad (8.3)$$

Straight forward calculation obviously leads to the same result as given by Eq. (2.14), but we have now to take into account in integrating Eq. (8.2) or Eq. (8.3) the fact that the energy is not a continuous quantity. Thus, since the energy cannot take all the values continuously, p as well as q accordingly cannot assume every arbitrary value either.

Denoting the amount of the energy quantum by  $\varepsilon$ , the energy E can take only those values which are integral multiples of  $\varepsilon$  and hence p and q can take only those values which satisfy

$$E = \alpha p^2 + \beta q^2 = n\varepsilon$$
,  $n = 0, 1, 2, ...$  (8.4)

Geometrically speaking, the points on a set of ellipses defined by Eq. (8.4) on the q-p plane, as shown in Fig. 10, correspond to the

allowed values for p and q. Thus the integration of Eq. (8.2), or of Eq. (8.3), is to be understood not as a surface integral over the whole q-p plane but as a sum of the line integrals on these ellipses.

In accordance with this understanding, the calculation proceeds as follows. We first change the variables



lows. We first change the variables Fig. 10. Allowed values for p and q. from q and p to x and y by  $\sqrt{\beta} q = x$ 

and  $\sqrt{a} \phi = y$ , respectively. The energy E is then given by  $E = x^2 + y^2$ and the integrations to be performed become

$$I_1 = \frac{1}{\sqrt{\alpha\beta}} \int \int (x^2 + y^2) e^{-(x^2 + y^2)/kT} dx dy$$

and

$$I_2 = \frac{1}{\sqrt{\alpha\beta}} \int \int e^{-(x^2 + y^2)/kT} \, dx dy \, .$$

In order to carry out these integrations, we use as the integration variables the energy E itself and the angle  $\theta$ , defined by  $\theta = \tan^{-1}(y/x)$ , instead of x and y. The above integrals then become

$$I_{1} = \frac{1}{2\sqrt{\alpha\beta}} \int \int E e^{-E/kT} dE d\theta$$
$$I_{2} = \frac{1}{2\sqrt{\alpha\beta}} \int \int e^{-E/kT} dE d\theta.$$

Since  $\theta$  can take any value from zero to  $2\pi$ , whereas E has to satisfy Eq. (8.4), we first integrate with respect to  $\theta$  and obtain

$$I_1 = \frac{\pi}{\sqrt{\alpha\beta}} \int E \, \mathrm{e}^{-E/kT} \, \mathrm{d}E$$

and

 $I_2 = \frac{\pi}{\sqrt{\alpha\beta}} \int e^{-E/kT} dE \, .$ 

and hence

$$\langle E \rangle = \frac{\int E e^{-E/kT} dE}{\int e^{-E/kT} dE}$$
 (8.5)

where the integration with respect to E is to be understood as the sum of the integrand for those values of E satisfying Eq. (8.4). Thus

$$\langle E \rangle = \frac{\sum_{n=0}^{\infty} n\varepsilon e^{-n\varepsilon/kT}}{\sum_{n=0}^{\infty} e^{-n\varepsilon/kT}}.$$
(8.5')

The summation of these series is not difficult. Since

$$\sum_{n=0}^{\infty} n\varepsilon e^{-n\varepsilon/kT} = -\frac{\partial}{\partial \left(\frac{1}{kT}\right)^{n=0}} e^{-n\varepsilon/kT}$$

we have only to calculate  $\sum_{n=0}^{\infty} e^{-n\epsilon/kT}$ , which in turn is nothing but a geometrical series the sum of which is readily given as

$$\frac{\mathrm{e}^{\mathfrak{e}/kT}}{\mathrm{e}^{\mathfrak{e}/kT}-1}\,.$$

From this result we obtain

$$\sum_{n=0}^{\infty} n\varepsilon \, \mathrm{e}^{-n\varepsilon/kT} = \frac{\varepsilon \, \mathrm{e}^{\varepsilon/kT}}{(\mathrm{e}^{\varepsilon/kT} - 1)^2} \, .$$

The substitution of these results into Eq. (8.5') gives

$$\langle E \rangle = \frac{\varepsilon}{\mathrm{e}^{\varepsilon/kT} - 1}$$
 (8.6)

or in terms of the function P(x) defined by Eq. (7.9)

$$\langle E \rangle = kTP\left(\frac{\varepsilon}{kT}\right).$$
 (8.7)

This is the distribution law when the atomic nature of the energy is taken into consideration. If the energy did not have such an atomic nature, Eq. (8.5) would be integrated in the usual manner instead of using Eq. (8.5') and we would get immediately Eq. (2.14).

The comparison of Eq. (8.7) with Planck's interpolation formula, Eq. (7.8), leads us to the following very important conclusion: If we take  $h\nu$  as the energy quantum, i.e.

$$\varepsilon = h \nu,$$
 (8.8)

r being the frequency of the oscillating system under consideration, our distribution law, Eq. (8.7), gives Eq. (7.8) and hence Planck's formula, Eq. (7.5). From Eq. (8.8) we can conclude that the energy value of the oscillator can take only integral multiples of this unit:

$$E = nhv$$
,  $n = 0, 1, 2, ...$  (8.9)

In this way Planck arrived at the epoch-making idea that not only material substances but also energy reflects the atomic nature of the physical world. This idea was the great discovery that the dawn of this century, 1900, witnessed.

According to Eq. (8.8) the larger the frequency  $\nu$ , the larger the energy unit. Therefore, the atomic nature of energy becomes more and more distinct with increasing frequency  $\nu$  of the oscillating system. Furthermore, when the amount of energy under consideration is sufficiently large we can neglect the atomic nature of energy just as

we can treat the material object of large scale as if it were continuous matter. This implies that, in considering the problem of thermal equilibrium, if kT is sufficiently larger than  $h\nu$  we can regard the energy as continuous and the energy equipartition law prevails. On the other hand, if  $h\nu$  is comparable with or larger than kT, the unit energy plays an important role and the energy equipartition law becomes invalid; then a share of even a single unit of energy to a degree of freedom would make its energy E greater than kT and consequently its probability  $e^{-E/kT}$  become small and such degrees of freedom hardly receive any energy at all. This explains why some degrees of freedom "die" at low temperatures. The reader may find an analogy to the situation in Fig. 11, which is self-explanatory.



Fig. 11. Distribution laws in classical and in quantum theory.

The value of Planck's constant h, as given by Eq. (7.4), has been determined by experiments of great precision. Its approximate value, however, can be estimated from our daily experiences together with the above considerations.

We observe that a heated object starts radiating a dim red light at around 500 °C. The color of the radiation becomes brilliant red at about 1000 °C and with still increasing temperature the object starts emitting a bright white light at about 2000 °C.

Since the wave length of red light is  $8 \times 10^{-5}$  cm its frequency is

$$v_{\rm red} = \frac{c}{\lambda_{\rm red}} = \frac{3 \times 10^{10}}{8 \times 10^{-5}} \simeq 0.4 \times 10^{15} \, {\rm sec}^{-1} \, .$$

The observed fact that at about 500 °C it is a dim red light and be-

comes brilliant red at about 1000 °C suggests that kT at 500 °C is still too small compared with  $hv_{red}$  to make the red oscillation very active, and that at 1000 °C kT becomes comparable to  $hv_{red}$  thus sufficiently activating the red oscillation. Hence we can expect the following relation to hold,

whence

$$h\nu_{\rm red} \cong k \ (1000 + 273)$$

$$h \simeq 1273 \times \frac{k}{\nu_{\rm red}}$$
.

Using the above value of  $v_{red}$  and Boltzmann's constant, we find

$$h = 4 \times 10^{-28} \text{ erg} \cdot \text{sec}$$
.

This estimated value of h is about one order of magnitude smaller than the correct value of  $6.62 \times 10^{-27}$  erg·sec, but in view of the very rough nature of our estimate it should be regarded as a fair agreement.

# § 9. The Quantum Theory of Specific Heat

We have now obtained Planck's distribution formula to replace the energy equipartition law. It is natural to think of constructing a new theory of specific heat based on this new formula. Einstein and Debye arrived at this idea almost at the same time and developed the theory of the specific heat of solids. In this section we follow Debye's method.

Since Planck's distribution law contains the frequency  $\nu$ , the energy of the thermal oscillations cannot be calculated only from the knowledge of the total number, f, of the oscillational degrees of freedom. We have to know the number of proper oscillations of various frequencies for this purpose. Mathematically speaking, we have to know the number of proper oscillations,  $Z(\nu) d\nu$ , of frequencies between  $\nu$ and  $\nu + d\nu$ . For an elastic continuum or for electromagnetic oscillations it is simple to calculate this  $Z(\nu) d\nu$ , and it is given by Eq. (4.4) or Eq. (4.5). It, however, is an extremely difficult problem to calculate this quantity for an actual crystalline substance which has a complex lattice structure. A simple model of a crystal is treated in Appendix II.

Still, Debye arrived at a sensible result by using the following approximation. He replaced the actual crystalline substance by an

elastic continuum in which both modes of oscillation, longitudinal and transverse, are possible as they are in actual substances. These two oscillations have different velocities of propagation,  $c_1$  and  $c_4$ , in actual substances.

Following the argument as given in § 4, we obtain

$$Z(\nu) \,\mathrm{d}\nu = 4\pi L^3 \left(\frac{1}{c_1^3} + \frac{2}{c_t^3}\right) \nu^2 \,\mathrm{d}\nu \tag{9.1}$$

where the factor 2 in the  $c_t^3$  term comes from the fact that a transverse wave has two possible directions of polarization.

The actual crystal is different from a continuum in that it has a finite number of degrees of freedom f = 3N, where N is the number of atoms in the crystal. Consequently, the frequency of oscillation of the crystal has an upper limit  $v_{\max}$  and Z(v) for v larger than  $v_{\max}$  is to be defined as zero. In order to determine this  $v_{\max}$ , we can use the requirement that the total number of proper oscillations must be equal to 3N and thus

$$\int_{0}^{\nu_{\max}} Z(\nu) \, \mathrm{d}\nu = 3N \tag{9.2}$$

As the result, we get

$$Z(\nu) = \begin{cases} 3GL^3\nu^2, & \text{for } \nu \le \nu_{\max} \\ 0, & \text{for } \nu > \nu_{\max} \end{cases}$$
(9.3)

where

$$G = \frac{4\pi}{3} \left( \frac{1}{c_1^3} + \frac{2}{c_t^3} \right).$$
(9.4)

Since the amount of energy distributed to a proper oscillation is given by Planck's result, Eq. (7.8), we can readily calculate the total energy of the crystal, which is

$$U = \int_{0}^{\nu_{\max}} kTP\left(\frac{h\nu}{kT}\right) Z(\nu) \, \mathrm{d}\nu \,. \tag{9.5}$$

The introduction of Eq. (9.3) into Eq. (9.2) gives for  $v_{max}$ 

$$\boldsymbol{\nu}_{\max} = \left(\frac{3N}{GL^3}\right)^{\frac{1}{3}}.$$
(9.6)

#### CH. I, § 9] THE QUANTUM THEORY OF SPECIFIC HEAT

Using this result, U becomes

$$U = 3NkTD\left(\frac{\Theta}{\overline{T}}\right) \tag{9.7}$$

where D(x) and  $\Theta$  are defined by

$$D(x) = \frac{3}{x^3} \int_0^x \xi^2 P(\xi) \, \mathrm{d}\xi = \frac{3}{x^3} \int_0^x \frac{\xi^3 \mathrm{d}\xi}{\mathrm{e}^\xi - 1} \tag{9.8}$$

and

$$\Theta = \frac{h v_{\max}}{k}.$$

The quantity  $\Theta$  has the dimension of temperature and is called the Debye temperature. It has different values for different substances, depending on  $c_1$ ,  $c_1$  and  $N/L^3$ . On the other hand, the function D(x), which is called the Debye function, has a universal nature and does not depend on the properties of the substance.

For a unit mole of crystal, Nk can be replaced by the gas constant R. Then differentiation with respect to temperature T of Eq. (9.7) yields for the mole specific heat of the crystal

$$C = 3R \left\{ 4D \left( \frac{\Theta}{T} \right) - \frac{3\Theta/T}{e^{\theta/T} - 1} \right\}$$
(9.9)

in which only  $\Theta$  depends on the properties of the crystal.

According to Eq. (9.8), the function D(x) has values close to unity for sufficiently small x and decreases as x increases. It approaches zero with  $x \to \infty$ .

With this behavior of the function D(x) in mind, we find from inspection of Eq. (9.9) that when the temperature T is sufficiently larger than the Debye temperature  $\Theta$  the specific heat C is close to Dulong-Petit's value of 3R whereas when T is considerably smaller than  $\Theta$  it approaches zero. This fact leads us to overcome the difficulty encountered previously in Fig. 3 of § 2. We reproduce here again the experimentally obtained relationship between the specific heat and temperature as well as the theoretical one given by Eq. (9.9). As can be seen in the figure, with the proper choices of  $\Theta$  for various substances the agreement between the experimental values and the theoretical

4I

ones is remarkably good, though it is difficult to calculate the value of  $\Theta$  theoretically for a given substance. The values of  $\Theta$  for various

TABLE 2Debye temperatures for<br/>several substancesSubstanceØ

398

215

474

88

Al

Ph

Ag

CaF,

substances are shown in Table 2. The good
agreement observed in Fig. 12 clearly in-
dicates that Planck's idea of the energy
quantum is well applicable not only to
the problem of electromagnetic oscilla-
tions but also to that of material oscil-
lations.

A similar consideration is expected to solve the difficulty encountered with the specific heat of a gas. However, in this case, the problematical point is in the ro-

tational energy of diatomic molecules as was explained in connection with Fig. 2 of § 2.



Fig. 12. The relations between the molar specific heat of various solids and the temperature. The abscissa shows temperature (in units of Debye temperature), the ordinate molar specific heat (cal/R).

We now know, thanks to Planck, about the quantum of oscillational energy, but we have no idea what value the quantum of rotational energy is to take. Hence, this problem has to be left for later discussion. At any rate, also in this case, the rotational degrees of freedom die out at low temperature as a result of the existence of the quantum of rotational energy and consequently the value of the specific heat decreases. The reader is referred to § 24 for discussion of these points.

# CHAPTER 2

# THE CORPUSCULAR NATURE OF LIGHT

# § 10. The Hypothesis of the Light Quantum

From what has been explained in Chapter 1 it should have become clear to the reader how the introduction of the concept of the energy quantum removed the difficulties associated with black body radiation and specific heat. Logically it appears very simple to get to Planck's formula by this approach. Nevertheless it took nearly 10 years for Planck's theory to be generally accepted. This was not only because it differs fundamentally from Newtonian mechanics and from Maxwell's electrodynamics but also because there were scarcely any experimental data which would demonstrate in a direct manner, without making any statistical manipulation, the existence of the energy quantum.

It may appear that the concept of the energy quantum is not in severe contradiction with the classical theory. For example, consider an oscillating system, a pendulum to be specific; then according to mechanics the coordinate of the pendulum varies sinusoidally with time. The existence of an elementary quantum of energy simply requires that the amplitude of oscillation can not take arbitrary values. In so far as the amplitude has a certain allowed value, can we not consider that the pendulum is moving sinusoidally according to the laws of mechanics? Similar questions can be raised also in the case of the oscillation of a string or of the black body radiation. Namely, can we not consider that the proper oscillation of the hollow cavity obeys Maxwell's equation so long as its amplitude has a certain allowed value.

This idea, however, does not work. It would work for oscillating systems which are at all times isolated without any contact with the environment. When the hollow cavity is exchanging energy with the environment through a small opening in its wall, the energy cannot flow in or out continuously through the opening since the energy in each proper oscillation of the hollow cavity consists of elementary quanta of energy. The energy of the hollow cavity has to change by an amount  $h\nu$  discontinuously at some instant and this is to be compensated by the same discontinuous change in the energy of the environment. Could this happen inside the framework of Newtonian mechanics and Maxwell's electrodynamics?

It is quite evident from the following consideration that the concept of the energy quantum is utterly incompatible with classical concepts.



Fig. 13. A hollow cavity with a small internal cavity.

Consider a small chamber of volume v in a sufficiently large hollow cavity of volume V as is shown in Fig. 13. These two chambers have a small opening as a connecting link and they are in thermal equilibrium at temperature T. Let one of the proper oscillations of the chamber have frequency v. The probabilities for its energy to take values  $0, hv, 2hv, \ldots$  are not zero since  $e^{-nhv/kT}$  never becomes zero for any values of *n*.Hence the energy of the small chamber v is

[CII. 2, § 10

zero at some time,  $h\nu$  or  $2h\nu$  at some other time, thus changing its value incessantly. An increase of the energy by an amount  $h\nu$  has to occur at some instant or other. Since there is no permissible intermediate value, this increase has to occur within an instant of time; i.e., a finite amount of energy has to flow in from the large chamber V through the small opening within an instant of time. This, however, is utterly impossible according to Maxwell's theory. For the energy in this theory is supposed to be distributed all over the space and to be transported with a finite velocity. Therefore, unless the small opening has a kind of magical power, that of collecting energy from a distance within an instant of time, the above situation can never occur.

In order for an energy quantum to jump in through the opening within an instant of time, the energy of the electromagnetic field has to be considered as spreading over the space discontinuously in lumps of finite amounts rather than as distributed continuously. In other words, the propagation of light of frequency v has to be considered as the flight of a particle of energy hv. The first who noted the corpuscular nature of light in this sense was Einstein and he called this particle "the quantum of light". Whether the energy is distributed continuously in the space or lumped into amounts of hv can be decided by observing the temporal variation of the energy contained in the small chamber v. Einstein noticed that even without a detailed ob-

# CH. 2, § 10, 11] THE ENERGY IN A HOLLOW CAVITY

servation of this temporal variation the decision can be made to a certain extent by investigating the fluctuation of the energy in the chamber v. He furthermore discovered that this fluctuation could be calculated from the known formula of the radiation spectrum, even without detailed knowledge of the laws governing the radiation phenomena, provided that certain basic assumptions could be made. Thus he showed that if one once accepted Planck's formula one had also to accept the corpuscular nature of light. This was in 1905.

## § 11. The Fluctuation of the Energy in a Hollow Cavity

We follow Einstein in considering the corpuscular nature of light mentioned in the preceding section. To start with we take Boltzmann's principle, Eqs. (2.1) and (2.2), as the basic assumption. It is quite difficult to imagine the break-down of this principle since it is closely related to the second law of thermodynamics.

We consider the radiation of temperature T inside a hollow cavity of volume V which has no exchange of energy with the surroundings; the temperature T of the hollow cavity, by the way, can be attained by providing a small window for the hollow cavity and immersing it in the thermal reservoir of temperature T for a sufficient length of time before one closes the window again. Then the spectral distribution of the radiation inside should be given by Eq. (7.5). We now consider a small chamber of volume v in this hollow cavity as shown in Fig. 13. It is evident that the spectral distribution of radiation inside this small chamber v is also given by Eq. (7.5) and hence the energy contained in the radiation of frequencies between v and v + dv in this volume is

$$E = v U(v) \, \mathrm{d}v = \frac{8\pi h v}{c^3} \frac{1}{\mathrm{e}^{h v/kT} - 1} v^3 \, \mathrm{d}v \,. \tag{11.1}$$

This is to be accepted as an experimentally established law whether one accepts Planck's idea or not. This energy, E, is of course to be understood as the average energy,  $\langle E \rangle$ , which is experimentally to be observed.

We next assume that the average energy of the radiation is the sum of average energies corresponding to each independent degree of freedom of various frequencies. In view of the fact that light waves of various frequencies can exist in a space independently of each other, this assumption seems well justified. Even if one does not believe in the wave theory of light and does not accept such things as frequencies, there must be some parameter which distinguishes the colour of light. Let this parameter be denoted by v and called conventionally "frequency". The radiation energy given by Eq. (11.1) is then to be considered as the sum of the average energies corresponding to a number of degrees of freedom with frequencies between v and v + dv. Denoting the energy of the s-th degree of freedom by  $E_s$ , we have

$$vU(v) \,\mathrm{d}v = \sum_{s=1}^{f} \langle E_s \rangle = \langle E \rangle \tag{11.2}$$

where  $\sum_{s=1}^{f}$  means the summation over the number of degrees of freedom having frequencies between  $\nu$  and  $\nu + d\nu$ .

When the volume v of the small chamber is sufficiently small as compared with the total volume V, the large chamber V can be considered as a thermal reservoir relative to the small chamber v. We then can apply Boltzmann's principle to the small chamber and

$$\langle E_s \rangle = \frac{\iint E_s \,\mathrm{e}^{-E_s/kT} \,\mathrm{d}p_s \,\mathrm{d}q_s}{\iint \,\mathrm{e}^{-E_s/kT} \,\mathrm{d}p_s \,\mathrm{d}q_s} \,, \tag{11.3}$$

where, if necessary, the integrals may be replaced by sums. Though the average energy of the small chamber is given by Eq. (11.1), the actual instantaneous energy fluctuates around this average value. The measure of the fluctuation is usually given by its mean square; i.e., by the average of the square of the difference between the actual energy  $\sum_{s=1}^{f} E_s$  and its average  $\sum_{s=1}^{f} \langle E_s \rangle$ . Then, denoting this mean square by  $\langle \Delta E^2 \rangle$ , we have from the definition

$$\langle \Delta E^2 \rangle = \left\langle \left( \sum_{s=1}^f E_s - \sum_{s=1}^f \langle E_s \rangle \right)^2 \right\rangle$$

$$= \sum_{s=1}^f \sum_{s'=1}^f \langle (E_s - \langle E_s \rangle) (E_{s'} - \langle E_{s'} \rangle) \rangle.$$

$$(11.4)$$

Noting that in Eq. (11.4) the energies of mutually independent degrees of freedom are fluctuating without inter-correlations, and

# CH. 2, § 11] THE ENERGY IN A HOLLOW CAVITY

hence that  $\langle (E_s - \langle E_s \rangle) (E_s - \langle E_{s'} \rangle) \rangle$  for  $s \neq s'$  in Eq. (11.4) is zero, we obtain

$$\langle \Delta E^2 \rangle = \sum_{s=1}^{f} \langle (E_s - \langle E_s \rangle)^2 \rangle = \sum_{s=1}^{f} \{ \langle E_s^2 \rangle - \langle E_s \rangle^2 \}$$
(11.5)

where  $\langle E_s^2 \rangle$  is to be given, according to Boltzmann's principle, by

$$\langle E_s^2 \rangle = \frac{\iint E_s^2 e^{-E_s/kT} dp_s dq_s}{\iint e^{-E_s/kT} dp_s dq_s}.$$
 (11.6)

We shall now prove that, once  $\langle E \rangle$  is given as a function of T, the mean square  $\langle \Delta E^2 \rangle$  of Eq. (11.5) can be obtained from this function as follows:

$$\langle \Delta E^2 \rangle = \frac{\partial \langle E \rangle}{\partial (-1/kT)} = \frac{\partial U(\nu)}{\partial (-1/kT)} \nu \, \mathrm{d}\nu \,.$$
 (11.7)

We have, first of all,

$$\frac{\partial \langle E \rangle}{\partial (-1/kT)} = \sum_{s=1}^{f} \frac{\partial \langle E_s \rangle}{\partial (-1/kT)}$$

which becomes, by the introduction of Eq. (11.3),

$$= \sum_{s=1}^{f} \left\{ \frac{\iint E_s^2 e^{-E_s/kT} dp_s dq_s}{\iint e^{-E_s/kT} dp_s dq_s} - \frac{(\iint E_s e^{-E_s/kT} dp_s dq_s)^2}{(\iint e^{-E_s/kT} dp_s dq_s)^2} \right\}.$$

Then, noting Eq. (11.3) and Eq. (11.6), this is nothing but

$$= \sum_{s=1}^{f} \left\{ \langle E_s^2 \rangle - \langle E_s \rangle^2 \right\},\,$$

thus completing the proof of Eq. (11.7).

This important theorem, Eq. (11.7), holds quite generally irrespective of whether the radiation obeys Maxwell's equations or not, or of whether the energy has a quantum nature or not. The only underlying assumptions are Boltzmann's principle, the mutual independence of different degrees of freedom of radiation, and the requirement that the spectral distribution of energy is a function of temperature only. The last of the above three assumptions is necessary since, in the case of gaseous substances for example (where the energy is a function not only of temperature but also of density), the gas density, and hence the number of degrees of freedom f, in the small chamber is also fluctuating around the average, and as a result such a simple treatment is not allowed.

Substituting the experimentally verified Eq. (11.1) for  $U(v)v \, dv$  into the general formula, Eq. (11.7), we obtain the mean square fluctuation  $\langle \Delta E^2 \rangle$  without making any specific assumptions as to the nature of radiation.

To begin with, however, we use the Rayleigh-Jeans formula instead of Eq. (11.1). Namely we put

$$vU(v) \,\mathrm{d}v = \frac{8\pi v}{c^3} \,kT \,v^2 \,\mathrm{d}v$$

in Eq. (11.7). We then get by differentiating with respect to  $-\frac{1}{kT}$ 

$$\langle \Delta E^2 \rangle = \frac{8\pi v}{c^3} k^2 T^2 v^2 dv$$

which on elimination of T becomes

$$\langle \Delta E^2 \rangle = \frac{\{ U(\nu)v \, \mathrm{d}\nu \}^2}{\frac{8\pi v}{c^3} \nu^2 \, \mathrm{d}\nu} = \frac{c^3}{8\pi \nu^2} \frac{\langle E \rangle^2}{v \, \mathrm{d}\nu}.$$
 (11.8)

On the other hand the use of Wien's formula,

$$U(\mathbf{v})\mathbf{v}\,\mathrm{d}\mathbf{v}=\frac{8\pi vh}{c^3}\,\mathrm{e}^{-h\mathbf{v}/kT}\,\mathbf{v}^2\,\mathrm{d}\mathbf{v}\,,$$

gives

$$\langle \Delta E^2 \rangle = h \nu \frac{8\pi v h}{c^3} e^{-h \nu/kT} \nu^2 d\nu = h \nu \langle E \rangle.$$
 (11.9)

Finally the use of Planck's formula, Eq. (11.1), gives

$$\langle \Delta E^2 \rangle = h\nu \langle E \rangle + \frac{c^3}{8\pi\nu^2} \frac{\langle E \rangle^2}{v \, \mathrm{d}\nu}.$$
 (11.10)

We investigate in the following sub-sections the significance of these mean square fluctuations in the three cases.

## (i) THE FLUCTUATION WITH THE RAYLEIGH-JEANS FORMULA

Since the Rayleigh-Jeans formula is derived from Maxwell's theory, we can expect that the fluctuation formula, Eq. (11.8), reflects the wave nature of light. In fact, that there should be fluctuations due to the wave nature of light can be expected from the following consideration. In a hollow cavity filled with light, although the light is distributed uniformly on the average in the space, it is possible that at a certain point and a certain time the light wave of a certain frequency interferes, constructively or destructively, with that of a slightly different frequency, since the light waves are intermingled randomly in the space.

Due to a kind of beat resulting from these interferences, the energy of the small chamber v fluctuates becoming smaller at one time and larger at some other time than its average value. A detailed calculation of the fluctuation due to wave nature is given in Appendix VII.

Concerning the energy fluctuation due to such beat phenomenon, the following statement can be made without refering to specific calculations. We consider the radiation having a given average energy  $\langle E \rangle$  and denote its mean square fluctuation by  $\langle \Delta E^2 \rangle$ . For another radiation whose amplitude is twice as large everywhere in the frequency range dv under consideration, its average energy  $\langle E \rangle$  is four times as large as before, since the energy is proportional to the square of the amplitude. Meanwhile, since all the amplitudes are doubled the beat amplitude becomes also doubled, thus giving  $\langle \Delta E^2 \rangle$  16 times as large as before. This implies that  $\langle \Delta E^2 \rangle$  has to be proportional to the square of  $\langle E \rangle$ . On the other hand, when we double the number of degrees of freedom,  $8\pi v v^2 dv/c^3$ , by doubling the frequency range dv under consideration, by multiplying the frequency  $\nu$  itself by  $2^{\frac{1}{2}}$ , or by doubling the volume v of the small chamber, the average energy  $\langle E \rangle$  is doubled since it is the sum of energies in each degree of freedom and is proportional to the number of degrees of freedom.  $\langle E \rangle^2$ becomes then four times as large as before, while  $\langle \Delta E^2 \rangle$ , being the sum of fluctuations in each degree of freedom as given by Eq. (11.5), is proportional to the number of degrees of freedom and becomes only twice as large as before. Therefore, the ratio  $\langle \Delta E^2 \rangle / \langle E \rangle^2$  has to be inversely proportional to the number of degrees of freedom  $8\pi v \nu^2 d\nu/c^3$ , i.e.

$$\langle \Delta E^2 \rangle = \operatorname{const} \frac{c^3}{8\pi\nu^2} \frac{\langle E \rangle^2}{v \, \mathrm{d}\nu},$$
 (11.11)

where the proportionality constant is dimensionless and depends on neither  $\langle E \rangle$ , v, v nor dv. We can in this way understand the implication of Eq. (11.8). It can be ascertained by a more detailed calculation that this proportionality constant actually takes the value unity in conformity with Eq. (11.8).

## (ii) THE FLUCTUATION WITH WIEN'S FORMULA

From the consideration of the preceding sub-section, we see that the wave nature of light leads necessarily to  $\langle \Delta E^2 \rangle$  proportional to  $\langle E \rangle^2$  and inversely proportional to  $8\pi v v^2 dv/c^3$  and hence that Eq. (11.9) is not obtainable if we start with the assumption that light has a wave nature. Rewriting Eq. (11.9), we get

$$\frac{\langle \Delta E^2 \rangle^{\frac{1}{2}}}{\langle E \rangle} = \frac{(h\nu)^{\frac{1}{2}}}{\langle E \rangle^{\frac{1}{2}}} \,. \tag{11.12}$$

This relationship implies that the ratio of the fluctuation to the average energy is inversely proportional to the square root of the average energy. Hence, in contrast to Eq. (11.11) which is based on the wave nature of light, the larger the average energy is, the less important the fluctuation becomes in this case. This kind of fluctuation we often encounter in statistics, in dealing with the cases where many independent fluctuations are superposed. To be more specific, we treat in the following the case of the fluctuation in the theory of a gas.

We consider a large volume V filled with a great many gas molecules moving independently inside, and let the number of these molecules be denoted by N. The quantity N/V is then the number of molecules per unit volume, or the density  $\rho$ , thus

$$\frac{N}{V} = \rho . \tag{11.13}$$

We now consider a very small volume v within this large volume assuming that  $v \ll V$ . Then, the probability for a particular molecule to be found in the small volume v is given by

$$w = \frac{v}{N} = \frac{\rho v}{N} \tag{11.14}$$

## CH. 2, § 11] THE ENERGY IN A HOLLOW CAVITY

and the probability for it not to be found there is (1 - w). We now can obtain the probabilities of finding in the small volume v no molecule, one, two, ..., n molecules respectively. They are as follows:

the probability of finding no molecule in $v$	$=(1-w)^N$
the probability of finding one molecule in $v$	$= \begin{pmatrix} N \\ 1 \end{pmatrix}  w(1 - w)^{N-1}$
the probability of finding two molecules in $v$	$= \binom{N}{2} w^2 (1 - w)^{N-2}$
three molecules	$= \binom{N}{3} w^3 (1-w)^{N-3}$
<i>n</i> molecules	$= \binom{N}{n} w^n (1-w)^{N-n}$

Considering the fact that V is much larger than v and hence that the probability w is very much smaller than unity, as well as the fact that N is very large, we can approximately rewrite the above probabilities for n very much smaller than N to get;

the probability of	f finding no molecu	the in $v$	$= e^{-\rho v}$	
the probability o	of finding one mo	lecule in v	$= e^{-\rho v} \frac{\rho v}{1!}$	
two molecules			$= e^{-\rho v} \frac{(\rho v)^2}{2!}$	(11.15)
three molecules			$= e^{-\rho v} \frac{(\rho v)^2}{3!}$	
n molecules			$= \mathrm{e}^{-\rho v} \frac{(\rho v)^n}{n!}$	

where use is made of  $e^{-\alpha} = \lim_{x \to 0} (1-x)^{\alpha/x}$ . Since these probabilities are derived under the assumption that  $V \gg v$  and  $N \gg 1$ , they can be applied also for the case of infinitely many molecules moving in a infinitely large volume where, however, the density N/V is kept at  $\rho$ .

With the help of these probabilities, we can easily calculate the average number of molecules in the small volume v and we find

$$\langle n \rangle = e^{-\rho v} \sum_{n=0}^{\infty} n \frac{(\rho v)^n}{n!} = \rho v .$$
 (11.16)

Furthermore, using the above result, the mean square of the number of molecules is

$$\langle n^2 \rangle = e^{-\langle n \rangle} \sum_{n=0}^{\infty} n^2 \frac{\langle n \rangle^n}{n!}$$
  
=  $\langle n \rangle + \langle n \rangle^2$ .

Hence, we obtain

$$\langle \Delta n^2 \rangle = \langle n^2 \rangle - \langle n \rangle^2 = \langle n \rangle$$

or, in another form,

$$\frac{\langle \Delta n^2 \rangle^{\frac{1}{2}}}{\langle n \rangle} = \frac{1}{\langle n \rangle^{\frac{1}{2}}}.$$
(11.17)

In this way we see that when a great number of independent molecules are moving randomly in a very large volume, or when we have an infinite number of molecules in an infinitely large volume, the ratio of the fluctuation to the average number of molecules in a small volume is inversely proportional to the square root of the average number. The importance of fluctuations thus becomes less with increasing average number. This is the very reason why, in the ordinary case of gases consisting of a large number of molecules, we can safely say, "the number of molecules in the small chamber is (v/V)N" and restrict ourselves to the consideration of average behavior without taking the fluctuations into account.

In this model gas, if each molecule had an energy  $h\nu$ , Eq. (11.17) could be transformed, by using the relations  $\langle E \rangle = h\nu \langle n \rangle$ ,

$$\langle E^2 \rangle = (hv)^2 \langle n^2 \rangle$$
, and  $\langle \Delta E^2 \rangle = (hv)^2 \langle \Delta n^2 \rangle$ , into

$$\frac{\langle \Delta E^2 \rangle^{\frac{1}{2}}}{\langle E \rangle} = \frac{1}{\langle n \rangle^{\frac{1}{2}}} = \frac{(h\nu)^{\frac{1}{2}}}{\langle E \rangle^{\frac{1}{2}}}.$$
 (11.18)

Comparison of Eq. (11.18) with Eq. (11.12) makes it clear that the fluctuation obtained from Wien's formula is of the same nature that

would be obtained if the light were considered as independently moving particles of energy  $h\nu$ .

The fundamental difference between the fluctuation due to the wave nature and that due to the corpuscular nature explained in (i) and (ii), respectively, is as follows.

In the first case (wave nature), the energy in the small volume, E, as a function of time will resemble, for example, that shown in Fig. 14 (a) and the square of the corresponding energy,  $E^2$ , that given in Fig. 14 (b). The average energy  $\langle E \rangle$  and the mean square energy  $\langle E^2 \rangle$  can be obtained from these two figures. What change would



Fig. 14. The time variation of the energy in a small cavity (according to the wave theory). In Fig. (2a) and Fig. (2b) the light intensity is twice that of Fig. (a) and Fig. (b), respectively. These graphs are drawn for the following numerical values.

these  $\langle E \rangle$  and  $\langle E^2 \rangle$  suffer if the light were to be intensified by a factor of two? According to the wave theory of light, this intensification implies that the amplitude is made larger by a factor of  $\sqrt{2}$ . Therefore, for the light with doubled intensity, E and  $E^2$  are represented by Figs. (2a) and (2b), respectively. Fig. (2a) is obtained by expanding the Fig. (a) in the direction of the ordinate by a factor of two while Fig. (2b) results from a similar procedure but with a factor of four.

One can readily see from the figures that  $\langle E \rangle$  and  $\langle E^2 \rangle$  become larger by a factor of two and four, respectively, and furthermore that  $\langle \Delta E^2 \rangle$ , proportional to  $\langle E \rangle^2$ , becomes four times larger. The point here is that the intensification of the light by a factor of two results in Fig. (a) being expanded, in the direction of ordinate, by the same factor.

In the second case (corpuscular nature), the energy E in the small chamber can be represented as in Fig. 15 (a). In this case, since each particle has the same amount,  $h\nu$ , of energy, the "*E*-curve" takes the form of a partially broken comb. In Fig. 15 (a), however, the quantity



Fig. 15. The time variation of the energy in a small cavity (according to the particle theory). In Fig. (2a) and Fig. (2b) the light intensity is twice that of Fig. (a) and Fig. (b), respectively. These graphs are drawn for the following numerical values.

$$\langle E \rangle = 0.08h\nu, \langle E \rangle^2 = 0.006(h\nu)^2, \qquad \langle E \rangle = 0.16h\nu, \langle E \rangle^2 = 0.026(h\nu)^2, \\ \langle E^2 \rangle = 0.08(h\nu)^2, \qquad \langle E^2 \rangle = 0.16(h\nu)^2, \\ \langle \Delta E^2 \rangle = \langle E \rangle - \langle E \rangle^2 = 0.07(h\nu)^2, \qquad \langle \Delta E^2 \rangle = \langle E^2 \rangle - \langle E \rangle = 0.13(h\nu)^2, \\ \langle \Delta E^2 \rangle = \langle E \rangle^2 = 11, \qquad \qquad \langle \Delta E^2 \rangle = \langle E \rangle^2 = 5, \\ \langle \Delta E^2 \rangle = \langle E \rangle^2 \sim \frac{1}{\langle E \rangle^2}$$

 $E/h\nu$ , i.e., the number of particles in the small chamber is used, instead of E itself, for the ordinate.

For the sake of simplicity, we have considered in the figure the case of small density so that the probability of finding two particles simultaneously in the chamber was extremely small.

The figure for the corresponding  $(E/h\nu)^2$  is shown in Fig. 15 (b) which has the same shape as that of (a) since  $E/h\nu$  is either zero or one. In case there is an instant when two particles are simultaneously in the small chamber, the shape of these two figures will become different, but, in our simple case of a small density,  $\langle E \rangle / h\nu$  is thus equal to  $\langle E^2 \rangle / (h\nu)^2$ . We now intensify the light by a factor of two. Since in the corpuscular theory of light the energy carried by each particle of light is always hv irrespective of whether the light is intense or weak, the intensification of light implies a corresponding increase of the density of these particles. Accordingly, the time variation of  $E/h\nu$ now takes the form as shown in Fig. 15 (2a). The difference between Fig. 15 (a) and Fig. 15 (2a) is not simply an ordinate scale difference as in the case of the wave theory, but instead a difference in the number of teeth of the broken comb. The same comparison exists between Fig. 15 (b) and Fig. 15 (2b). Consequently,  $\langle E^2 \rangle$  of Fig. 15 (2b) is only twice as large as that of Fig. 15 (b), while it was four times as large in the wave theory, and hence  $\langle E^2 \rangle / (h\nu)^2$  is proportional to  $\langle E \rangle$ , not to  $\langle E \rangle^2$ ; i.e., in our simplified case  $\langle E^2 \rangle / (h\nu)^2 = \langle E \rangle / h\nu$ . This is the reason why  $\langle E^2 \rangle / \langle E \rangle^2$  is inversely proportional to  $\langle E \rangle$ ; in our simplified case,  $\langle E \rangle / h v \ll 1$ ,

$$\langle \Delta E^2 \rangle = (h\nu)^2 \left\{ \frac{\langle E^2 \rangle}{(h\nu)^2} - \left( \frac{\langle E \rangle}{h\nu} \right)^2 \right\} \simeq (h\nu)^2 \frac{\langle E^2 \rangle}{(h\nu)^2} = h\nu \langle E \rangle$$

and hence  $\langle \Delta E^2 \rangle / \langle E \rangle^2 = h\nu / \langle E \rangle$  in agreement with (11.18).

The essential feature of the corpuscular theory is that the intensification of light implies an increase by the corresponding factor of the number of particles in the small chamber and not the elongation of the scale of the ordinate as was the case in the wave theory.

## (iii) THE FLUCTUATION WITH PLANCK'S FORMULA

The fluctuation with Planck's formula, as given by Eq. (11.10), consists of two terms, of which one is of the Rayleigh-Jeans type while the other is of Wien's type. We, therefore, see that the actual light can

neither be considered as a pure wave nor as a pure particle. The phenomena of interference and diffraction of light to be discussed in the later sections also contradict the pure particle interpretation of light. Actual light has both natures and in the case where Rayleigh-Jeans' formula holds approximately it behaves like a wave, while in the case where Wien's formula holds it behaves predominantly like a particle.

The same sort of argument as applied to the energy fluctuation can be applied to more realistic problems. In fact, Einstein has given the following case as an example. When a particle is floating immersed in black body radiation, the radiation exerts some pressure on it. On the average the net force due to this radiation pressure is zero since the pressure is, on the average, the same in all directions. The actual instantaneous pressure, however, fluctuates around the average value; and hence, for example, the pressure from the right at one time is larger than that from the left and at some other time the pressure from above is larger than that from below. Consequently, the particle cannot stay at rest but moves slightly in a random fashion. The calculation of this random motion shows that it is of the type due to the action of random waves when the Rayleigh-Jeans formula is applied while it is of the type due to the random impacts of particles when Wien's formula is applied. In the latter case the motion resembles the Brownian motion of colloidal particles. The use of Planck's formula gives a random motion consisting of both of the above types. Therefore, the action of light on this particle also shows that the light behaves like a wave or like a particle according as Rayleigh-Jeans' or Wien's formula holds approximately and that it behaves as if it has both of these natures when Planck's formula holds.

What, then, is light which has a wave nature on the one hand and a particle nature on the other? The waves and the particles that we are acquainted with are of such a nature that one cannot simultaneously be the other. It is quite obvious that if something is a wave it cannot be a particle and vice versa. The solution of this problem leading to a proper theory of light, finally came about 20 years after Einstein's proposal in 1905 of his corpuscular theory of light. This theory of light involves a large amount of advanced mathematics. A penetrating view through the mathematical barriers will reveal the way in which light can simultaneously have the dual natures of wave and particle. Furthermore, quantum mechanics shows that this dual nature of wave
and particle is a property not only of light but is quite universal and applies also to matter. Hence the particle aspect of light signified by "the quantum of light" does not have any essential difference from other ordinary particles and the name "photon" begins to replace "the quantum of light" and classes light as a particle just like an electron, proton, neutron or meson. We shall in this volume gradually explain how all this comes about.

## § 12. The Photoelectric Effect

Since the electromagnetic theory of light has lost its firm basis, the phenomena in which light is absorbed or emitted by matter, or more generally in which light energy is transformed to other forms of energy, are now unexplained and we have no clue for finding the laws which they obey. It was made clear in § 10 that the continuous transformation, according to Maxwell's theory, of light energy into other forms is entirely incompatible with the assumption of a light quantum. It has been found that light, when it approximately obeys Wien's formula, has the nature of a particle; more specifically, when the light falls on a small colloidal particle the fluctuation of the radiation pressure is of the type that one would expect if the pressure were due to colliding particles, and hence the colloidal particle shows a motion similar to Brownian motion. This is an extremely important finding since it gives a promising clue for the solution of the problem. In ordinary light sources such as an arc, the temperature is not very high and, hence, Wien's formula holds quite well for the ultraviolet frequency region. Therefore, one can expect that, when ultraviolet radiation from such light sources falls on matter, its effect on matter will be just as if it were a stream of particles.

It is known experimentally that when light, ultraviolet radiation, or X-rays of still higher frequencies, impinge on a material surface, such as a metal surface, electrons are emitted. This effect is called the photoelectric effect and has been known since the end of the last century. But the detailed investigations of this effect were first done by Lenard at the beginning of this century. Lenard's experimental findings are as follows:

1. The energies of individual electrons emitted are independent of the radiation intensity.

2. The intensification of the radiation results in an increase in the

[CH. 2, § 12

number of emitted electrons, more specifically the number per unit time, or in other words the frequency of events.

3. The energy of an individual electron depends on the color of the irradiating light; the shorter the wave length of the radiation, the more the energy of the emitted electron.

It is extremely difficult to explain these three experimental facts by the wave theory of light. If the energy of the emitted electron is acquired from incident radiation, this energy is expected to increase with increasing intensity of the radiation unless some extremely artificial, or unrealistic, mechanism is introduced to avoid it. If one imagines what happens in the intra-atomic domain, it would be as follows. The electrons would be bound in the atom under the influence of some sort of force such as an elastic force. When a light wave comes, the electron moves synchronously with this incident wave and the amplitude of motion becomes larger and larger, the electron acquiring energy from the wave until finally it jumps out of the atom, i.e., out of the range of the binding elastic force. This is supposed to be the electron emitted in the photoelectric effect. Considering in this way, the energy of the electron emitted is to be due to the last one or one half cycle of the light wave just before it jumps out of the atom, since the earlier cycles would simply increase the amplitude of oscillation inside the atom. The energy available in this last one or one half cycle is naturally larger when the amplitude of the incident light wave is larger. This brings up a difficulty of the wave theory since it not only contradicts the first of Lenard's findings but also is unacceptable in view of the amount of energy the emitted electron carries. For Lenard observed that the energy of an emitted electron from the irradiation by an ordinary arc light is of the order of  $5 \times 10^{-12}$  erg. In order for the electron to acquire this much energy during a short period of one or two cycles, the intensity, in the sense of the wave theory, must be immensely strong; a quantitative estimate of this intensity will be given at the end of this section. Lenard has thus concluded that some new and more intricate mechanism has to be introduced for the theoretical interpretation of the experimental results.

Einstein, on the other hand, considered that the above three findings of Lenard were nothing but a clear exposition of the corpuscular nature of light. In the preceding section we have discussed the essential differences of the wave theory and the corpuscular theory of light in relation to the fluctuation of radiation; see Figs. 14 and 15. According

to these discussions, the intensification of light by a factor of two in the wave theory implies that the amplitude of a light wave is enlarged by a factor of  $\sqrt{2}$  and, consequently, that the force acting on the electron is strengthened by the same factor of  $\sqrt{2}$ . On the other hand the same intensification of light in the corpuscular theory does not imply that the energy of an individual photon is increased but instead that the number of photons is increased proportionally. Therefore the first and the second of the experimental findings of Lenard that only the number of emitted electrons increases with increasing intensity of the irradiating light while their energies remain unchanged correspond precisely to the corpuscular nature of light. According to Einstein, not only the first and the second but also the third of Lenard's experimental findings can be understood on the basis of the corpuscular theory since the energy of light quantum is  $h\nu$ , and hence light of shorter wave length consists of larger energy quanta. In fact, the wave length of violet light is around  $4 \times 10^{-5}$  cm and its frequency  $\nu_{violet}$  is then given by

$$v_{\rm violet} = 0.8 \times 10^{15} \, {\rm sec}^{-1}$$

which leads to

$$h\nu_{\rm violet} = 5.2 \times 10^{-12} \, {\rm erg}$$

for the energy quantum of violet light.

This value is in good agreement with the experimental value of  $5 \times 10^{-12}$  erg. This agreement then furnishes the following interpretation for the nature of the photoelectric effect.

The photoelectric effect is the phenomenon where a photon strikes an atom giving its energy  $h\nu$  to the atom from which an electron, as a consequence, is emitted with approximately the same amount of energy as carried by the original photon.

In this expression, however, the phrase "approximately the same amount of energy" requires further specification. In order for an electron to be ejected from inside an atom into the region outside the atom and further through the surface of the object into the outer free space, a certain amount of separation energy must be supplied. Hence, the observed energy of an electron, E, is smaller by this amount, W, than that of the photon, hv, and we have the relation

$$E = hv - W . \tag{12.1}$$

Eq. (12.1) is a very simple one but nevertheless is very important and expresses the corpuscular nature of light in a most simple and clear-cut fashion.

The experimental confirmation of this relation, the so-called Einstein's relation, had to wait, however, for another ten years or so. It was accomplished by Millikan's experiment in 1916. He separated various spectral lines of a mercury lamp and used them individually as monochromatic light sources, while in Lenard's experiment the light sources, being arcs between various electrodes, were not quite monochromatic. Millikan then measured the maximum energy of the emitted electrons for each of these monochromatic spectral lines and obtained its dependence on the frequency of light. The result is shown in Fig. 16.

The figure is for lithium as the irradiated metal. The light sources are the spectral lines of a mercury lamp as explained above. Five spectral lines altogether are used and the results with these lines are indicated by the five points in the figure, where the frequency of the



Fig. 16. The relationship between the energy of a light quantum and the frequency of light. Abscissa shows the frequency (1/sec), ordinate the energy of the light quantum (voltage of Li-plate).

spectral line is plotted on the abscissa while on the ordinate is plotted the maximum energy of the emitted electron. This is measured in the following manner. In front of the lithium plate is placed a device to receive the emitted electrons, a Faraday cage to be specific. The number of electrons can be determined by observing the current through the cage. The maximum energy of the electrons, which we are after in this experiment, can be measured by giving a certain known electric potential V to the lithium plate while the cage is kept earthed. Since the electron has a negative charge -e, it cannot reach the cage if its energy is less than eV, this giving no current through the cage. Consequently, by adjusting the potential V so that the current through the cage just stops, we can measure the maximum energy of the electron by the relation

$$E = eV. (12.2)$$

The energy E, measured in this way, is to be compared with the electron energy in Einstein's relation, Eq. (12.1). According to this equation, all the electrons are considered to have the same energy, because it is assumed that there is no energy loss other than the separation energy. Most of the electrons however are liable to lose part of their energy, for instance, by giving a certain amount of excitation energy to the atom on their way to the Faraday cage. Since there are no conceivable processes which give rise to an increase in the electron energy, Eq. (12.1) gives the maximum amount of energy with which the electron can arrive at the cage and only that fraction of the electrons which luckily did not suffer an energy loss on their way can have this maximum energy.

We thus have, from Eqs. (12.1) and (12.2),

$$eV = h\nu - W \,. \tag{12.3}$$

This linear relation between V and v is shown by the straight line in Fig. 16. The slope of this line gives a determination of h. Millikan obtained from this experiment

$$h = 6.58 \times 10^{-27} \, \text{erg} \cdot \text{sec}$$
.

This value is in good agreement with that of Planck. Einstein's hypothesis of light quantum has herewith become an unquestionable fact.

It would be worthwhile at this point to show a little more quantitatively how difficult it is to explain this experimental fact from the standpoint of the wave theory of light. The amount of energy flowing out of a light source of 1 candle is of the order of 1 erg per second through an area of 1 cm<sup>2</sup> placed at a distance of 3 meters. Assuming that the radiation consists mostly of yellow light, the frequency of which is about  $0.5 \times 10^{15}$ /sec, we find that the light quantum hv is  $3 \times 10^{-12}$  erg and that approximately the same amount of energy is to be imparted to the electron ejected by the photoelectric process. This is the fact verified by Lenard's other experiments. Now according to the wave theory of light, the energy is considered to be distributed continuously in space. Hence it takes quite a long time for the atom to absorb this much energy, unless the atom has a sort of supernatural method of collecting energies instantly from an extended region of space. In fact, considering an atom placed at a distance of 3 meters from a unit candle light source, the amount of energy flowing through the cross-section of this atom, approximately  $10^{-8}$  cm in radius, is about  $10^{-16}$  erg per second. Therefore, it takes at least around  $30\ 000$  second for this atom to absorb  $3 \times 10^{-12}$  erg of light energy, even if the atom is assumed to absorb the entire light energy which passes through its cross-section. The experimental fact, however, is that the photoelectric effect occurs instantaneously with the incidence of the light.

A similar consideration applies also to photo-chemical reactions. When light is shining on a photographic plate, the light quanta are impinging on a AgBr molecule, giving their energy  $h\nu$  to dissociate it. The dissociation of a AgBr molecule requires a certain minimum amount of energy and hence radiation which has too small a frequency can not react with the photographic plate. Furthermore, since this minimum energy is of the order of  $10^{-12}$  erg, it takes, according to the wave theory of light, an extremely long exposure time for the light of ordinary intensity to be recorded on the photographic plate. The same argument holds also for human vision and, if light did not have a corpuscular nature, one could never, throughout the whole course of his life, register faint star-light with his eyes.

In connection with photo-chemical reactions, it is possible to estimate the magnitude of Planck's constant h from common daily phenomena. It is well-known that ultraviolet rays are most effective in causing photo-chemical reactions. We know that if we face a redburning stove for a long period of time our faces do not become sunburned; but only five minutes or so on a high mountain or at a summer beach is enough to give a good sun-burn. Sun-burning implies that a certain kind of photo-chemical reaction has taken place in the skin. When the minimum energy required for this reaction exceeds hv of the irradiating heat rays, sun-burning does not occur no matter how large the amount of irradiation may be. On the other hand, since hv of the ultraviolet rays in the sun-shine is sufficiently large, this reaction can readily occur. This is the reason why we get sun-burnt on high mountains or at beaches. Now an estimate of the minimum energy can be made by observing a few simple chemical reactions. For instance, the threshold energies of the following reactions are known;

$$O_2 = O + O - 117$$
 kcal  
AgBr = Ag + Br - 24 kcal

The first of these equations, which plays a role in the photo-chemical formation of ozone from oxygen, implies that a unit mole of molecular oxygen requires 117 kcal of energy for its dissociation into atomic oxygen. The number of molecules per unit mole is  $N = 6.02 \times 10^{23}$  and 1 kcal is  $4.18 \times 10^{10}$  erg. The threshold energy for a molecule to be dissociated is therefore approximately

$$E = \frac{4.18 \times 10^{10} \times 117}{6.02 \times 10^{23}} \simeq 8.0 \times 10^{-12} \text{ erg}.$$

The wave length of ultraviolet rays being roughly 2  $\times$  10<sup>-5</sup> cm, its frequency is 1.5  $\times$  10<sup>15</sup> sec<sup>-1</sup> and

$$h\nu_{\text{violet}} = h \times 1.5 \times 10^{15} \text{ erg}$$
.

Since the reaction is known to start in the ultraviolet region, we equate the above two and

$$E \approx h v_{violet}$$

from which we obtain

$$h \approx 5.4 \times 10^{-27} \,\mathrm{erg} \cdot \mathrm{sec}$$
.

The close agreement of this value with that of Planck,  $6.62 \times 10^{-27}$  erg·sec, should be regarded as fortuitous in view of our rather rough method of estimation, but the approximate agreement can be considered meaningful.

It is now clear that the corpuscular nature of light is neither a pure academic subject in which only a group of queer physicists are indulging themselves nor a phenomenon which requires some gigantic machine for observation. It is revealed in almost every instance of our daily life.

# § 13. The Compton Effect

Another phenomenon in which the corpuscular nature of light is clearly indicated is the Compton effect discovered by Compton in 1923. The phenomenon itself had been known some time before and Compton discovered that it was due to the corpuscular nature of light.

It is concerned with the scattering of X-rays. X-rays are known to be scattered by material objects. When, using a monochromatic X-ray source, we measure the wave length of the scattered X-ray accurately, we find that there exist some X-rays of longer wave lengths than that of the incident one, together with the one of equal wave length. Theoretically Thomson had treated this scattering on the basis of Maxwell's theory. According to this theory, the scattered X-ray should have exactly the same wave length as the incident X-ray. For, according to Maxwell's theory, the scattering occurs in the following fashion. When the electromagnetic wave of incident X-rays hits the atom, the electrons in the atom undergo forced oscillations due to the periodic electric force, the non-periodic force of the radiation being negligibly small for practically available X-ray amplitudes, because the latter is (like radiation pressure) proportional to the square of the amplitude while the former is proportional to the amplitude itself. The electrons then emit spherical waves as the result of these forced oscillations, and these waves are recognized as the scattered X-rays. The period of the oscillatory motions of the electrons will be the same as that of the incident wave; the change of frequency of the scattered wave is possible only when the electrons undergo translational motion by the non-periodic part of the force but this is negligibly small. This is what Thomson's theory of X-ray scattering essentially is.

The experimental result, however, is that a part of the scattered X-rays has smaller frequencies, while the frequency of the other part remains the same. Compton performed a very precise experimental examination of this effect and showed that his results can most simply be explained by the corpuscular theory of light. Almost simultaneously with Compton's finding, Debye also published the same idea. Before giving the details of the theory of Compton and Debye, we must understand something about the momentum of a light quantum.

# (i) THE MOMENTUM OF A LIGHT QUANTUM

We have shown that a light quantum is a particle carrying an energy hv. This particle must also carry momentum, since it is known to

сн. 2, § 13]

65

excrt some pressure on any wall by which it is reflected. In the framework of the wave theory, we showed that the pressure P is related to the energy density U by

$$P = \frac{1}{3}U\tag{13.1}$$

provided that the light is incident isotropically on the wall. This relationship was not only verified directly by experiments but also was the basis of our derivation of Stefan's law and Wien's displacement law in previous sections. We, therefore, have to regard this relationship as a correct one.

We now determine the momentum of a light quantum in such a way that this relationship is arrived at also in the corpuscular theory.

The simplest method of doing this is as follows. We consider a cubic cavity of side L and begin with only one photon of energy hv in it. This photon will perform a zig-zag motion with the light velocity c inside the cavity as shown in Fig. 17. Whenever it is reflected by one of the walls, AB to be specific, it exerts a force on the wall.

Taking the x-axis normal to the wall AB as in Fig. 17 and denoting the momentum of the photon, its x-component, and the x-component of the velocity by p,  $p_x$  and  $c_x$ , respectively, we find that the number of collisions per unit time is  $c_x/2L$  and that the change of particle momentum in each collision is  $2p_x$ . Hence, the amount of momentum transfered per unit time by the photon to the wall AB is  $p_x c_x/L$ . This further implies that the average force on the wall AB is  $p_x c_x/L$ . The pressure on the wall due to this photon is then



Fig. 17. The motion of a light quantum in a cavity.

$$P_1 = \frac{p_x c_x}{L^3}$$

where the subscript 1 on P indicates that the pressure is due to one photon. We now consider the case where a large number of photons, N, are moving in the cavity. The pressure due to all these particles can be obtained by adding the pressures due to each individual particle and we get

$$P = \frac{1}{L^3} \sum p_x c_x \, .$$

Denoting the value of  $p_x c_x$  averaged over all the particles by  $\langle p_x c_x \rangle$ , it becomes

$$P = \frac{N}{L^3} \langle \dot{p}_x c_x \rangle . \tag{13.2}$$

The relationship between the particle number N and the energy density U in the cavity is

$$N=\frac{UL^3}{h\nu}\,.$$

On the other hand, it is natural to consider that the momentum of a particle is in the same direction as its velocity and hence

$$p_x = p \frac{c_x}{c} \, .$$

Using these relations together with  $\langle c_x^2 \rangle = \frac{1}{3}c^2$ , which implies the nonpreference of either the x, y or z axis, Eq. (13.2) can be transformed to

$$P = \frac{U}{3} \frac{c\phi}{h\nu} \,. \tag{13.3}$$

In order for this to coincide with the experimentally verified relation, Eq. (13.1), p must be given by

$$p = \frac{h\nu}{c} \,. \tag{13.4}$$

We have thus determined the momentum of the photon of energy hv. Namely the momentum of a photon is equal to its energy hv divided by the light velocity c.

# (ii) THE QUANTUM THEORY OF X-RAY SCATTERING

We have learned already how to treat X-ray scattering in the framework of Maxwell's theory. In this section we consider the problem on the basis of the corpuscular theory of light. Among the various possible ways of interpreting the phenomenon, the simplest will be the following.

When an X-ray particle collides with an electron, it will be scattered in a direction different from the original one. The change in the direction of motion necessarily involves a change of the momentum. If the total momentum is to be conserved in the collision, the struck electron

66

must take up the recoil momentum which exactly balances the difference of the X-ray momenta before and after the collision. The energy conservation, which must also be satisfied, requires the energy of the scattered X-ray, hv', to be smaller than that of the incident X-ray, hv, by the amount carried away by the recoil electron. Thus the scattered X-ray has a smaller frequency, i.e., a larger wave length, than that of the incident X-ray. This process is nothing but an elastic scattering of an X-ray particle with an electron.

We now proceed to treat the problem more quantitatively. Suppose that an X-ray of frequency  $v_0$  collides with an electron of mass minitially at rest and gets scattered by it. A schematic representation of the problem is shown in Fig. 18. Let the frequency of the X-ray scattered in the direction  $\varphi$  be denoted by  $v_{\varphi}$ , and the momenta of the



Fig. 18. The elastic collision of an X-ray particle and an electron.

incident and scattered X-ray by  $h\nu_0/c$  and  $h\nu_{\phi}/c$ , respectively. The momentum of the recoil electron will then be the vector difference of these two momenta. Denoting the ratio of the velocity of the recoil electron to the light velocity c by  $\beta$ , and taking the theory of special relativity into account, the momentum of the electron is given by  $m\beta c/\sqrt{1-\beta^2}$ . The vector relation between the above three momenta then is expressed by

$$\left(\frac{m\beta c}{\sqrt{1-\beta^2}}\right)^2 = \left(\frac{h\nu_0}{c}\right)^2 + \left(\frac{h\nu_{\varphi}}{c}\right)^2 - 2\frac{h\nu_0}{c}\frac{h\nu_{\varphi}}{c}\cos\varphi . \quad (13.5)$$

The conservation of energy implies that the energy of the scattered X-ray,  $h\nu_{\varphi}$ , is equal to that of the incident one,  $h\nu_{0}$ , minus the relativistic kinetic energy of the recoil electron,

$$mc^2\left(\frac{1}{\sqrt{1-\beta^2}}-1\right).$$

Thus

$$h\nu_{\varphi} = h\nu_0 - mc^2 \left(\frac{1}{\sqrt{1-\beta^2}} - 1\right).$$
 (13.6)

Solving the two simultaneous equations, Eq. (13.5) and Eq. (13.6), for the two unknown quantities,  $\beta$  and  $\nu_{\varphi}$ , we obtain

$$\nu_{\varphi} = \frac{\nu_0}{1 + \frac{2h\nu_0}{mc^2}\sin^2\frac{1}{2}\varphi}$$
(13.7)

for the frequency of the X-ray scattered in the direction  $\varphi$ . Eq. (13.7) implies that  $\nu_{\varphi}$  never exceeds  $\nu_0$ ; the frequency of the scattered X-ray can never be larger than that of the incident X-ray.

In terms of the wave lengths, Eq. (13.7) takes a simpler form

$$\lambda_{\varphi} = \lambda_0 + \frac{2h}{mc} \sin^2 \frac{1}{2}\varphi \qquad (13.8)$$

where  $\lambda_0$  and  $\lambda_{\varphi}$  signify the wave lengths of the incident X-ray and of the X-ray scattered in the direction  $\varphi$ , respectively.

The energy of the recoil electron is now readily obtained from Eqs. (13.6) and (13.7) as

$$E_{kin} = h(\nu_0 - \nu_{\varphi}) = h\nu_0 \frac{\frac{2h\nu_0}{mc^2} \sin^2 \frac{1}{2}\varphi}{1 + \frac{2h\nu_0}{mc^2} \sin^2 \frac{1}{2}\varphi}.$$
 (13.9)

Furthermore, the angle  $\Theta$  between the directions of motion of the recoil electron and the incident X-ray is given by

$$\cot \Theta = -\left(1 + \frac{hv_0}{mc^2}\right) \tan \frac{1}{2}\varphi. \qquad (13.10)$$

The change of wave length in the X-ray scattering,  $\Delta\lambda$ , is, according to Eq. (13.8), expressed as

$$\Delta \lambda = \lambda_{\varphi} - \lambda_{0} = \frac{2h}{mc} \sin^{2} \frac{1}{2}\varphi . \qquad (13.11)$$

This quantity  $\Delta \lambda$  depends only on the scattering angle  $\varphi$  and not on the wave length of the incident X-ray. Therefore, the spectrum of the

scattered X-ray at a certain angle is expected to be uniformly shifted from that of the incident ray by a constant amount. The quantity h/mc in this formula has the dimension of length and is generally called the Compton wave length. The numerical value of the Compton wave length is

$$\frac{h}{mc} = 2.42 \times 10^{-10} \text{ cm} = 0.0242 \text{ Å}.$$

# (iii) COMPARISON WITH EXPERIMENTS

The scattering by graphite of the X-rays from an X-ray tube with a molybdenum anti-cathode has been investigated. In Fig. 19 the solid

line shows the spectral intensity of the scattered X-rays in the direction  $90^{\circ}$  to the incident X-ray. This is to be compared with the spectral intensity of the incident X-ray which is also shown in the same figure by the dotted line. The figure clearly indicates that the spectrum of the scattered X-ray is shifted from that of the incident uniformly toward larger wave lengths by a constant amount.

From this experiment, the amount of the shift is estimated to be

$$\Delta \lambda_{exp} = 0.022 \text{ Å}$$

which is in good agreement with the theoretical value, Eq. (13.11),

$$\Delta \hat{\lambda}_{\text{theory}} = 0.024 \text{ Å}$$
.

#### (iv) THE RECOIL ELECTRON

Although Compton's theory was found to be in agreement with experiment, a conclusion necessarily resulting from this theory, i.e., the existence of the recoil electrons, had not yet been verified. However, several months after Compton proposed his theory, Wilson and Bothe independently observed the recoil electrons.

We now briefly describe the experimental method used by them, since their method, using a cloud chamber, is extremely important in studies of atomic physics.

Electrons and  $\alpha$ -particles are so small in size that they cannot be



Fig. 19. The spectrum of the scattered X-ray in the Compton effect.

observed directly even with the finest microscope. There, however, is an experimental method by which we can visualize the passage in space of these tiny charged particles. This is the famous method of the cloud chamber discovered by Wilson in 1911.

The main part of a cloud chamber is a glass cylinder with a piston inside. The cylinder is filled with air, or perhaps with another gas, generally at about one atmosphere in pressure, and saturated with water vapor or with the vapor of some organic substance such as alcohol. When the piston is suddenly pulled out, the air is expanded, and the temperature of the air decreases suddenly because of this expansion. Since the air is saturated with water vapor, the sudden decrease of the air temperature gives rise to fog formation in the cylinder. A droplet of fog is formed around some small particle in the air, a dust particle for instance, which acts as a nucleus for the fog formation. When the air is extremely purified, the fog cannot be formed in this way. It is said, by the way, that the famous London fog is due to small particles in the smoke from stoves.

Not only the dust particles but also the ions in the air are known to serve as nuclei for the formation of fog.

Wilson noticed this fact and considered that the passage of an electrically charged particle, electron or  $\alpha$ -particle, in such a cloud chamber will leave as a result of collisions with the air atoms, a train of ions along the line of its motion and that these ions will act as nuclei for fog formation. The motion of this particle will then be made visible by these fog-droplets. This idea has been fruitfully realized and the traversal of electron or  $\alpha$ -particle has, in fact, been seen in the cloud chamber as a train of droplets, when the piston is pulled out and intense illumination is given to the chamber from the side. We call this train of droplets the track of the particle.

Wilson and Bothe used this apparatus to observe the recoil electron. They sent X-rays through a cloud-chamber. X-rays will then collide with the electrons of the air atoms and leave the recoil electrons, the tracks of which are observable in the cloud-chamber by the mechanism explained above.

Examples of electron tracks in a cloud chamber are shown in Phot. I, where they are observed as white twisted lines. The lines of the electron motion are so zig-zag and twisted because they frequently collide with the air atoms and change their direction of motion. In the photograph are seen two kinds of electron tracks, one comparatively long and the other shorter. The shorter ones are the tracks of the recoil electrons under consideration, while the longer ones are due to the photo-electric electron ejected from the air atoms as explained in the previous sections. The electrons emitted in the photoelectric effect have a larger energy than the recoil electrons and accordingly are able to travel a longer distance in the air. This is obvious when one recalls the fact that the recoil electron receives only a part of the energy of the incident X-ray in the Compton effect while in the photoelectric effect all the energy is transfered to the emitted electron.

In going through the air, the electron collides continually with the air atoms giving up a part of its energy to them in each impact and its energy is gradually decreased until it finally comes to rest. The distance an electron can traverse before it comes to rest depends on the energy it has at the beginning. Conversely, one can estimate the initial energy of the electron from the observed length of its track by means of the theoretically calculated relationship between the two quantities. The actual estimate of the electron energy by this method revealed that the electrons of longer path had, in fact, an energy  $hv_0$  at the beginning while those of shorter path had an energy satisfying Eq. (13.9).

Compton's theory thus received an additional confirmation.

#### (v) THE COMPTON-SIMON EXPERIMENT

Compton in collaboration with Simon carried out another experiment in order to still more firmly establish his theory. When an electron is hit by an X-ray and suffers a recoil, the X-ray particle itself gets deflected. Unfortunately, the X-ray particle, unlike an electron, does not leave a visible track in the cloud-chamber and one can not identify the direction of its motion. In certain favorable cases, however, the scattered X-ray will give rise to another Compton effect before it leaves the chamber. The observation of these successive Compton effects will enable us to infer the direction of the scattered X-ray. We then can estimate the three quantities, (1) the velocity of the recoil electron, from its track length, (2) the direction of the recoil electron, which is given by the tangential direction of the beginning of its track, and (3) the direction of the scattered X-ray, from the position of the second Compton effect relative to the first. Since the direction and the frequency, and hence the momentum and the energy, of the incident X-ray are known, we can experimentally verify the vector relation shown in Fig. 18.

A photograph of such an event is shown in Phot. II. In Fig. 20 a schematic drawing of the experiment is given, the two wavy lines



Fig. 20. The relation between successive re-Compton-Simon experiment. A shows X-ray § 14. The Photon Gas tube, B filter, C Pb screen, and D cloud chamber.

indicating the tracks of the two recoil electrons. From the beginning part of the track near the center, one finds that the first recoil electron was ejected in the direction of angle  $\Theta$  with respect to the incident X-ray. The scattering angle of the X-ray particle,  $\varphi$ , is experimentally measured by the position of the second Compton effect relative to the first. In the figure is shown also the scattering angle  $\varphi'$  theoretically expected from Eq. (13.10). The two angles  $\varphi$  and  $\varphi'$  were found to coincide within the small experimental error  $\Delta \varphi$ .

In this way, the corpuscular nature of an X-ray is clearly shown by its behavior in a collision with an electron, this being similar to coil electrons in the the behavior of billiard balls on a billiard table.

According to the preceding sections the corpuscular nature of light now seems to be well establi-

shed. In the corpuscular theory, light is considered to consist of particles flying in space with the velocity of light and the difference in the color of the light is ascribed to the difference in the energy of the corresponding particles. For instance, red light consists of the particles of energy  $hv_{\rm red} = 2.6 \times 10^{-12}$  erg and violet light of those of energy  $h\nu_{\text{violet}} = 5.0 \times 10^{-12}$  erg.

From this standpoint, the problem of black body radiation is nothing but the ordinary problem of the energy distribution of gas molecules. Since black body radiation can be considered simply as a photon gas in a box immersed in the reservoir of temperature T, the problem of determining the spectrum of black body radiation in thermal equilibrium is the same as asking for the energy distribution of this photon gas. When the ideal gas of temperature T is confined in a box of volume V, the number of gas molecules of momenta between p and p + dp in magnitude is known to be given by

$$dn = A e^{-\frac{1}{kT} E(p)} p^2 dp$$
 (14.1)

#### THE PHOTON GAS

where A is the normalization factor to be determined so as to make the total number of the gas molecules equal to the given N, i.e., to satisfy the relation

$$N = A \int_{0}^{\infty} e^{-\frac{1}{kT} E(p)} p^{2} dp. \qquad (14.2)$$

In the case of an ordinary gas, we have

$$E(p) = \frac{p^2}{2m}$$

and Eq. (14.2) gives

$$A=\frac{4\pi N}{(2\pi m kT)^{\frac{3}{2}}}\,.$$

When the theory of special relativity is taken into account, i.e. when  $E(p) = c \sqrt{p^2 + m^2 c^2}$ , this becomes however

$$A = \frac{N}{\int_{0}^{\infty} \exp\left(-\frac{c}{kT}\sqrt{p^{2} + m^{2}c^{2}}\right)p^{2} dp}.$$
 (14.3)

In either case, Eq. (14.1) is well known as Maxwell's distribution law.

Unfortunately, the elementary method which is usually used to derive Maxwell's distribution law cannot be applied to the case of the photon gas, since the total number of photons in the box is not constant due to the absorption and the emission by the wall while in the case of ordinary gases it always is kept at a given constant value. The instantaneous total number in the photon gas case not only fluctuates around its average value but also the average value itself changes with changing temperature as will be given below by Eq. (14.6).

We do not, therefore, try to develop the kinetic theory of the photon gas but, instead, will just demonstrate in the following that Wien's formula corresponds to Maxwell's distribution law for ordinary gases.

Wien's formula gives for the energy density of light in the frequency range from v to v + dv

$$U(\nu) \, \mathrm{d}\nu = \frac{8\pi h}{c^3} \,\mathrm{e}^{-h\nu/kT} \,\nu^3 \,\mathrm{d}\nu \;. \tag{14.4}$$

In order to compare this formula with Maxwell's distribution law, Eq. (14.1), we should convert it to an expression giving the number of

photons in the same frequency range. This conversion can be accomplished by multiplying Eq. (14.4) by the volume of the Box, V, and then by dividing by  $h\nu$ . The result is,

$$dn = \frac{8\pi V}{c^3} e^{-h\nu/kT} \nu^2 d\nu. \qquad (14.5)$$

The total number, N, of photons in the box is then given by

$$N = \frac{8\pi V}{c^3} \int_0^\infty e^{-h\nu/kT} \nu^2 \, \mathrm{d}\nu = \frac{16\pi V}{c^3} \left(\frac{kT}{h}\right)^3.$$
(14.6)

Introducing Eq. (14.6) into Eq. (14.5), we obtain an expression which does not involve V, i.e.,

$$dn = \frac{Nh^3}{2(kT)^3} e^{-h\nu/kT} \nu^2 d\nu . \qquad (14.7)$$

This is to be compared with Maxwell's distribution law, Eq. (14.1).

We already have the relativistic form of the distribution law, Eq. (14.1) together with Eq. (14.3). In these formulae, we put m = 0 to make them applicable to the case of the photon gas. One recalls that for a photon E = hv and p = hv/c so that E = cp. This correspond to putting m = 0 in  $E = c\sqrt{p^2 + m^2c^2}$ . Eq. (14.3) then becomes

$$A = \frac{Nc^3}{2(kT)^3}.$$
 (14.8)

Introducing this into Eq. (14.1), we obtain

$$dn = \frac{Nc^3}{2(kT)^3} e^{-E/kT} p^2 dp . \qquad (14.9)$$

This Eq. (14.9), Maxwell's distribution law for the gas consisting of zero mass molecules, is easily shown to be identical to Wien's formula, Eq. (14.7), when one makes use of Einstein's relationship for a photon: hv = E and hv/c = p.

From this consideration, we clearly understand that Wien's formula is based entirely on the corpuscular theory of light, while, as explained before, the Rayleigh-Jeans formula is obtained from the wave theory of light. The distribution law which actually is observed in nature however is neither of these but instead is given by Planck's formula. Therefore, we can consider light neither as of entirely corpuscular nature nor as of totally wave nature.

# § 15. The Corpuscular Nature and the Wave Nature

In the last several sections, we described various phenomena which preferentially revealed the corpuscular aspect of light. It also has frequently been stressed that we cannot consider light as purely of corpuscular nature and that light has the dual nature of particle and wave. In this section, we describe how inconceivable and how full of inconsistencies this dual nature of light seems to us.

The wave interpretation of light was introduced primarily to explain the interference effects with light. To be specific, we recall Young's

experiment with a double slit as shown in Fig. 21. A monochromatic light source, which for simplicity we regard as a point source, is placed at the position O in the figure. A screen with two slits,  $S_1$  and  $S_2$ , is placed between the light source and the photographic plate P. They are all firmly fixed on a bench. The

photographic plate then shows a characteristic pattern of interference. In order to obtain a clear and distinct interference pattern the distance between the two slits,  $S_1$  and  $S_2$ , must not be too large. But this point is not essential to our consideration here.

The appearance of the interference pattern implies that the light ray passing through one of the slits is affected by the presence of the other slit. In fact, the photograph of the interference pattern with the two slits both open is entirely different from the superposition of the diffraction patterns of the two slits open one at a time. In Phot. III, (a) shows the diffraction pattern due to one slit while the interference pattern due to two slits is labelled (b). It is quite obvious that one can not obtain the pattern (b) from any superposition of two (a)'s. This remarkable fact of interference is well known to be explainable without any artificial hypotheses by considering light as a wave.

But if we were to assume that light is corpuscular, one half of the light particles will pass through the slit  $S_1$  and the other half through the slit  $S_2$  but one particular light particle has to pass through only one of the two slits, either  $S_1$  or  $S_2$  but not both. This is because the



Fig. 21. Double-slit experiment.

particle is an entity of energy hv which can not be further divided. It is then impossible to have a part of the single particle passing through S<sub>1</sub> while the other part passes through  $S_2$ . If one places a cloud-chamber behind the slits and observes the photoelectric effect or the Compton effect, he will never find a fraction of the energy  $h\nu$  causing these effects. The existence or non-existence of the slit, S<sub>2</sub>, therefore, seems to have no effect on the particles passing through  $S_1$ . It is then expected that the particles passing through the slit  $S_1$  will give rise to a pattern on the photographic plate irrespective of whether the other slit S<sub>2</sub> is open or not. The same will be true for the particles passing through the slit S2. Hence, the interference pattern due to the two slits  $S_1$  and  $S_2$  is to be nothing but a superposition of the patterns due to  $S_1$  and  $S_2$  open one at a time. This difficulty may seem to be overcome by considering that a particle passing through one slit interacts with another particle passing through the other slit in such a way as to produce the observed interference pattern. However, the interference pattern can be observed also with such a faint light source that at most only one particle can exist at one time between O and P in Fig. 21, provided a sufficiently long exposure is given. It is then quite unnatural to consider that a particle interacts with another still unborn at O or with another already absorbed at P. Hence, this kind of interpretation of the interference pattern has to be abandoned.

Let us consider another example of interference phenomena. This is the experiment with Michelson's interferometer as shown in Fig. 22.



The interferometer consists of a light source, O, a semi-transparent mirror H, two mirrors  $M_1$  and  $M_2$ , and a photographic plate P all properly fixed on a solid frame, see Fig. 22. The light from the source O is split by the semi-transparent mirror H, a part reflected in the direction of the mirror  $M_1$  and the rest through H in the direction of the mirror  $M_2$ . They are both reflected by  $M_1$  and  $M_2$ , respectively. Again at H, a part of the light from  $M_1$  goes to the photographic plate P and the rest to O and

a similar splitting occurs also with the light from M<sub>2</sub>. The light arri-

ving at P and O is accordingly a superposition of the light rays from  $M_1$  and  $M_2$ . The intensities of light at P and O, therefore vary, depending on the difference between the two light paths, HM1H and HM<sub>2</sub>H; according as the path difference is an even or odd multiple of the wave length, the two light rays going to P may constructively interfere with each other while those going to O destructively, or vice versa. Therefore, it may happen that all the light emitted by O arrives at P, or that it is all sent back to O giving no light at P. The occurrence of one of these possibilities is determined by a very delicate condition specifying the position of the mirror M<sub>2</sub> within an accuracy of a wave length. In this experiment, it does not matter, in principle, how remote the mirrors  $M_1$  and  $M_2$  are from H. They can even be several hundred meters, say, away from H. If the light were particles, how can one expect an interference effect? For the particle at M<sub>1</sub>, what effect can be expected from a minute displacement, of the order of the wave length, of M<sub>2</sub> which is located some hundred meters away? The answer is obviously none and the same is true for the particles at M<sub>2</sub>. There is therefore no other way but to consider that always a part of the light goes to P and the rest back to O and that the shares are the same irrespective of the positions of  $M_1$  and  $M_2$ .

As revealed in the considerations above, the corpuscular theory of light is entirely powerless to explain interference effects just as the wave theory is incapable of explaining the photoelectric effect. Light must have the dual nature of particle and wave.

Our concept of particle and wave however does not allow the possibility that something is both a particle and a wave. If something is a particle, it can not be a wave at the same time, and vice versa. The correct explanation of this dual nature was given later, in 1926–1927. Until then physicists had to take the inconsistent attitude of considering light as a wave in some cases and as a particle in other cases. It is said that during those years physicists considered light as a wave on Monday, Wednesday and Friday, and as a particle on Tuesday, Thursday and Saturday.

#### CHAPTER 3

# THE OLD QUANTUM MECHANICS

#### § 16. The Structure of the Atom

Toward the end of the last century, physicists began, as a result of studies of vacuum discharge phenomena and of radioactivity, to consider that atoms were after all not indivisible ultimate objects but instead had complex structures. First of all, the studies of cathode rays in vacuum discharges established that cathode rays consisted of a stream of particles each having a negative electric charge. The electric charge, -e, and the mass, m, of the particle were measured accurately and turned out to be

 $e = 4.8 \times 10^{-10}$  e.s.u.,  $m = 9.04 \times 10^{-28}$  gram.

Along with these studies, it gradually began to be accepted that this particle, the electron, was present inside the atom as one of its constituents. The mass of an electron is very small and, indeed, is only  $\frac{1}{1836}$  of the mass of a hydrogen atom which is the lightest of all the atoms. Therefore, there must be some other constituent much heavier than the electron inside the atom in order to make up this difference. This heavier substance probably should have a positive electric charge counterbalancing the electron's negative charge so that we get an electrically neutral atom. Our knowledge of the intratomic structure gradually became clearer than this rather vague picture by a coordination of theoretical and experimental studies. We follow this development in this chapter.

#### (i) ZEEMAN EFFECT

The emission of cathode rays from atoms is not by itself proof enough for us to conclude that the electrons, the cathode ray particles, are actually present inside the atom. There is however another phenomenon called the Zeeman effect which gives an additional strong support to this picture. Though we do not know exactly by what mechanism electrons are bound inside the atom or in what states of motion they are, the constituent with positive charge mentioned above must be in a state of mechanical equilibrium with the electrons since the atom is very stable and shows always the same characteristics. The fact that the atom can eventually emit or absorb light would suggest that the electrons can oscillate around this equilibrium state. The Zeeman effect to be described here allows us to conclude that the agent of emission and absorption is actually the same as the cathode ray particle, the electron.

It is well known that the light emitted by the atom has a characteristic spectrum consisting of many lines. For the moment, we consider one of these lines. In some cases, close investigation reveals that, as is observed in the case of the well-known D line of the Na atom, an apparent single line consists of two or more lines but, for the sake of simplicity, we take the line which is really single, or stated more rigorously, one which belongs to a certain singlet term as will be explained in § 27. We now place the atom is a magnetic field and let it emit light. The spectral line which is ordinarily single is now split into three lines; one has the same frequency as before and the other two are displaced from it in opposite directions but by equal amounts. The magnitude of this frequency displacement is found to be proportional to the strength of the applied magnetic field. This phenomenon was found by Zeeman in 1897 and is called the Zeeman effect. Phot. IV is a photograph of the Zeeman effect of the 4680 Å line of Zn taken by Zeeman himself.

Theoretically, this phenomenon is interpreted in the following manner. The charged particle in the atom responsible for the emission and absorption of light undergoes some periodic motion with a certain characteristic frequency determined by the intratomic forces. When the external magnetic field H is applied to this atom, there will be an additional force

$$F = \frac{-e}{c} \left( \mathbf{v} \wedge \mathbf{H} \right) \tag{16.1}$$

acting on the charged particle whose charge and velocity arc denoted by -e and v, respectively.

This law, Eq. (16.1), has been well established experimentally by using the cathode rays. The problem of explaining the Zeeman effect is now reduced to finding the difference in the motion of the particles due to

this additional force, especially the difference in the characteristic frequencies of motion before and after the application of H.

Let the cylindrical coordinates describing the charged particles in the atom be denoted by  $(r_n, \varphi_n, z_n)$ , where the subscript *n* refers to the coordinates of the *n*-th particle. The equations of motion in the absence of the magnetic field for the *n*-th particle are given by

$$\ddot{r}_{n} - r\dot{\varphi}_{n}^{2} = -\frac{1}{m} \frac{\partial V}{\partial r_{n}},$$

$$2\dot{r}_{n}\dot{\varphi}_{n} + r_{n}\ddot{\varphi}_{n} = -\frac{1}{m} \frac{1}{r_{n}} \frac{\partial V}{\partial \varphi_{n}},$$

$$\ddot{z}_{n} = -\frac{1}{m} \frac{\partial V}{\partial z_{n}}.$$
(16.2)

where *m* stands for the mass of the particle and *V*, the potential of the force governing the interatomic world, is a function of  $r_1$ ,  $\varphi_1$ ,  $z_1$ ,  $r_2$ ,  $\varphi_2$ ,  $z_2$ ,... Since the atom is considered to be spherically symmetric, the simultaneous displacement of the angular coordinates,  $\varphi_n$ , of all the particles leaves the potential *V* unchanged, i.e.

$$V(r_1, \varphi_1 + \Delta, z_1; r_2, \varphi_2 + \Delta, z_2; r_3, \varphi_3 + \Delta, z_3; \dots)$$
  
=  $V(r_1, \varphi_1, z_1; r_2, \varphi_2, z_2; r_3, \varphi_3, z_3; \dots)$ . (16.3)

When the external magnetic field H is applied in the direction of the z-axis, the equations of motion are obtained by adding the additional force, Eq. (16.1), to the right hand of Eq. (16.2) and are given by

$$\ddot{r}_{n} - r_{n} \dot{\varphi}_{n}^{2} = -\frac{1}{m} \frac{\partial V}{\partial r_{n}} - \frac{eH}{mc} r_{n} \dot{\varphi}_{n} ,$$

$$2\dot{r}_{n} \dot{\varphi}_{n} + r_{n} \ddot{\varphi}_{n} = -\frac{1}{m} \frac{1}{r_{n}} \frac{\partial V}{\partial \varphi_{n}} + \frac{eH}{mc} \dot{r}_{n} , \qquad (16.2')$$

$$\ddot{z}_{n} = -\frac{1}{m} \frac{\partial V}{\partial z_{n}} .$$

The approximate solution of Eq. (16.2') is easily obtained if we know the solution of Eq. (16.2). Denoting the solution of Eq. (16.2) by

#### сн. 3, § 16]

$$r_{n} = r_{n}^{0} ,$$

$$\varphi_{n} = \varphi_{n}^{0} + \frac{eH}{2mc} t ,$$

$$z_{n} = z_{n}^{0} .$$
(16.4)

where the above property, Eq. (16.3), of the potential V has been made use of.

Eq. (16.4) implies the simple fact that the motion of a particle in the presence of the magnetic field H is obtained from its motion in the absence of the field by simply superposing a rotation with the angular velocity of eH/2mc around the field direction. We call this rotation the Larmor precession and the angular velocity  $\omega = eH/2mc$ the Larmor angular velocity. In terms of the rectangular coordinates, Eq. (16.4) takes the form

$$x_n = x_n^0 \cos \omega t - y_n^0 \sin \omega t ,$$
  

$$y_n = x_n^0 \sin \omega t + y_n^0 \cos \omega t , \qquad (16.4')$$
  

$$z_n = z_n^0 .$$

We thus have obtained the relationship between the motions with and without the magnetic field. The difference in the frequencies of the two motions can now easily be obtained. In view of the fact that the atom emits a line spectrum of light, the coordinates of the particle,  $x_n^0$ ,  $y_n^0$  and  $z_n^0$ , would, in the absence of the magnetic field, have a time dependence like  $\cos(2\pi\nu t)$ , this being the required periodic nature. The frequency  $\nu$  should be the same as the frequency of the emitted light.

When the magnetic field is applied, the introduction of this form for  $x_n^0$ ,  $y_n^0$  and  $z_n^0$  into Eq. (16.4') shows that  $x_n$  and  $y_n$  now consist of the terms cos {  $(2\pi \nu \pm eH/2mc)t$  } while  $z_n$  remains unchanged.

In this way we see that the frequency of the particle motion in the presence of a magnetic field has, besides the original value of  $\nu$ , two additional ones,  $\nu \pm eH/4\pi mc$ . Accordingly, the frequency of the light

emitted by these particles in the atom will undergo the same modification. This is the theoretical interpretation of the Zeeman effect. By comparing this theoretical value of the frequency shift in the Zeeman effect,

$$\Delta v = \pm \frac{eH}{4\pi mc} \tag{16.5'}$$

with the experimental one, we can determine the ratio e/m of the particles responsible for the light emission of the atom. The comparison of this ratio with that of the cathode ray particles has established that they are same and hence that these particles in the atom are actually electrons.

## (ii) THE THOMSON MODEL AND THE NAGAOKA MODEL OF THE ATOM

It has now become clear that electrons exist in the atom. On the other hand, we know nothing about the heavier constituent with positive electric charge. Since this constituent is very heavy, it has nothing to do with the emission or absorption of light by the atom; in other words, there is for the moment no experimental way of obtaining information about it. Therefore, we must proceed with some hypothesis concerning this constituent.

As hypothesis of this kind, two different ones were proposed in the same year, 1904. One was due to Thomson and the other to Nagaoka.

Thomson considered the positive-charge constituent of the atom to be distributed uniformly over the atom; thus the size of the atom can be said to be the size of this sphere of positive charge. The electrons are considered to be embedded in this sphere. According to this model of the atom, there is a stable state of equilibrium where the electrons are at rest in the atom. This guarantees the stability of the atom. When, due to some cause, the electrons are displaced from this stable state, they start oscillating around their equilibrium positions. The atom would then emit light. In order for the emitted light to have the frequency in the usually observed range of the atomic spectra, a size of about  $10^{-8}$  cm in radius is theoretically calculated for the sphere of positive charge. This size is just about the one required by the theory of gases for the size of the atom.

Contrary to this idea, Nagaoka considered the positive charge to be more or less concentrated at the center of the atom and that the electrons surround this central nucleus of positive charge like the rings of Saturn. He showed that, when the electrons are properly arranged in this manner, there exists a dynamically stable configuration. In this state, however, the electrons are not at rest and hence according to Maxwell's theory they would, even in a dynamically stable state, emit light and lose energy. This state is, therefore, not really a stable one. This seems to be the reason why Nagaoka's model was not generally accepted. This model also does not explain why the atom has a size of about  $10^{-8}$  cm in radius, but we will come back later to this point.

#### (iii) SCATTERING EXPERIMENTS WITH *a*-PARTICLES

A clear-cut experimental decision concerning the above two models of the atom was not given until Rutherford made an epoch-making study in 1911. This was primarily because the nature of the positivecharge constituent of an atom was difficult to observe because of its quite heavy mass.

In 1911, Rutherford obtained an important clue to the structure of the atom in the following way.

Consider high-velocity  $\alpha$  particles, emitted by radium, passing through a material substance. The  $\alpha$  particle has an electric charge 2eand has approximately the same mass as the He atom. When this particle passes through an atom, the direction of its motion will be deflected by the forces due to the electrons and to the positive charge in the atom. However, since the mass of electron is extremely small, it is hardly probable that the trajectory of the much heavier  $\alpha$  particle will be changed by interaction with the electron. The deflection, therefore, would mainly be due to the influence of the positive charge in the atom.

The description of this deflection can be obtained by theoretical calculations when the electric field due to the distribution of positive charge in the atom is known. Even without detailed calculation, we can see the general tendency that the Nagaoka model, in which the positive charge is concentrated at the center, will make large angle deflections more probable than the Thomson model in which the positive charge is spread uniformly over the atom. This is so because in the former model there would be a very strong electric field near the center which strongly deflects the  $\alpha$  particle passing by the center.

When  $\alpha$  particles traverse a material substance, they will be scattered. The accurate experimental investigation of this scattering process, therefore, will tell much about the electric field in the atom. A very detailed experiment of this kind was performed by Geiger and Marsden in 1909.

From the results of this experiment, it is seen that an  $\alpha$  particle is sometimes deflected by more than 90° while the majority suffer only small angle deflections. Thomson, meanwhile, had shown through a calculation based on his model that large angle deflections of the  $\alpha$ particles could not happen in his model of the atom. The large-angle deflection of an  $\alpha$  particle observed by Geiger and Marsden was then generally considered as due to successive occurences of many small angle deflections.

Rutherford however realized the importance of this fact and noticed that large-angle deflections occured much more frequently than would be given by the statistical theory for successive small-angle deflections. The large-angle deflection is then to be considered as due to a single act and if it is actually so one has to conclude that there exists a much stronger electric field in the atom than is considered by Thomson.

### (iv) RUTHERFORD'S FORMULA

Led by the above experimental fact, Rutherford carried out the calculation of the scattering process of  $\alpha$  particles assuming that the positive charge was concentrated as a nucleus at the center of an atom. This calculation by Rutherford is a typical example of the method of calculation which is applied to the scattering of atoms, molecules, and elementary particles, or even to other phenomena. We hence describe it in some detail.

We consider an  $\alpha$  particle, with charge + 2e, passing through in an atom. As was stated before, we can neglect the influence of the electrons without causing a noticeable error. Our problem is, then, simply that of two positive charges interacting through the electric force which is inversely proportional to the square of their relative distance. Furthermore, since the nucleus of the atom is considerably heavier than the  $\alpha$  particle, we can regard this nucleus without appreciable error as a center of force fixed in space. The problem is now the simple Kepler problem. The only difference from the problem of the planets is in that we are dealing here with a repulsive force, the orbits being hyperbolic; the  $\alpha$  particle comes close to the atom from infinity and after deflection moves out again to infinity.

Suppose an  $\alpha$  particle is passing by the positive charge Ze con-

centrated in the atomic nucleus as shown in Fig. 23. In the figure, S is the center of the atom. The particle then moves on an hyperbola with S as the outer focus. We take the x-axis through S and parallel to the

line, POP', along which the  $\alpha$  particle is approaching the atom from the left. The hyperbola then has this line, POP', as one of its asymptotes. If we denote  $\overline{P}$ the other asymptote by QOQ', this gives the direction at infinity after the scattering. The scattering angle is accordingly given by  $\angle OOP' = \theta$ .

The distance from the x-axis of the particle deflected by a nucleus.  $\alpha$  particle at infinity when it is approaching the atom, i.e., the distance be-



Fig. 23. The path of an  $\alpha$ -S is the nucleus, PAQ the path of the  $\alpha$ -particle.

tween the line POP' and the x-axis in Fig. 23, is denoted by p. This distance  $\phi$  is a measure of how close the  $\alpha$  particle comes to the atom and is an important quantity in this kind of calculation. Hence this distance p is given the name of "impact parameter". When p is very large, the  $\alpha$  particle will pass the atom at a great distance and accordingly suffer hardly any deflection. When, on the contrary, p is zero, the  $\alpha$  particle will make a head-on collision with the atom and suffer the maximum deflection which, from symmetry considerations, amounts to 180°. The scattering angle  $\theta$  is in general a function of p, the form of which we can determine in the following manner.

Let the velocities of the  $\alpha$  particle at infinity and at the point of closes approach to the atom, i.e. at A in Fig. 23, be denoted by v and u, respectively. Then conservation of angular momentum gives the relation

$$pv = \overline{SA} \cdot u$$
,

where SA stands for the length of the line SA. On the other hand, the conservation of energy is expressed by

 $\frac{1}{2}mv^2 = \frac{1}{2}mu^2 + \frac{2Ze^2}{SA}.$ 

Putting now

$$\frac{4Ze^2}{mv^2} = b \tag{16.6}$$

we eliminate u from the above two equations and obtain

$$p^2 = \overline{SA} (\overline{SA} - b) . \tag{16.7}$$

Introducing the angle  $\varphi$ , as shown in Fig. 23, by

$$\varphi = \angle \text{SOP}' = \frac{1}{2}\pi - \frac{1}{2}\theta \tag{16.8}$$

we get

we get

$$SO = p \operatorname{cosec} \varphi$$
. (16.9)

Let the normal to the x-axis from the other focus S' be S'N. Then from the known characteristic of a hyperbola,

$$\overline{SN} = \overline{SA} - \overline{S'A} = 2\overline{OA} ,$$
$$\overline{OA} = p \cot \varphi . \qquad (16.10)$$

Adding Eq. (16.10) to Eq. (16.9), we obtain

$$SA = p \cot \frac{1}{2} \varphi$$
.

Substitution of this relation into Eq. (16.7) and use of Eq. (16.8) give

$$\cot \frac{1}{2}\theta = \frac{2p}{b} \tag{16.11}$$

which is the relation between  $\theta$  and  $\phi$  that we have been seeking.

We can now obtain the deflexion angle  $\theta$  of the  $\alpha$ -particle when the impact parameter p is given. However, in actual experiments, we do

having this given p value, but instead we observe the situation in which an atom is placed in an uniform stream of  $\alpha$ -particles, i.e. in which there are  $\alpha$ -particles of various p-values incident on the atom. Each of these  $\alpha$ -particles is deflected through an angle defined by its impact parameter and we observe the scattering of Fig. 24. The scattering of an  $\alpha$ -beam.  $\alpha$ -rays as shown in Fig. 24. In this

not follow a particular  $\alpha$ -particle

experiment, what is measured is then the number of particles scattered into the angles between  $\theta$  and  $\theta + d\theta$  i.e., the angular distribution of the scattered particles.



сн. 3, § 16]

The angular distribution of scattering can be derived in the following way. Consider that n incident particles per unit time are flowing in through unit area of the plane normal to the stream. Then the particles with impact parameters between p and p + dp will be coming in at the rate of

$$\mathrm{d}n = n2\pi p \,\mathrm{d}p \tag{16.12}$$

per unit time, since these particles are flowing in through the area  $2\pi\rho d\rho$  of the above plane.

According to Eq. (16.11), these particles are deflected into the angles between  $\theta$  and  $\theta + d\theta$  to be determined, respectively, by

$$\cot \frac{1}{2}\theta = \frac{2p}{b}$$

and

$$\cot\left(\frac{\theta+\mathrm{d}\theta}{2}\right) = \frac{2(p\,+\mathrm{d}p)}{b}\,.\tag{16.13}$$

Eq. (16.13) can be transformed to

$$\frac{1}{2} \frac{1}{\sin^2 \frac{1}{2\theta}} d\theta = \frac{2}{b} dp .$$
 (16.13')

In other words, when the angles  $\theta$  and  $\theta + d\theta$  are specified, the particles scattered in this angular range have their impact parameters between  $\phi$  and  $\phi + d\phi$  satisfying, respectively, Eqs. (16.11) and (16.13').

From this point of view, the number of particles scattered into these angles can be obtained by eliminating p and dp from Eqs. (16.12), (16.11) and (16.13'). The result is

$$\mathrm{d}n = n \, \frac{\pi b^2}{4} \, \frac{\cot \, \frac{1}{2} \theta}{\sin^2 \, \frac{1}{2} \theta} \, \mathrm{d}\theta$$

or by using Eq. (16.6)

$$dn = n \cdot 4\pi \frac{Z^2 e^4}{m^2 v^4} \frac{\cot \frac{1}{2}\theta}{\sin^2 \frac{1}{2}\theta} d\theta .$$
 (16.14)

When the incident particles are coming at the rate of one particle per unit time per unit area of the plane normal to the stream, the number of particles scattered per unit time in the direction between  $\theta$  and  $\theta + d\theta$  is denoted by dQ and is called the scattering cross-section. Obviously, dQ is obtained by dividing Eq. (16.14) by n and

$$dQ = 4\pi \frac{Z^2 e^4}{m^2 v^4} \frac{\cot \frac{1}{2}\theta}{\sin^2 \frac{1}{2}\theta} d\theta .$$
 (16.15)

This quantity has the dimension of area and the name "cross-section" is given to it for the following reason.

The situation in which incident particles are deflected by collision with an atom is similar to the one we encounter in shooting a target for a prize. We can consider that a target is placed in front of the atom and normal to the incident particle stream. On this target we draw two concentric circles with radii p and p + dp as shown in Fig. 25(a). Then the particles which hit the area between these two concentric



Fig. 25. The target in Rutherford scattering. The concentric circles in Fig. (b) correspond, from the inner to the outer ones, to the scattering angles  $127^{\circ}$ ,  $90^{\circ}$ ,  $53^{\circ}$ ,  $28^{\circ}$ ,  $11.4^{\circ}$ ,  $5.7^{\circ}$ , respectively. The ratios of their radii are 0.25:0.5:1:2:5:10.

circles will be deflected by an angle between  $\theta$  and  $\theta + d\theta$ . The deflexion angle can be considered as the amount of prize. Naturally, the area nearer to the axis will correspond to the larger-prize angle. The only inadequacy in this analogy is in that we do not try to hit the

center of the target but instead we shoot at random at the rate of one particle per unit time per unit area. In this state of affairs, the number of particles which hit the area between any two concentric circles, and which accordingly are deflected into the corresponding angular range, is equal to the dimension of this area and hence the name "cross-section" results.

In more general cases, e.g., when a certain process occurs whenever the particle hits a certain area of the target, we can in a similar sense define the cross-section for the process. Furthermore, even in the case where such a concept of target area cannot be introduced, we can also define the cross-section by the rate of occurrence per unit time of the process for the incident flux of one particle per unit time per unit area. Fig. 25(b) shows the concentric circles corresponding to the deflexion angles,  $5.7^{\circ}$ ,  $11.4^{\circ}$ ,  $28^{\circ}$ ,  $53^{\circ}$ ,  $90^{\circ}$ ,  $127^{\circ}$  and  $180^{\circ}$ . The radii of the circles have been calculated from Eq. (16.11).

The formula given by Eq. (16.15) is generally called the "Rutherford scattering formula". The above consideration refers to the scattering of  $\alpha$ -particles by a single atom. In the actual case of  $\alpha$ -particles traversing through matter, they undergo such a scattering repeatedly by many atoms. However, as can be seen from Fig. 25(b), in order to be deflected by a large angle, say 90° or more, the particle has to hit a very small area near the centre and the probability for such a large angle deflexion to occur repeatedly many times is extremely small. Hence, such a large-angle scattering is to be considered actually as a single large-angle deflexion. On the other hand, small-angle deflexions can easily happen repeatedly in matter. With the help of the theory of probability, we can calculate the frequency of occurrence of such small angle-deflexions as well as the resulting scattering angle. We do not go into the details of this but instead will, in the following, confine ourselves to the results of Rutherford's considerations.

### (v) RUTHERFORD'S CONCLUSION

Rutherford showed, from the analysis of the experiment by Geiger and Marsden, that his theory is actually in agreement with the experiment. Therefore, the positive charge in the atom is not distributed uniformly over the entire atom, as in Thomson's model, but instead is concentrated in the form of a nucleus at the center. Moreover, he succeeded in determining the charge of this central nucleus and concluded that, when this charge is denoted by Ze, Z is, to within an error of 20%, equal to one half of the atomic mass of the target matter. Van den Broek later showed that this quantity Z is to be set equal to the atomic number itself.

It has now become clear, due to the above arguments of Rutherford, that at the center of the atom there is an atomic nucleus which weighs approximately the same as the whole atom, that the nucleus has an electric charge equal to the atomic number Z times e, and that there exist Z electrons around this nucleus neutralizing the atom as a whole.

The fact that we could thus obtain knowledge of the atomic structure on the base of a good experiment has been a decisive step in the development of quantum mechanics.

#### (vi) DIFFICULTIES IN RUTHERFORD'S ATOM MODEL

Now we have a heavy nucleus of positive charge at the center and light electrons of negative charge around this nucleus. Since the attractive force, inversely proportional to the square of the relative distance, is acting between the nucleus and the electrons, the latter must revolve around the former in order that they not collapse to the center. The situation then closely resembles the solar system.

But what would determine the size of the atom? The orbit of this micro-solar system can take any size depending on the initial conditions, and, accordingly, the size, the energy and the period of motion of the atom can take any value. What then will explain the fact that an atom of a given material has a unique size, unique spectrum and unique energy depending only on the material. This difficulty is seen also in Nagaoka's model. In the fundamental equations of the theory in Rutherford's model as well as in Nagaoka's, the electric charge eand mass m naturally appear as the constants which determine the structure of the atom. The mass of the nucleus does not come into play since we assume that it is very large and hence the nucleus may be assumed to be at rest. Then any conclusion of our theory must be expressed in terms of combinations of these two quantities, e and m. However, as can be seen from their dimensions, it is impossible to construct a quantity having the dimension of length by combining the above two quantities. This explains why these two models cannot allow a determination of the size of the atom. The removal of the assumption that the nucleus is at rest and hence the introduction of the nuclear mass M in the theory does not remedy the situation, for obvious dimensional reasons.

There exist still more difficulties. Even if by some unknown law, like Bode's law for the solar system, the sizes of the electronic orbits are determined, the electrons in these orbits cannot help but emit electromagnetic waves in accordance with Maxwell's theory. Maxwell's theory requires that a charged particle in accelerated motion necessarily emits electromagnetic waves which carry away energy per unit time at the rate of

$$S = \frac{2}{3} \frac{e^2}{c^3} |\dot{\mathbf{v}}|^2 \tag{16.16}$$

where e, v and c are the charge of the particle, its velocity, and the velocity of light, respectively. Due to the energy loss resulting from this process, the electron should eventually collapse into the nucleus. We have an analogous situation when sound is being emitted by a freely vibrating body in air; it eventually quiets down because of the loss of energy into the air in the form of sound waves.

We take a hydrogen atom which of all the atoms has the simplest structure and show how rapidly the energy degrades by the emission of electromagnetic waves. Since the atomic number Z of hydrogen is unity, its center nucleus has charge e and around this nucleus there is one electron of charge -e. Among the various possible motions of the electron in this micro-solar system, we take the simplest one of circular motion with radius r. The acceleration  $\dot{v}$  in this circular motion is to satisfy the equation of motion

$$m \mid \dot{\mathbf{v}} \mid = \frac{e^2}{r^2}.$$

Substituting this  $|\dot{\mathbf{v}}|$  into Eq. (16.16), we obtain

$$\frac{\mathrm{d}W}{\mathrm{d}t} = -\frac{2}{3} \frac{e^6}{m^2 c^3 r^4} \tag{16.17}$$

for the energy loss per unit time of the atom by the emission of electromagnetic waves. On the other hand, the energy of an electron in this circular motion, kinetic energy plus potential energy, is

$$W = -\frac{e^2}{2r}.$$
 (16.18)

From Eqs. (16.17) and (16.18), we obtain

$$\frac{\mathrm{d}r}{\mathrm{d}t} = -\frac{4}{3} \frac{e^4}{m^2 c^3 r^2} \,,$$

the integration of which gives the time necessary for r to shrink from its initial value of a to zero. The result is

$$\tau_a = \frac{1}{4} \left( \frac{amc^2}{e^2} \right)^2 \frac{a}{c} \,. \tag{16.19}$$

If we start with the value  $a = 0.53 \times 10^{-8}$  cm which is generally accepted for the radius of hydrogen atom, we have

$$\tau_a = 1.4 \times 10^{-12} \, \mathrm{sec}$$
.

Thus a hydrogen atom of usual size should shrink to nearly a point within this very short period. This is obviously ridiculous.

Thomson avoided this type of model, an atom with a center nucleus, and took instead, the expended sphere of positive charge so that he could get around this difficulty. In his theory, besides m and e, there is another quantity specifying the size of this sphere as the additional structure constant. Therefore, it is quite natural that a quantity with the dimension of length results from his theory. As a matter of fact, the size of atom is nothing but this third structure constant, i.e., the size of the sphere of positive charge. Furthermore, the second difficulty in Rutherford's model also disappears in this model, since there is an equilibrium state in which the electrons are at rest and hence emitting no electromagnetic waves.

The third difficulty in Rutherford's model (not quite independent of the second) concerns the spectrum of light emitted by the atom. Let us calculate the frequency  $\nu$  of the electronic motion in the circular orbit considered above. Denoting the angular velocity of this circular motion by  $\omega$ , we have

$$mr\omega^2 = \frac{e^2}{r^2}$$
,

expressing the balance of the centrifugal force  $mr\omega^2$  and the attractive electric force  $e^2/r^2$ . The frequency  $\nu$  then, satisfies
$$\frac{1}{\nu^2} = \left(\frac{2\pi}{\omega}\right)^2 = \frac{4\pi^2}{e^2} mr^3.$$
 (16.20)

Thus  $\nu$  is proportional to  $r^{-\frac{3}{2}}$ , a result corresponding to Kepler's third law for the solar system. Accordingly, this frequency increases as rshrinks from a to zero. But then the frequency of light emitted during this motion cannot be constant and we cannot understand at all why a hydrogen atom emits a line spectrum of light having a constant frequency.

Finally, even if we neglect this variation of frequency with the loss of energy, this model by Rutherford still has a feature inconsistent with experimental fact. For in this model, the electrons undergo a periodic motion and, hence, their coordinates are expressible generally by Fourier series. This means that the term  $\cos(2\pi \nu t)$  of the series is in general accompanied by the terms  $\cos(2\pi\tau vt)$  with  $\tau$  an integer. In other words, the existence of a fundamental frequency  $\nu$  in the spectrum of light emitted by this electronic motion implies in general the co-existence of higher harmonics with frequencies which are integral multiples of v. In our previous case of a circular orbit, the higher harmonics did not appear but this is an exceptional case. In fact, if we treat the elliptic orbits of the hydrogen atom, we get this type of spectrum. However, in the experimentally observed spectrum of hydrogen, we do not observe anything similar to the situation where the fundamental frequency appears together with its higher harmonics. In a still more complex case where the electron in the atom undergoes some multi-periodic motion with fundamental frequencies  $v_1, v_2, \ldots$ ,  $v_{k}$ , the coordinates of the electron are given as a multiple Fourier series of terms like  $\cos \{ 2\pi(\tau_1\nu_1 + \tau_2\nu_2 + \ldots + \tau_k\nu_k)t \}$  with  $\tau_1, \tau_2, \ldots,$  $\tau_k$ , integers or zero. Consequently, in this case, these fundamental frequencies are to be accompanied by higher harmonics with the frequencies  $(\tau_1\nu_1 + \tau_2\nu_2 + \ldots + \tau_k\nu_k)$ . The actually observed spectrum of the emitted light, however, does not show any regularity of this kind.

This last difficulty is inherent not only in Rutherford's model but in all the models in which the electron in periodic motion is considered to be responsible for the emission of light.

Now, then, what kind of regularity is experimentally observed in the spectrum of the hydrogen atom, or of still more complex atoms?

93

#### § 17. The Spectrum of Atoms

By examining with a spectrometer the light from Geisler tubes of various gases or from arcs with various substances as electrodes, we can observe the spectrum characteristic of the substance used. Of a spectrum thus observed, the part called the line spectrum is due to the emission by the atom of the substance. Since this line spectrum shows a particular pattern characteristic of the kind of atom under consideration, it is the most direct source of information about the structure of the atom. Therefore, it is extremely important to find the regularities appearing in the patterns of this line spectrum.

Among the various substances, the hydrogen atom shows a line spectrum with the simplest pattern. Phot. V gives the visible to ultraviolet part of the spectrum of hydrogen, and this apparently has some sort of regularity. The numerical expression for this regularity in the spectrum of hydrogen was successfully obtained by Balmer, then a high-school teacher in Switzerland, in 1885. He observed that the four lines in the visible part,  $H_{\alpha}$ ,  $H_{\beta}$ ,  $H_{\gamma}$  and  $H_{\delta}$ , have wave lengths given, respectively, by

$$\frac{9}{5}h$$
,  $\frac{16}{12}h$ ,  $\frac{25}{21}h$ , and  $\frac{36}{32}h$ ,

where h = 3645.6 Å, and furthermore that these wave lengths are given by the general formula

$$\lambda = \frac{n^2}{n^2 - 4} h, \qquad n = 3, 4, 5, 6.$$
 (17.1)

The constant h here, by the way, has nothing to do with Planck's h. It was later found that this Balmer formula is applicable to all the lines in the ultraviolet region as observed in Phot. V. However, Balmer did not succeed in finding adequate formulae for the spectra of atoms other than hydrogen.

Following this finding by Balmer, Rydberg of Sweden investigated the spectra of other atoms and found the formulae for these more complex spectra. He found that, despite the apparent complexities, the spectrum of any atom can be classified into several series and that the lines in each of these series are expressible by the formula

$$\frac{1}{\lambda} = \frac{1}{\lambda_{\infty}} - \frac{R}{(n+b)^2}$$
(17.2)

сн. 3, § 17]

with *n* a positive integer, and  $\lambda_{\infty}$  signifying the limiting wave length of the series under consideration, i.e., the wave length toward which the lines of the series converge with ever increasing density, like  $H_{\infty}$  in the spectrum of hydrogen shown in Phot. V. In Phot. VI is shown as an example the spectrum of Na. This spectrum is considerably more complex than the spectrum of hydrogen in Phot. V. But when it is classified into principal series, first and second subordinate series, as indicated in Phot. VI, the lines in each of these series obcy formulae of the type given by Eq. (17.2). Thanks to Rydberg it has been found that the use of the reciprocal of the wave length, instead of the wave length itself as used by Balmer, simplifies the formula and that the frequency,  $c/\lambda$ , probably has a more fundamental meaning than the wave length. Another important achievement of Rydberg was his discovery that the constant R in his formula does not depend on the kind of atom but is a universal constant. This R is called Rydberg's constant and has the value

$$R = 109677.691 \text{ cm}^{-1}$$
.

This quantity is related to Balmer's h by R = 4/h.

On the other hand, the constant b in Rydberg's formula takes different values for different substances, or for different series, but within a certain series it has an approximately constant value. These findings by Rydberg were made in 1890.

Rydberg pushed his experimental studies still further and found that the series-limits of various series had also a similar regularity. Thus the term  $1/\lambda_{\infty}$  in the above formula can be expressed as  $R/(m + a)^2$ where *m* is a positive integer. He thus found that the spectrum in general could be expressed by the formula

$$\frac{1}{\lambda} = \frac{R}{(m+a)^2} - \frac{R}{(n+b)^2},$$
(17.3)

which is now called Rydberg's formula. Expressed in this form, the Balmer formula becomes

$$\frac{1}{\lambda} = \frac{R}{2^2} - \frac{R}{n^2},$$

which is a special case, m = 2, of a more general formula

$$\frac{1}{\lambda} = \frac{R}{m^2} - \frac{R}{n^2} \,. \tag{17.4}$$

In fact, if we extend the observation of the hydrogen spectrum to a still wider range of wave lengths, there are, besides the Balmer series with m = 2, also the series with m = 1 which was found by Lyman in 1906 and named after him, the Paschen series of m = 3 which was discovered by Paschen in 1908, and the series of m = 4 which was later found by Brackett in 1922. In Fig. 26, a schematic drawing of



Fig. 26. The spectrum of hydrogen. There should exist near v/c = 0 many more complex series, but those left of the Brackett series (m = 4) are omitted to avoid confusion.

these series is presented. For the sake of simplicity, the series of m larger than 4 have been omitted from the figure; these however exist in the low frequency region of the spectrum.

When Rydberg discovered the formula of Eq. (17.3), he did not forget to stress at the same time a fact which was later found to be extremely important for the development of quantum mechanics. He noted that the frequencies of a spectrum were given as the difference of two terms each of which contained an integer. He called these terms the spectral terms. Since this is so, it is simpler, in classifying the spectral lines, to work with the terms rather than with the frequencies themselves, the latter being found by taking differences of two terms. Since Ritz in 1908 used this fact extensively in his analysis of spectral lines, this is generally called Ritz's Combination law.

Rydberg's formula for spectral lines is not only very simple and elegant, but is also firmly verified experimentally. However, as Ritz stressed, this formula cannot be understood at all in the realm of classical theories. For according to classical ideas, it is considered that the light from an atom is emitted by the atomic electrons which are in some sort of periodic motion and produce waves in the surroun-

### BOHR'S THEORY

ding electromagnetic field. According to this point of view, the spectrum of light in general has to contain, together with the fundamental modes, the higher harmonics the frequencies of which have the form of a sum of integral multiples of these fundamental frequencies. This will be explained in more detail in §28, but, as was stated at the end of the last section, the reason for this is that the dipole moment of an electron in periodic motion is in general expressible as a multiple Fourier series with respect to time. Each term of this series has a frequency of the form  $(\tau_1\nu_1 + \tau_2\nu_2 + \ldots + \tau_k\nu_k)$  and the light is emitted with this frequency.

# § 18. Bohr's Theory

While many important results were being derived by various people, an event of enormous importance to quantum mechanics occurred in 1913. In this year Bohr produced the theory of structure of atom, which will be our subject in the following sections.

#### (i) THE FUNDAMENTAL IDEA

Despite the fact that the Rutherford model, contrary to others, is based on firm experimental results, there are various difficulties associated with it as was explained previously. One of these is that it does not contain a structure constant (as does Thomson's model) which can be used for the explanation of the size and stability of an atom. We have given some dimensional considerations on this point.

This difficulty however arose because we assumed the validity of Newtonian mechanics for the motion of the electron and the validity of Maxwell's theory for the emission of electromagnetic waves. On the other hand, since Planck's discovery of h, these classical theories have been shown to be merely approximate in nature. Now then, the difficulty in Rutherford's model might well be due to the neglect of this h and not due to some defect of the model itself.

If Planck's constant h enters in the law governing the atomic world, the situation becomes quite different. As was explained previously, we cannot construct a constant with the dimension of length from any combination of the two constants e and m. This was a clearcut indication of the difficulty. But if the law contains h, it is possible for the results of the theory to contain some quantity involving this htogether with e and m. There exists one quantity with the dimension of length derivable from the combination of these three constants, namely

$$l = \frac{h^2}{me^2}.$$
 (18.1)

It would then become possible for us to explain the size of atom theoretically without arbitrarily introducing a certain object of predetermined size in the theory such as was done in Thomson's theory. In fact, using the known numerical values of e, m and h, the above lbecomes  $2.1 \times 10^{-7}$  cm, a value which is roughly of the order of atomic size. Then our expectation that Planck's constant enters is considerably strengthened, since, if a certain quantity of the dimension of length could be derived from such a theory, it would be the above constant l multiplied by some numerical factor not much different from unity such as  $\pi$ , 2, 3,  $\sqrt{2}$  etc. This numerical factor, by the way, will later be shown to be  $1/(2\pi)^2$ , see Eq. (18.13'). Now that the introduction of Planck's, or Einstein's quantum concept seems to solve all the difficulties in our present problem, would it not be natural to proceed in the following manner?

Planck considered that, in quantum theory, the energy, and hence the amplitude, of a dynamical system in sinusoidal oscillation, to be called hereafter simply a harmonic oscillator, is not allowed to take any values continuously but takes only a set of discrete values given by

$$E_n = h\nu n$$
,  $n = 0, 1, 2, \dots$  (18.2)

Should it not then be that the energy of an atom also is allowed, in quantum theory, to take only a set of discrete values? Assuming this to be so, we designate the energy values by  $W_n$ , with *n* being integer. The discreteness of the energy requires that the electron orbits in the atom have discrete sizes, the smallest among which will then correspond to the size of the stable atom.

Meanwhile, Einstein showed that the photoelectric effect can be well understood by considering that the electron, on absorbing the energy of incident light of frequency  $\nu$ , increases its energy by  $h\nu$ . Though in the photoelectric effect the energy is transfered from the incident light to an electron, would not the quantum theoretical situation be the same also in the reverse process where an atom emits light? Thus, when an atom originally in a state of energy  $W_n$  changes

#### BOHR'S THEORY

its state for some reason to that of energy  $W_m$ , where we assume  $W_n > W_m$ , can we expect that a light quantum of energy  $h\nu$  satisfying  $h\nu = W_n - W_m$  will be emitted? If so, the frequency of the emitted light must satisfy

$$\nu = \frac{W_n - W_m}{h}, \qquad (18.3)$$

which we call Bohr's frequency condition.

Such an idea is entirely different from that of Maxwell. For now the emission of light is not considered as due to something like the acceleration of electrons in the state  $W_n$  but instead as due to an abrupt change of state,  $W_n \rightarrow W_m$ . As long as an atom is in one of the allowed energy states, it does not emit light at all. It was for this reason that Bohr called such states stationary states. Furthermore, such an abrupt change of state as  $W_n \rightarrow W_m$  is entirely beyond comprehension within the framework of Newtonian mechanics. We call this abrupt change a quantum jump or transition. As will soon be shown, the frequency given by Eq. (18.3) has an entirely different value from that of the electron's orbital motion. Despite this, Bohr's idea is quite convincing because we have seen in the photoeffect that the interaction between a light wave and an electron occurs in a pattern entirely different from Maxwell's.

According to this new idea, the wave length of light emitted from an atom satisfies

$$\frac{1}{\lambda} = \frac{\nu}{c} = \frac{W_n}{hc} - \frac{W_m}{hc} \,. \tag{18.3'}$$

This result,  $1/\lambda$  being given by the difference of two terms, is in fact the Rydberg-Ritz combination law itself.

Eq. (18.3') gives a physical meaning to what was called the term when we talked about spectral lines; aside from an additive constant of unknown magnitude it is nothing but the atom's allowed energy, divided by hc.

Thus, Bohr introduced the following two hypotheses as the basis of his theory.

Hypothesis (I): An atom can not take continuous energy values as in the classical theory, but instead it is allowed to take only certain discrete energy values,  $W_1, W_2, \ldots$ , characteristic of each atom. In these allowed states, the atom does not emit light. The term "atom" is to be understood here to include molecule, Planck's oscillator, and proper oscillations of the electromagnetic field, etc. We call the above allowed states and the allowed energy values,  $W_n$ , the "stationary states" and the "energy levels" of the atom, respectively.

Hypothesis (II): Emission or absorption of light by an atom occurs when the atom makes a transition from one to another of its stationary states. When this happens, monochromatic light of the frequency given by Eq. (18.3) is emitted or absorbed depending on whether  $W_n > W_m$  or  $W_n < W_m$ .

The case of proper oscillations of the electromagnetic field is here excluded from the meaning of the term "atom".

#### (ii) THE ENERGY LEVELS OF HYDROGEN

It has already been stated that the generalized Balmer formula holds for the spectrum of hydrogen; i.e.,

$$\frac{1}{\lambda} = \frac{R}{m^2} - \frac{R}{n^2}$$

By comparing this with Eq. (18.3'), we find that the allowed energy values  $W_n$  are given by

$$W_n = -\frac{Rhc}{n^2} \,. \tag{18.4}$$

The negative sign for  $W_n$  is due to the fact that the electron in the hydrogen atom is bound, requiring then a certain amount of work in order to be separated from the atom.

In this way, the energy levels of the hydrogen atom have been determined. They correspond, in the hydrogenic atom case, to the levels of Planck's oscillator given by Eq. (18.2). Various spectral lines observed in experiments are emitted when the atom makes a transition from one to another of these energy levels. For instance, the lines belonging to the Balmer series are emitted in the transitions from the levels  $W_3, W_4, W_5, \ldots$ , to the level  $W_2$ , while the ones belonging to Lyman series are emitted in the transitions from those levels to the level  $W_1$ . In Fig. 27 are shown, in units of  $\nu/c$ , these energy levels of a hydrogen atom. The arrows between two levels indicate the transitions

100

сн. 3, § 18]

between them and their length gives the reciprocal of the wave length of the light emitted in the corresponding transition.

In the case of Planck's oscillator, that is, when a mass m is oscillating under the influence of an elastic force

with elasticity coefficient  $\kappa$ , the frequency of the oscillation is given by

$$v=rac{1}{2\pi}\sqrt{rac{\kappa}{m}}$$
 ,

then Eq. (18.2) becomes, by using these structure constants together with h,

$$E_n = \frac{h}{2\pi} \sqrt{\frac{\kappa}{m}} n . \qquad (18.5)$$

Would it now be possible to transcribe Eq. (18.4) for a hydrogen atom into a similar form by using the structure constants, m and e, of our dynamical system together with h? Namely, the problem is to express the Rydberg constant R as a combination of m, e and h. If this were accomplished, the formula would

become more than a mere empirical formula and this would be a great step forward.

To achieve this we proceed as follows. As n increases in Eq. (18.4),  $W_n$  increases and approaches zero. The increase of  $W_n$  in turn implies the expansion of the atomic orbit, provided that Newtonian mechanics is still valid for an electron in a stationary orbit. From our experience of phenomena involving larger dimensions, we expect that the behavior of the electron in orbits much larger than the size of atom can be described by the classical theory. Accordingly the frequency of the light emitted in transitions from  $W_n$ 's of very large n should, in this limit, coincide with the classical one.

Meanwhile, as has been repeatedly stated, the frequency of light emitted by the electron in a periodic motion is the same as, or an integral multiple of, the fundamental frequency of the electron's orbital motion. In Rutherford's model, a hydrogen atom is similar to a simple solar system and we can calculate the frequency of the orbital motion by





using Kepler's Third Law, Eq. (16.20). By eliminating r from Eq. (16.20) and Eq. (16.18), we obtain

$$\nu = \sqrt{\frac{2}{\pi^2 m e^4}} | \dot{W} |^{\frac{3}{2}}, \qquad (18.6)$$

where W is the energy in this orbital motion. Hence the classical frequency of light emitted by this motion is given by

$$v_{\tau} = \sqrt{\frac{2}{\pi^2 m e^4}} |W|^{\frac{3}{2}} \tau , \qquad (18.6')$$

where  $\tau = 1$  and 2, 3, 4, ..., this equation giving respectively the fundamental frequency and the higher harmonics.

On the other hand, in Bohr's theory, the frequency is given by

$$v_{n \to m} = -\frac{Rc}{n^2} + \frac{Rc}{m^2}$$
, (18.7)

where  $v_{n \to m}$  is the frequency of the light emitted in the transition  $n \to m$ .

Putting  $m = n - \tau$ , Eq. (18.7) can be approximated for n's very much larger than  $\tau$  as

$$\nu_{n\to n-\tau}=\frac{2Rc}{n^3}\,\tau\,,$$

or, by the use of Eq. (18.4),

$$v_{n \to n-\tau} = \frac{2}{\sqrt{Rch^3}} |W_n|^{\frac{3}{2}} \tau$$
,  $\tau = 1, 2, 3, \dots$  (18.8)

From this formula we see that, in the case of very large n, even in Bohr's theory the frequencies of the light emitted by an atom consists of a certain "fundamental frequency"

$$\frac{2}{\sqrt{Rch^3}} \|W_n\|^{\frac{3}{2}}$$

and its integral multiples just as in the classical theory. If we take the view-point that, for very large n, the result of Bohr's theory coincides

with the classical theory, we should equate the two fundamental frequencies, i.e.,

$$\frac{2}{\sqrt{Rch^3}} = \sqrt{\frac{2}{\pi^2 me^4}},$$

$$R = \frac{2\pi^2 me^4}{ch^3}.$$
(18.9)

from which results

The correctness of this expression can be verified by introducing in it the numerical values of m, c, c, and h. This gives for the right hand side of Eq. (18.9) the value  $1.09 \times 10^{5}$  cm<sup>-1</sup>, which is in good agreement with the experimentally determined value of R, 109 677.691 cm<sup>-1</sup>.

We have thus succeeded in expressing the Rydberg constant R in terms of the structure constants, m, e and h, together with the light velocity c.

Though this expression contains c, it does not mean that the light velocity c is also one of the structure constants in the Rutherford theory. As a matter of fact, introducing Eq. (18.9) into Eq. (18.4), we get

$$W_n = -\frac{2\pi^2 m e^4}{h^2} \frac{1}{n^2} , \qquad (18.10)$$

which does not contain the light velocity c. This formula, Eq. (18.10), corresponds to Eq. (18.5) for Planck's oscillator.

The idea that the results of quantum theory should coincide with those of the classical theory in the limit of very large n has enabled us to express R in terms of the structure constants; this very idea in fact may be generalized to become Bohr's correspondence principle which played the role of a guiding principle in the days of the old quantum mechanics, i.e. until quantum mechanics was finally completed as a theory. We shall come back to this important principle later but at present we make only the following remarks.

When we say that the results of quantum theory coincide with those of classical theory, it is only in the sense that they agree in numerical values. In fact, there is a fundamental difference in the ways that the two theories produce these results. In the classical theory, light is emitted by an electron during the process of its motion and higher harmonics are emitted simultaneously with the fundamental mode in

[CH. 3, § 18

the same process; in quantum theory light is emitted in transitions from one to another of the stationary states and either the fundamental mode or one of the higher harmonics is emitted in one transition, but not all of them simultaneously in one process. This concept of a transition has no counter-part in the classical theory.

Another remark is that, when n is not large, there is no relation whatsoever between the quantum theoretical frequency of the emitted light and that of the electron's orbital motion. The former is given, making use of Eq. (18.2) and Eq. (18.10), by

$$v_{n \to n-\tau} = \frac{2\pi^2 m e^4}{h^3} \left\{ \frac{1}{(n-\tau)^2} - \frac{1}{n^2} \right\}, \qquad (18.11)$$

and the latter, making use of Eq. (18.6') and Eq. (18.10), by

$$\nu_{\tau} = \frac{2\pi^2 m e^4}{h^3} \frac{2}{n^3} \tau \,. \tag{18.11'}$$

In Fig. 28 is given a comparison of the two sets of frequencies with n = 5, i.e. the spectrum  $\nu_{5 \to m}$  of light actually emitted from hydrogen atoms in the state of n = 5, and the spectrum  $\nu_r$ , expected if emission



Fig. 28. Comparison between quantum and classical frequencies.

occurred in accordance with Maxwell's theory by the electron motion in the same state, n = 5. The two spectra are obtained by putting n = 5 and  $\tau = 1$ , 2, 3 and 4 in Eq. (18.11) and Eq. (18.11').

Needless to say,  $\nu_{\tau}$  consists of the fundamental frequency

$$\frac{4\pi^2 m e^4}{h^3} \frac{1}{5^3}$$

and its integral multiples while, in the spectrum  $v_{5 \rightarrow 5-\tau}$ , lines with  $\tau = 1, 2, 3$ , and 4 belong to the Lyman, Balmer, Paschen, and Brackett series, respectively.

## BOHR'S THEORY

#### (iii) THE SIZE OF THE ATOM AND ITS MAGNETIC MOMENT

Bohr thus gave the clue for understanding the empirical law of atomic spectra on the basis for a new theory which is entirely different in nature from the old one. This new theory is related to the old one only in that they give the same result in the limit of large n.

Now we inquire as to the behavior of atomic electrons in a stationary state. We already know that it has an energy  $W_n$ . But, can we gain further theoretical insight into the details of the electron's behavior? If not we would not be able to derive theoretically the various characteristics of the atom, e.g. its size. It turns out, however, that the above two hypotheses are not sufficient for this purpose. We have to introduce, besides these two, an additional hypothesis concerning the behavior of electrons. Since we have at present no experimental clues as to the content of this hypothesis, we shall, according to Bohr, tentatively take it to be the following.

Hypothesis (III). The electron in a stationary state obeys the laws of classical mechanics.

With this hypothesis, we can calculate the size of an atom in the following manner.

For the sake of simplicity, we first assume that the electron is describing a circular orbit of the atomic micro-solar system. Let r be the radius of this orbit and let the electron energy in this orbital motion be given by Eq. (16.18). Equating this with the  $W_n$  as given by Eq. (18.10), we obtain

$$r = \frac{h^2}{4\pi^2 m e^2} n^2$$
,  $n = 1, 2, 3, \dots$  (18.12)

This result implies that not only the energy but also the size of the atom cannot assume any arbitrary value but instead can take only a set of discrete values. The state of the lowest energy, and consequently the most stable one, is that of n = 1 and its radius is denoted specifically by a, i.e.,

$$a = \frac{h^2}{4\pi^2 m e^2} \,. \tag{18.13}$$

The numerical value of a is  $0.528 \times 10^{-8}$  cm which in turn is also the value required by the gas kinetic theory for the size of the atom.

The quantity l which was introduced as the fundamental length at the beginning of this section by a dimensional consideration is related to this radius a by

$$a = \frac{i}{(2\pi)^2}.$$
 (18.13')

We call this the Bohr radius.

Since this state of n = 1 is the one with the lowest energy among the various possible states of the hydrogen atom, a hydrogen atom, when isolated, eventually ends up in this state and stays there indefinitely. In this sense, we call this state the ground state, while states with higher values of n are called excited states. The stability of the atom, which was one of the main difficulties of Rutherford's model, is now assured, since an atom, once it comes to the ground state, stays there forever. Its radius therefore also remains constant as given above.

It is further to be noted that the introduction of the third hypothesis together with the assumption of a circular orbit allows a simplification of the relationship given by Eq. (18.10). Namely, in terms of the angular momentum M of the electron's circular motion, its energy W is expressed as

$$W = -\frac{M^2}{2r^2m} \, .$$

Eliminating r and W from this equation by the aid of Eqs. (18.10) and (18.12), we obtain

$$M = \frac{h}{2\pi}n, \quad n = 1, 2, 3, \dots$$
 (18.14)

This relation should be particularly noted since it does not contain the structure constants m and e, this then suggesting its validity for central forces in general. It will be shown later that this is actually the case. Eq. (18.14) shows that under the assumption of ordinary mechanics for the electron's motion in a stationary state with a circular orbit, the angular momentum is an integral multiple of Planck's constant devided by  $2\pi$ . It can take, just like the energy, only certain discrete values.

In the ground state, a hydrogen atom then has an angular momentum equal to  $h/2\pi$ . (This statement, however, will later be revised somewhat

in § 42.) The circular motion of an electron with the electric charge -e and with this angular momentum produces a circular electric current and hence the atom has a magnetic moment  $\beta$  given by

$$\beta = \frac{eh}{4\pi mc}.$$
 (18.15)

The atom in a state other than its ground state has a magnetic moment equal to an integral multiple of this  $\beta$ , and hence the magnetic moment is also allowed to take only discrete values, i.e., integral multiples of this unit  $\beta$  which we call the Bohr magneton.

Hypothesis III thus leads to the discontinuous nature of various characteristics of the atom. It is a characteristic feature of the quantum theoretical approach that only certain discrete values are allowed for a physical quantity which from the classical view point could take on a continuous range of values. We describe this situation by saying that this quantity is quantized and call the integer n appearing in this quantization the quantum number.

In the above discussion, we assumed a circular orbit in addition to Hypothesis III, but a more general case will later be considered.

# § 19. The Quantum Condition

In the last section, we learned that the energy of a hydrogen atom in a stationary state is given by Eq. (18.10) while that of Planck's oscillator is given by Eq. (18.5). We now ask what law, if any, determines in general the quantum theoretically allowed states of a given dynamical system. Without such a general law we cannot treat atoms more complicated than the hydrogen atom, and cannot treat molecules. We have to know the recipe for selecting the quantum theoretically allowed states from the continuous manifold of states of a general dynamical system. Hereafter we call this recipe simply the quantum condition.

The general quantum condition was obtained historically in the following manner.

#### (i) EHRENFEST'S ADIABATIC HYPOTHESIS

The quantum laws are in their nature entirely different from the laws of classical mechanics or of electrodynamics and they cannot be under-

stood in terms of classical ideas. We learned in the last section that the first two hypotheses could not take us beyond the law of spectral lines and the third hypothesis was introduced tentatively in order to treat the dynamical system in a stationary state by classical mechanics. On the other hand, the transition, being a typically quantum effect, remained entirely inaccessible to treatment by classical methods.

In view of this feature of the theory, we have to rely on the classical theory in some parts of the problem while in the other parts we introduce quantum theoretical ideas. This attitude naturally is very unsatisfactory and is to be considered merely as a temporary one. It should at least be made clear at what point we have to make the switchover from the classical ideas to quantum theoretical ones. Our Hypothesis III was to meet this requirement. We now ask if it is possible to extend the domain of classical theory beyond the limit as specified by Hypothesis III. In partial response to this question, Ehrenfest introduced Hypothesis IV on the basis of considerations to be described in the following. This Hypothesis IV plays an important role in leading us to a quantum condition of more generality and is called Ehrenfest's adiabatic hypothesis.

Our derivation of Wien's displacement law in § 5 was based entirely on Maxwell's classical idea. Thus, for a hollow cavity which was compressed adiabatically and at the same time infinitely slowly, we used this classical idea in deriving the variation of amplitude and energy of each proper oscillation in the hollow cavity. Nevertheless, this displacement law holds also in Planck's formula which is entirely of a quantum nature. This is a very remarkable fact.

In the meantime, when we recall the quantum theoretical derivation of Planck's formula in § 8, we find that the requirement  $\varepsilon = a$  universal constant times  $\nu$  was necessary and sufficient for the derivation of a formula which satisfies the displacement law despite the existence of the energy quantum  $\varepsilon$ : with this substitution for  $\varepsilon$  the function  $P(\varepsilon/kT)$  of Eq. (8.7) became a function of  $\nu/T$  as required by the displacement law. Thus we can conclude that if, following Planck, we accept the existence of the elementary unit of energy for electromagnetic oscillations, the necessary and sufficient condition for the derived result to satisfy the displacement law is that this elementary unit is proportional to the frequency, say  $\varepsilon = h\nu$ , and that the energy of each proper oscillation is an integral multiple of this unit:  $E = h\nu \cdot n$ .

Now a deeper meaning for this noteworthy relationship and its

origin will emerge if we go back to § 5 where we proved Wien's law. There we used the procedure of compressing the hollow cavity adiabatically and at the same time infinitely slowly. During this compression, the energy as well as the frequency of each proper oscillation of the light varies continuously. Since, however, the ratio  $E_s/v_s$  is an adiabatic invariant, the quantum condition  $E_s = hv_s u_s$  ( $n_s = 0, 1, 2, 3, ...$ ) for each proper oscillation, once satisfied initially, will be satisfied at all times during and after the process of this compression. Thus, the quantum condition being always satisfied during the compression, our proof of Wien's law can naturally be expected to be valid also in quantum theory. This is the very origin of the above connection between the energy unit and the displacement law.

In this argument, however, a basic assumption is tacitly made. Namely that the results of classical mechanics on infinitely slow adiabatic variations, especially the result about adiabatic invariants, hold also in quantum theory and, in addition, if the system is in one of the quantum theoretically allowed states before the variation, it will stay there during as well as after the variation.

This assumption which led to Wien's law in the case of the black body radiation has been explicitly extended by Ehrenfest to include dynamical systems in general. The Hypothesis IV below is thus called Ehrenfest's adiabatic hypothesis.

Hypothesis (IV). When a variation is given externally to a dynamical system, and if this variation is performed infinitely slowly, then the behavior during the variation of this dynamical system can be described by conventional mechanics and, in this case, the system which is initially in a quantum theoretically allowed state will remain there during as well as after the variation.

This hypothesis can be applied to practical cases where an external electric or magnetic field is gradually and slowly applied to an atom. But when the external field is varied suddenly, the variation of the system does not follow the mechanical description and the atomic system will be observed to make a quantum transition just as when it was hit by incident light. In fact, the incidence of light can be regarded as an external electromagnetic field which varies very rapidly, and in such a case, the interaction between the atom and the light wave occurs in a completely different way.

The adiabatic hypothesis can be applied not only in practical cases as above, but also for a "Gedankenexperiment" as in the proof of Wien's law.

This latter way of using the hypothesis will be shown in the following to give a powerful method of determining the quantum condition for general dynamical systems.

#### (ii) THE GENERALIZED ADIABATIC INVARIANT

The procedure of giving an infinitely slow variation to a general dynamical system can be described mathematically as follows.

Let the coordinates and their conjugate momenta describing the dynamical system under consideration be  $q_1, q_2, \ldots, q_f$ , and  $p_1$ ,  $p_2, \ldots, p_f$ , respectively. The expression for the energy H of this system contains, besides these q's and p's, a certain number of parameters,  $a_1, a_2, \ldots$ , which characterize the system. In the case of a pendulum for example, the length of the string will appear in H. By changing the numerical values of these parameters infinitely slowly, we can describe mathematically the process of giving an infinitely slow variation to the dynamical system. The  $\alpha_n$ 's and  $\beta_n$ 's appearing in Appendix VI are nothing but these parameters  $a_1, a_2, \ldots$ .

Suppose that the initial values  $a_1, a_2, \ldots$  of these parameters are varied infinitely slowly to  $a'_1, a'_2, \ldots$ , and that a certain motion denoted by Mot  $(a_1, a_2, \ldots)$  before the variation is changed to Mot  $(a'_1, a'_2, \ldots)$ . We call the two motions adiabatically related when an infinitely slow variation can change one of them to the other, say Mot (a) to Mot (a').

When there exists some quantity which is kept at a constant value throughout this change from Mot (a) to Mot (a'), we call it a generalized adiabatic invariant. The value of such a quantity can be used to characterize the relationships among the motions.

### (iii) THE ADIABATIC INVARIANT FOR THE SYSTEM OF ONE DI-MENSIONAL PERIODIC MOTION

When an external deformation changes the frequency of a dynamical system which is performing a sinusoidal oscillation, for instance the pendulum in § 5 or Planck's oscillator in Appendix VI, we have proved that  $E/\nu$  is an adiabatic invariant.

Would such an invariant exist also for more general deformations, say from a sinusoidal motion to a more general periodic motion or from one general periodic motion to another? Such a deformation can result if, for instance, the law governing the force changes from Hooke's law to that of a more general nature. сн. 3, § 19]

#### THE QUANTUM CONDITION

Before we go into the details of this question, it would be advisable to transform the adiabatic invariant of a sinusoidal motion into a general form which does not involve the structure constants of the system, such as v.

The Hamiltonian function of a dynamical system making a sinusoidal motion is given in general by

$$H = \alpha p^2 + \beta q^2 , \qquad (19.1)$$

while the frequency of motion  $\nu$  is

$$\nu = \frac{1}{\pi} \sqrt{\alpha \beta} . \tag{19.2}$$

The Hamiltonian, of course, has the physical meaning of the energy of the system.

We now consider the motion where the value of the energy H is E. In this motion, both q and p are functions of time and they describe a curve, q = q(t) and p = p(t) on the q-p plane. We call this curve the trajectory of the motion. In the present case where H is given by Eq. (19.1), the trajectory is an ellipse given by

$$E = \alpha p^2 + \beta q^2 , \qquad (19.3)$$

as was explained in  $\S 8$ .

The area, J, inside this ellipse on the q-p plane is readily calculated to be

$$J = \frac{\pi E}{\sqrt{\alpha\beta}} . \tag{19.4}$$

We thus obtain the relationship

$$\frac{E}{\nu} = J. \tag{19.5}$$

Thus, in the case of a dynamical system performing a sinusoidal motion, the area enclosed by the trajectory on the q-p plane is just the adiabatic invariant. This conclusion is remarkable in the sense that it does not involve the structure constant of the system explicitly and that, accordingly, it might be expected to hold also for dynamical systems undergoing more general periodic motions. In the following we show that this is actually the case.

Lest we should be confused by too much generality and unable to observe the logical structure of the argument, we confine ourselves to the case of a point mass m under the influence of a potential Vwhich can be of any form as long as it gives a periodic nature to the motion. The Hamiltonian of the system is then given by

$$H = \frac{p^2}{2m} + V(q, a) , \qquad (19.6)$$

[сн. 3, § 19

where the dependence of V on some parameter a is shown explicitly. By giving an infinitely slow change of this parameter with time, we can effect an adiabatic change on the system. Naturally, the number of parameters in general need not be only one and the mass m, for instance, may also be considered as a variable parameter. For the sake of simplicity, however, we treat here the case of only one parameter a.

At first the parameter a is considered to be kept at a constant value. Then we consider a motion with a constant energy E. Solving the equation

$$\frac{p^2}{2m} + V(q, a) = E$$
(19.7)

for p, we get

$$p = \pm \sqrt{2m} \{ E - V(q, a) \}, \qquad (19.8)$$

which gives the trajectory of the motion. When the system is performing a periodic motion, the trajectory describes a closed curve since the



Fig. 29. Trajectory of a periodic motion.

identical motion is to recur in a certain period of time. In our case, the trajectory being given by Eq. (19.8), the closed curve is symmetric with respect to the q axis of the q-p plane and looks, for instance, like the one shown in Fig. 29. The intersections of the curve with the q axis,  $q_1$  and  $q_2$  are given by the condition p = 0, i.e.

$$E = V(q, a)$$
. (19.9)

These are the points where the energy of the system consists entirely of potential energy and where the point mass comes to rest and starts

# сн. 3, § 19]

moving in the opposite direction. A closed trajectory as in Fig. 29 results when E - V(q, a) is positive for  $q_1 \leq q \leq q_2$  and negative otherwise. Then p is real for  $q_1 \leq q \leq q_2$  and imaginary otherwise. The motion, therefore, is also restricted within the interval  $q_1 \leq q \leq q_2$  and becomes periodic.

We now vary the parameter a with time very slowly. The trajectory would then look like the one given in Fig. 30. In each period the

trajectory would describe a curve very similar in nature to the one of Fig. 29 but would change its shape gradually. We will show that the area enclosed by each cycle of the trajectory remains constant despite the change of shape.

When a is changing with time the trajectory will not be a closed one completely in each cycle. We therefore redefine J as follows. Namely, the quantity J at time  $\tau$  is



Fig. 30. Adiabatic deformation of a trajectory.

defined as the area enclosed by a hypothetical trajectory which is obtained by fixing the parameter a at its value at time  $\tau$  and which accordingly describes a completely closed curve. In fact, the quantity  $E/\nu$  in the case of sinusoidal motion followed this definition.

We are now ready to start to prove the constancy of J. The proof is somewhat mathematical but it will help the reader to refresh his mathematics.

The Hamiltonian function of Eq. (19.6) is generalized to the form

$$H = H(p, q, a)$$
, (19.10)

where the parameter a is given as a slowly varying function of time, a(t). The existence of such a time varying parameter in the Hamiltonian function does not alter the fact that q(t) and p(t) are determined by Hamilton's equations of motion, i.e.,

$$\dot{q} = \frac{\partial H}{\partial \dot{p}}, \qquad \dot{p} = -\frac{\partial H}{\partial q}.$$
 (19.11)

If we introduce the q and p determined by these equations as functions of time into H of Eq. (19.10), the value of H thus obtained gives the instantaneous energy of the system in this state of motion. The energy is not a constant but a function of time which we denote by E. Namely we have

$$H(p, q, a) = E$$
, (19.12)

which can be solved for p as

$$p = p(q, E, a)$$
 . (19.13)

We now consider the hypothetical dynamical system which would result if the parameter a were fixed at the value which it has at a certain instant  $\tau$ . It is further assumed that the values of p and q of this system coincides at the time  $\tau$  with  $p(\tau)$  and  $q(\tau)$  of the original dynamical system, respectively. We denote this motion of the hypothetical system by  $p_{\tau}(t)$  and  $q_{\tau}(t)$ . The energy of this system  $E_{\tau}$ , which is naturally a constant, is to coincide with E of the original system at time  $\tau$ , i.e.  $E_{\tau} = E(\tau)$ . In the meantime, since this hypothetical system is completely periodic, its trajectory is a closed curve and is obviously given by the equation

$$p_{\tau} = p(q_{\tau}, E_{\tau}, a_{\tau}) . \tag{19.13'}$$

The area, J, enclosed by this curve is a function of  $\tau$  and is calculated by

$$J = 2 \int_{q_{\tau}^{(1)}}^{q_{\tau}^{(2)}} p(q_{\tau}, E_{\tau}, a_{\tau}) \, \mathrm{d}q_{\tau} , \qquad (19.14)$$

where  $q_{\tau}^{(1)}$  and  $q_{\tau}^{(2)}$  are the two roots of the equation  $p(q_{\tau}, E_{\tau}, a_{\tau}) = 0$ . Let the total change of J due to the change of the parameter a be denoted by  $\Delta J$ . Then our aim is to prove

$$\Delta J = 0 \tag{19.15}$$

for a sufficiently slow variation of a as a function of time.

The proof goes as follows. First of all,

$$\frac{\mathrm{d}J}{\mathrm{d}\tau} = 2 \left\{ p(q_{\tau}^{(2)}, E_{\tau}, a_{\tau}) \frac{\mathrm{d}q_{\tau}^{(2)}}{\mathrm{d}\tau} - p(q_{\tau}^{(1)}, E_{\tau}, a_{\tau}) \frac{\mathrm{d}q_{\tau}^{(1)}}{\mathrm{d}\tau} \right\} \\ + 2 \int_{q_{\tau}^{(1)}}^{q_{\tau}^{(2)}} \left\{ \left( \frac{\partial p}{\partial E} \right)_{\tau} \frac{\mathrm{d}E_{\tau}}{\mathrm{d}\tau} + \left( \frac{\partial p}{\partial a} \right)_{\tau} \frac{\mathrm{d}a_{\tau}}{\mathrm{d}\tau} \right\} \mathrm{d}q_{\tau}$$

where the symbol  $()_r$  implies that the variables, q, E, a (and sometimes p, too) inside the bracket assume their respective values at

#### сн. 3, § 19]

time  $\tau$ , i.e.  $q_{\tau}$ ,  $E_{\tau}$ ,  $a_{\tau}$ ,  $(\text{and } p_{\tau})$ . The first term on the right hand side of the above equation vanishes since  $q_{\tau}^{(1)}$  and  $q_{\tau}^{(2)}$  are both solutions of  $p(q_{\tau}, E_{\tau}, a_{\tau}) = 0$ . In the meantime, p(q, E, a) has been obtained by solving H(p, q, a) = E and hence we have

$$\left(\frac{\partial p}{\partial E}\right)_{\tau} = \frac{1}{\left(\frac{\partial H}{\partial p}\right)_{\tau}}, \quad \left(\frac{\partial p}{\partial a}\right)_{\tau} = -\frac{\left(\frac{\partial H}{\partial a}\right)_{\tau}}{\left(\frac{\partial H}{\partial p}\right)_{\tau}}.$$

Furthermore  $\partial H/\partial p$  is equal to dq/dt because of the equation of motion, Eq. (19.11). Introducing these relations into the expression for  $dJ/d\tau$  given above, we finally obtain

$$\frac{\mathrm{d}J}{\mathrm{d}\tau} = \int \left\{ \frac{\mathrm{d}E_{\tau}}{\mathrm{d}\tau} - \left( \frac{\partial H}{\partial a} \right)_{\tau} \frac{\mathrm{d}a_{\tau}}{\mathrm{d}\tau} \right\} \mathrm{d}t , \qquad (19.16)$$

where the integration with respect to t is to be carried out over one period of the hypothetical system, e.g. from  $\tau - 1/2\nu_{\tau}$  to  $\tau + 1/2\nu_{\tau}$ ,  $1/\nu_{\tau}$  being this period. We notice here that  $(\partial H/\partial a)_{\tau}$ , as a function of t, is

$$\left(\frac{\partial H}{\partial a}\right)_{\tau} = \frac{\partial H(p_{\tau}(t), q_{\tau}(t), a_{\tau})}{\partial a_{\tau}}.$$

In the meantime, for the motion of the actual dynamical system, not the hypothetical one, we have the relationship

$$\frac{\mathrm{d}E}{\mathrm{d}t} - \frac{\partial H}{\partial a} \frac{\mathrm{d}a}{\mathrm{d}t} = 0$$

where

$$\frac{\partial H}{\partial a} = \frac{\partial H(p(t), q(t), a(t))}{\partial a(t)}.$$

Then we get

$$\frac{\mathrm{d}E_{\tau}}{\mathrm{d}\tau} = \frac{\partial H(p(\tau), q(\tau), a_{\tau})}{\partial a_{\tau}} \frac{\mathrm{d}a_{\tau}}{\mathrm{d}\tau},$$

and substituting this into (19.16) we obtain

$$\frac{\mathrm{d}J}{\mathrm{d}\tau} = \int_{\tau-\frac{1}{2v_{\tau}}}^{\tau+\frac{1}{2v_{\tau}}} \frac{\partial H(p(\tau), q(\tau), a_{\tau})}{\partial a_{\tau}} \frac{\mathrm{d}a_{\tau}}{\mathrm{d}\tau} \mathrm{d}t - \int_{\tau-\frac{1}{2v_{\tau}}}^{\tau+\frac{1}{2v_{\tau}}} \frac{\partial H(p_{\tau}(t), q_{\tau}(t), a_{\tau})}{\partial a_{\tau}} \frac{\mathrm{d}a_{\tau}}{\mathrm{d}\tau} \mathrm{d}t .$$
(19.16')

In order to express mathematically the fact that the change of a occurs infinitely slowly, we replace a(t) by a(t/T) and let T tend to infinity. Then a(t/T) varies from a(0) to a(1) as t changes from 0 to T. If the values of a(0) and a(1) are fixed, then the change is quick when T is small and is slow when T is large.

Now the total change of J caused by changing a from a(0) to a(1) is given by integrating (19.16'), i.e.,

$$\Delta J = \int_{0}^{T} \mathrm{d}\tau \int_{\tau-\frac{1}{2\nu_{\tau}}}^{\tau+\frac{1}{2\nu_{\tau}}} \frac{\partial H(p(\tau), q(\tau), a_{\tau/T})}{\partial a_{\tau/T}} \frac{\mathrm{d}a_{\tau/T}}{\mathrm{d}\tau} \mathrm{d}t$$

$$-\int_{0}^{T} \mathrm{d}\tau \int_{\tau-\frac{1}{2\nu_{\tau}}}^{\tau+\frac{1}{2\nu_{\tau}}} \frac{\partial H(p_{\tau}(t), q_{\tau}(t), a_{\tau/T})}{\partial a_{\tau/T}} \frac{\mathrm{d}a_{\tau/T}}{\mathrm{d}\tau} \mathrm{d}t$$
(19.17)

where  $a_{\tau}$  is replaced by  $a_{\tau/T}$ . Our task is to show that  $\Delta J$  vanishes for infinitely large T.

For proving this it is convenient to use new integration variables defined by

$$\tau/T = \sigma$$

$$t/T = s$$
.

Eq. (19.17) is expressed as

$$\Delta J = T \left\{ \int_{0}^{1} d\sigma \int_{\sigma-\frac{1}{2Tv_{\sigma}}}^{\sigma+\frac{1}{2Tv_{\sigma}}} \frac{\partial H(p(T\sigma), q(T\sigma), a_{\sigma})}{\partial a_{\sigma}} \frac{da_{\sigma}}{d\sigma} ds - \int_{0}^{1} d\sigma \int_{\sigma-\frac{1}{2Tv_{\sigma}}}^{\sigma+\frac{1}{2Tv_{\sigma}}} \frac{\partial H(p_{T\sigma}(Ts), q_{T\sigma}(Ts), a_{\sigma})}{\partial a_{\sigma}} \frac{da_{\sigma}}{d\sigma} ds \right\}.$$
(19.15')

The domain of the double integration is a narrow band-shaped region along the straight line  $s = \sigma$  in the  $\sigma$ -plane extending from  $\sigma = 0$  to  $\sigma = 1$ , the width being of the order  $1/Tr_{\sigma}$ .

We now change the order of integration in the first integral in  $\{\}$  of Eq. (19.15') and replace the integration variable  $\sigma$  by s and s by  $\sigma$ . Then we get

the first integral = 
$$\int_0^1 d\sigma \int_{\sigma}^{\sigma} + \frac{1}{2Tv_{\sigma}} \frac{\partial H(p(Ts), q(Ts), a_s)}{\partial a_s} \frac{da_s}{ds} ds$$
.

Strictly speaking the domain of integration indicated in this expression is not exact, but the error is negligibly small, of the order of  $(1/T\nu)^2$ . In carrying through the integration over s we may replace  $a_s$  and  $da_s/ds$  in the integrand by  $a_\sigma$  and  $da_\sigma/d\sigma$  respectively, since the integration interval is very small, of the order of  $1/T\nu$ . The correction to this approximation is a small quantity of the order of  $1/T\nu$ . We may further replace p(Ts) and q(Ts), the motion of the actual system, by  $p_{T\sigma}(Ts)$  and  $q_{T\sigma}(Ts)$  of the hypothetical system. The correction is here again of the order of  $1/T\nu$ . After integration over s these correction terms will give rise to a term of the order of  $(1/T\nu)^2$ , the order of the integrand multiplied by the order of the area of the integration domain.

Having made these approximations it is now clear that in  $\{ \}$  of Eq. (19.15') the main term of the first integral is cancelled by the second integral and the remaining correction term is of the order of  $(1/T\nu)^2$ . We get then

$$\Delta J = T \left\{ \left( \text{finite term} \right) \left( \frac{1}{T \nu} \right)^2 \right\}$$

which vanishes for infinite T.

The proof given above is sufficiently general and can be extended even to the cases where the Hamiltonian function is not of the form given by Eq. (19.6), provided that the trajectory is a closed curve. It should, however, be noticed that the trajectory does not always become a closed one when the system performs a periodic motion. For instance, take the case of a simple pendulum, the motion of which is described by the angle of inclination  $\varphi$  as the coordinate. The pendulum makes an oscillatory motion in the interval  $\varphi_1 < \varphi < \varphi_2$ as long as its amplitude is small. When the amplitude gets larger and when finally the interval  $\varphi_1$  to  $\varphi_2$  is not wholly contained between  $180^{\circ}$  and  $-180^{\circ}$ , the oscillatory motion of the pendulum goes over to a circular motion. Such a circular motion is indeed one form of periodic motion of a pendulum but, since  $\varphi$  is not restricted within a finite interval  $\varphi_1$  to  $\varphi_2$  but increases steadily with time, the trajectory of the motion does not become a closed one but a wavy curve with a period of  $2\pi$  as shown in Fig. 31. It can be shown that in such a case J should be defined as the area enclosed by the trajectory, the  $\varphi$ -axis, and the lines  $\varphi = 0$  and  $\varphi = 2\pi$ .

• At this point, it seems necessary to make the following remarks. We have some freedom in choosing the coordinates and momenta to describe the motion of a dynamical system. When Q and P can describe the motion of the system just as well as q and p, i.e., when



rotation.

Q-P and q-p are mutually related by a canonical transformation, it can be shown that the I defined in terms of Q and P has the same value as the one in terms of q and p. Na-Fig. 31. Trajectory in the case of mely, it can be proved that the area enclosed by a trajectory is invari-

ant with respect to canonical transformations of coordinates and momenta.

It can sometimes happen that, while in a certain set of coordinates and momenta the trajectory is a closed one, the transformation to certain new variables can change the trajectory into a wavy one. Even in such cases, the invariance of J is preserved if the J in such variables is defined as explained above in connection with Fig. 31. This quantity J is occasionally referred to as the action variable. The general definition of J can be written as

$$J = \oint p \, \mathrm{d}q \,, \tag{19.18}$$

where the small circle on the integral sign indicates that the integration should be performed over one cycle of the motion.

Before concluding this section, another remark is necessary. A careful study of our proof for  $\Delta J = 0$  reveals the fact that there is an exceptional case where this conclusion docs not hold however slow the variation of the parameter may be. It is the case where the dynamical system under consideration passes at one time during the variation of the parameter through a state of motion with infinite period. In such a situation,  $1/T\nu$  is no longer infinitesimal however large T may be and our arguments no longer hold. There is no adiabatic invariant for such a variation of the parameter.

(iv) THE QUANTUM CONDITION FOR A ONE-DIMENSIONAL PERIODIC SYSTEM

It has now become clear that the quantity J is an adiabatic invariant for a one-dimensional periodic system in general. We consider a dynamical system for which J assumes the values  $J', J'', \ldots$ corresponding to each of its quantum theoretically allowed states. When we deform this dynamical system adiabatically to a different

one, our Hypothesis IV guarantees that J assumes the same set of values for the corresponding quantum theoretically allowed states of the latter system. On the other hand, the reduction of an arbitrary one-dimensional dynamical system to Planck's oscillator is in principle always possible. It is clear in this case, from Eq. (19.5) together with E = hv, that J is equal to nh for a stationary state. We, hence, can conclude that, for one-dimensional periodic systems, the quantum theoretically allowed states are characterized by

$$J = nh$$
,  $n = integers$  (19.19)

where

$$J = \oint p \, \mathrm{d}q \,. \tag{19.20}$$

We thus have arrived at the general quantum condition. The quantum condition in this form was derived independently by Wilson, Sommerfeld, and Ishihara.

#### (v) GENERALIZATION TO MULTI-PERIODIC SYSTEMS

When the dynamical system under consideration has many degrees of freedom, the situation becomes quite complicated and we shall content ourselves by simply quoting the results.

Let the coordinates of the dynamical system and their conjugate momenta be  $q_1, q_2, \ldots, q_f$  and  $p_1, p_2, \ldots, p_f$ , respectively. Furthermore, let the solutions of the equations of motion be given by the following k-fold Fourier series with respect to time t for p and q.

$$p_{s} = \sum_{\tau_{1},...,\tau_{k}=-\infty}^{+\infty} P_{\tau_{1}...\tau_{k}}^{(s)} \exp \left\{ 2\pi i \left( \tau_{1}w_{1} + \tau_{2}w_{2} + ... + \tau_{k}w_{k} \right) \right\}$$

$$q_{s} = \sum_{\tau_{1},...,\tau_{k}=-\infty}^{+\infty} Q_{\tau_{1}...\tau_{k}}^{(s)} \exp \left\{ 2\pi i \left( \tau_{1}w_{1} + \tau_{2}w_{2} + ... + \tau_{k}w_{k} \right) \right\},$$
(19.21)

where

$$w_1 = v_1 t$$
,  $w_2 = v_2 t$ , ...,  $w_k = v_k t$  (19.22)

and the  $P^{(s)}$ 's as well as the  $Q^{(s)}$ 's together with the  $\nu_s$ 's do not depend on time. Then it can be proved that there exist k action variables which are defined by

$$J_s = \int_0^1 \left( \sum_{r=1}^f p_r \frac{\partial q_r}{\partial w_s} \right) \mathrm{d}w_s \,, \qquad s = 1, 2, \dots, k \,, \qquad (19.23)$$

where Eqs. (19.21) and (19.22) are to be used for p and q and the integration with respect to  $w_s$  is to be carried out with the other w's kept constant. It may appear that J thus calculated depends on w's other than  $w_s$ , and hence on time through them, but it can be shown that the quantity J thus obtained is in fact independent of time. Eq. (19.23) can also be written as a line integral.

$$J_s = \oint \sum_{r=1}^{f} p_r \,\mathrm{d}q_r \tag{19.24}$$

where the path of integration is to be taken along the closed loop  $0 \le w_s \le 1$ , with the other w's fixed, in the q-p space, or more specifically in the 2*f*-dimensional space of  $q_1, q_2, \ldots, q_f, p_1, p_2, \ldots, p_f$ . The *J*'s thus defined can be shown by the aid of advanced mathematics to be adiabatic invariant quantities.

Guided by the analogy to the one-dimensional case, we set

$$J_1 = n_1 h$$
,  $J_2 = n_2 h$ , ...,  $J_k = n_k h$  (19.25)

as the quantum conditions for our dynamical system.

Even with these elaborate refinements, however, the actual determination of the stationary states from these quantum conditions and the calculation of the energy values to obtain the spectral terms can be carried out only in very simple cases.

One of the simplest is the case where the Hamiltonian function of the dynamical system under consideration consists of a sum of terms each depending on only one of the coordinates and its conjugate momentum as follows:

$$H(p_1, p_2, \ldots, p_f; q_1, q_2, \ldots, q_f) = \sum_{s=1}^f H_s(p_s, q_s) .$$
(19.26)

In this case, we can consider each degree of freedom of the whole system as an independent dynamical system and consequently the calculation of J can be carried out for each degree of freedom separately and this is nothing but the one-dimensional system treated in the last section. Thus, each individual  $H_s$  of Eq. (19.26) is conserved. Hence, the equations

$$H_s(p_s, q_s) = E_s$$
,  $s = 1, 2, ..., f$ , (19.27)

give separate trajectories for each degree of freedom and by solving these equations for the  $p_s$ 's we obtain

$$p_s = p_s(q_s, E_s)$$
,  $s = 1, 2, ..., f$ . (19.28)

The calculation of

$$J_s = \oint p_s \, \mathrm{d}q_s \tag{19.29}$$

together with the quantum condition

$$J_s = n_s h$$
,  $n_s = \text{integers}$  (19.30)

for each degree of freedom, s = 1, 2, ..., f, will determine the stationary states of the dynamical system under consideration.

Examples of this simple dynamical system are the oscillations of a crystal or of the electromagnetic field where the use of normal coordinate decomposes the Hamiltonian of the system into terms, one for each degree of freedom, which are all of the form of Eq. (2.13). Consequently, the quantum condition is given by  $E_s/\nu_s = n_s h$  for each degree of freedom and, hence, the energy in each degree of freedom becomes  $E_s = h\nu_s \cdot n_s$ .

The next simplest case will be the one in which there are, besides the total energy, (f-1) constants of motion,  $F_1, F_2, \ldots, F_{f-1}$ , each being a function of the p's and q's. Denoting the values of these constants of motion by  $K_1, K_2 \ldots K_{f-1}$ :

$$H = E$$
,  $F_1 = K_1$ ,  $F_2 = K_2$ , ...,  $F_{f-1} = K_{f-1}$ , (19.31)

we obtain f relations which relate the p's and q's and f constants of motion, E and the K's. We now regard these relations as a system of f equations for the f unknown p's. We then assume that these simultaneous equations are solved and, in addition, that each of the p's obtained by solving these equations depends, besides on E and the K's only on its corresponding coordinate q (independent of the other q's); i.e., that the solutions are of the form

$$p_{1} = p_{1}(q_{1}, E, K_{1}, K_{2}, \dots, K_{f-1})$$

$$p_{2} = p_{2}(q_{2}, E, K_{1}, K_{2}, \dots, K_{f-1})$$

$$\vdots$$

$$p_{f} = p_{f}(q_{f}, E, K_{1}, K_{2}, \dots, K_{f-1}).$$
(19.32)

Then we can use the same definition of J as in the one-dimensional case:

$$J_{s} = \oint p_{s} \,\mathrm{d}q_{s} \,. \tag{19.33}$$

The cases treated in the following sub-section and in the next section,  $\S$  20, will serve as examples of the situation explained above.

#### (vi) EXAMPLE

As one of the simplest examples, we consider Planck's oscillator, the Hamiltonian of which is

$$H = \frac{p^2}{2m} + \frac{1}{2}\kappa q^2 . \qquad (19.34)$$

In the state of energy E, the equation

$$\frac{p^2}{2m} + \frac{1}{2}\kappa q^2 = E \tag{19.35}$$

holds, and, as has been stated previously, this determines an elliptic trajectory in the q-p plane. The area, enclosed by this trajectory, J, is given by

$$J = \oint p \, \mathrm{d}q = 2 \int_{q_1}^{q_2} \sqrt{2m(E - \frac{1}{2}\kappa q^2)} \, \mathrm{d}q \,, \qquad (19.36)$$

where  $q_1$  and  $q_2$  are the two solutions of

$$E - \frac{1}{2}\kappa q^2 = 0$$
, (19.37)

namely,

$$q_1 = -\sqrt{\frac{2E}{\kappa}}$$

(19.38)

and

$$q_2 = +\sqrt{\frac{2E}{\kappa}}$$

The result of the integration is

$$J = 2\pi \sqrt{\frac{m}{\kappa}} E \,. \tag{19.39}$$

The same result can also be obtained directly from the consideration

(area of ellipse) =  $\pi \times$  (major radius)  $\times$  (minor radius)

$$=\pi \times \sqrt{2mE} \times \sqrt{\frac{2}{\kappa}E}$$
.

In any case, the application of the quantum condition

$$J = nh$$
,  $n = 0, 1, 2, ...,$  (19.40)

gives

$$E = \frac{1}{2\pi} \sqrt{\frac{\kappa}{m}} hn$$
,  $n = 0, 1, 2, ...$  (19.41)

Since the frequency of this oscillator is given by

$$\nu = \frac{1}{2\pi} \sqrt{\frac{\kappa}{m}} , \qquad (19.42)$$

Eq. (19.41) can be reduced to the well known formula

$$E = h v \cdot n$$
,  $n = 0, 1, 2, ...$  (19.41')

We can also determine the motion of the oscillator in these stationary states. The system naturally makes a sinusoidal motion

$$q = a\cos\left(2\pi v t + \delta\right), \qquad (19.43)$$

and the momentum is

$$p = mq = -2\pi m v a \sin (2\pi v t + \delta)$$
. (19.44)

In order to determine the amplitude a in Eq. (19.43) and in Eq. (19.44), we substitute these two expressions for q and p in Eq. (19.35) to obtain

$$E = \frac{1}{2}\kappa a^{2} + \frac{1}{2}(2\pi)^{2} m\nu^{2} a^{2}. \qquad (19.45)$$

Comparing this with Eq. (19.41') we get

$$a^{2} = \frac{1}{2\pi^{2}m\nu}nh$$
,  $n = 0, 1, 2, \dots$  (19.46)

The oscillator, therefore, can assume only a set of discrete values for the amplitude of its motion.

If we use, instead of Eq. (19.43), the expression

$$q = X e^{2\pi i \nu t} + X^* e^{-2\pi i \nu t}, \qquad (19.43')$$

Eq. (19.46) is then to be replaced, since  $X = \frac{1}{2} a e^{i\delta}$ , by

$$|X|^2 = \frac{1}{2(2\pi)^2 m\nu} nh$$
,  $n = 0, 1, 2, ...$  (19.46')

Eq. (19.46) or Eq. (19.46') is the condition imposed by the quantum theory on the amplitude of the oscillatory motion.

## § 20. The Stationary States of the Hydrogen Atom

In the last section, we derived the quantum condition which the motion of dynamical system has to satisfy in the stationary states. The quantum condition states that the quantity J of the dynamical system under consideration is to be an integral multiple of Planck's constant. We now show in this section that the determination of the stationary states of the hydrogen atom by means of this quantum condition does actually reproduce Bohr's relation

$$W_n = -\frac{2\pi^2 m e^4}{h^2} \frac{1}{n^2}, \quad n = 1, 2, 3, \dots$$
 (20.1)

Furthermore, the motion of the electron in these stationary states will be analyzed and we shall find allowed motions other than the circular ones considered previously.

Before we enter these main topics, however, we first show that the relation (18.14), obtained previously for the angular momentum M of circular motion,

$$M = \frac{h}{2\pi} n$$
,  $n = 1, 2, 3, ...,$  (20.2)

can be derived from the general quantum condition.

In the case of circular motion, we take as the coordinate of the

electron its angle of rotation  $\varphi$  and for this choice of the coordinate the conjugate momentum  $p_{\varphi}$  is the angular momentum M itself. Since the



Fig. 32. Trajectory of a circular motion in the plane defined by the angle and the angular momentum. angular momentum is conserved in this case, the trajectory in the  $\varphi - p_{\varphi}$  plane becomes a straight line parallel to the  $\varphi$ -axis as shown in Fig. 32. The definition of J in the case of circular motion implies that it is the area enclosed by the trajectory, the  $\varphi$ -axis, and the lines  $\varphi = 0$  and  $\varphi = 2\pi$ ; i.e., the shaded area in Fig. 32. We, hence, have

$$J=\int_0^{2\pi}M\mathrm{d}arphi=2\pi M\,\mathrm{d}$$

The application of the quantum condition, J = nh, to this result readily gives Eq. (20.2).

As is clear from this verification, Eq. (20.2) holds not only for the hydrogen atom but also for the general dynamical system which performs a circular motion under the influence of an arbitrary central force and it will soon be shown that its domain of validity can be extended still further to periodic motion in general in the field of an arbitrary central force.

We are now ready to discuss the general motion of the hydrogen atom. We can consider the nucleus to be at rest at the center since it is much heavier than the electron. The hydrogen atom then is nothing but the dynamical system in which a point mass moves around under the influence of a central force inversely proportional to the square of its distance from the center, a system quite familiar to us since the time of Kepler. Confining ourselves to consideration of the motion in two dimensions, we use the polar coordinates r and  $\varphi$  to describe the position of the point mass, the electron. The Hamiltonian of this dynamical system is then

$$H = \frac{1}{2m} \left( p_r^2 + \frac{p_{\varphi}^2}{r^2} \right) - \frac{e^2}{r}.$$
 (20.3)

One of the equations of motion with this Hamiltonian,

$$\dot{\varphi} = \frac{p_{\varphi}}{mr^2}$$
 or  $p_{\varphi} = mr^2\dot{\varphi}$ 

shows that the momentum conjugate to  $\varphi$  has the physical meaning of the angular momentum. Furthermore, another one of the equations of motion,

$$p_{oldsymbol{arphi}}=0$$
 ,

indicates that this angular momentum is conserved, namely that

$$p_{\varphi} = \text{const.} = M . \tag{20.4}$$

Obviously, the total energy of this dynamical system is conserved to give

$$\frac{1}{2m}\left(p_r^2 + \frac{p_{\varphi}^2}{r^2}\right) - \frac{e^2}{r} = E .$$
 (20.5)

Since we have, besides the total energy, another constant of motion, the method of the next simplest case in the last section can be applied. In fact, the solutions  $p_{\varphi}$  and  $p_r$  of Eq. (20.4) and Eq. (20.5) are

$$p_{\varphi} = M \tag{20.6}$$

and

$$p_r = \pm \sqrt{2mE + \frac{2me^2}{r} - \frac{M^2}{r^2}}$$
(20.7)

which indeed have the form of Eq. (19.32) and give two separate trajectories. When, however, E > 0 the trajectory in the *r*-*p*<sub>r</sub> plane does



not close upon itself. This corresponds to a hyperbolic orbit. On the other hand, when E < 0, the trajectory given by Eq. (20.7) becomes a closed curve as shown in Fig. 33. From Eq. (20.6) and Eq. (20.7), we obtain

$$J_{\omega} = 2\pi M$$

Fig. 33.  $r-p_r$  trajectory for and Kepler motion.

$$J_r = 2 \int_{r_1}^{r_2} \sqrt{2mE + \frac{2me^2}{r} - \frac{M^2}{r^2}} \, \mathrm{d}r \,, \qquad (20.8)$$

where  $r_1$  and  $r_2$ ,  $r_1 < r_2$ , are the solutions of  $p_r = 0$ , i.e.

$$2mE + \frac{2me^2}{r} - \frac{M^2}{r^2} = 0. \qquad (20.9)$$

Сн. 3, § 20]

#### THE STATIONARY STATES

Physically,  $r_1$  and  $r_2$  are the distances of the electron from the center when it is at each end of the major axis of the elliptic orbit. In the terminology of the solar system,  $r_1$  and  $r_2$  are the lengths of the radius vector at perihelion and at aphelion, respectively. The integration in Eq. (20.8) is elementary and gives

$$J_{r} = -2\pi \left( |M| - \frac{mc^{2}}{\sqrt{-2mE}} \right).$$
 (20.10)

Introducing now the quantum condition that

$$J_{\varphi} = kh$$
,  $k = 0, 1, 2, \dots$  (20.11)

and

$$J_r = n'h$$
,  $n' = 0, 1, 2, ...,$ 

we obtain

$$|M| = \frac{h}{2\pi}k \tag{20.12}$$

and, by the use of this.

$$n'h = -kh + rac{2\pi me^2}{\sqrt{-2mE}}$$
 ,

which is solved for E to give

$$E = -\frac{2\pi^2 m e^4}{h^2} \frac{1}{(n'+k)^2} \,. \tag{20.13}$$

Since both n' and k are integers, so is

$$n = n' + k , \qquad (20.14)$$

and with this definition of n, Eq. (20.13) is identical to Eq. (20.1). The case n = 0 however has to be avoided since it does not give a finite energy value. Eq. (20.12) is identical to Eq. (20.2). The case k = 0 is also to be excluded since it gives linear orbits through the origin which will be hindered by the presence of the nucleus at the origin. Furthermore, since the angular momentum M can in principle assume negative values, we should allow also negative values of k. However, since these values give just an inversion of the direction of the electron rotation, we may restrict ourselves to positive values of k without loss of generality.

We have seen that the quantum theoretically allowed states can be labelled by the two quantum numbers, n' and k, but it will be more convenient to use n and k instead of n' and k to label the stationary states since different values of n' or k gives the same energy value as long as (n' + k) is the same. For a given value of n, then, k assumes values from 1 to n. We call n and k the principal and the subsidiary quantum number, respectively.

We now investigate the shapes of the orbits. Readily obtained is the major axis of the elliptic orbit since it is equal to  $(r_1 + r_2)$ , which in turn, from Eq. (20.9), is given by

$$r_1 + r_2 = -\frac{e^2}{E} = \frac{h^2}{2\pi^2 m e^2} n^2$$
. (20.15)

The minor axis is equal to

$$2\left\{\left(\frac{r_{1}+r_{2}}{2}\right)^{2}-\left(\frac{r_{1}-r_{2}}{2}\right)^{2}\right\}^{4}=2\sqrt{r_{1}r_{2}}$$

$$=2\sqrt{-\frac{M^{2}}{2mE}}=\frac{h^{2}}{2\pi^{2}me^{2}}nk.$$
(20.16)

Hence, we get

$$\frac{\text{major axis}}{\text{minor axis}} = \frac{n}{k}.$$
 (20.17)

This result shows that the case k = 0 gives straight line orbits while the case n = k corresponds to circular orbits, the radius of which is given by

$$r=\frac{h^2}{4\pi^2me^2}\,n^2\,.$$

This radius is nothing but Eq. (18.12), and in terms of the Bohr radius a of Eq. (18.13), we can write

$$major axis = 2an$$

$$(20.18)$$

$$minor axis = 2ank$$

and

minor axis 
$$= 2ank$$
.

Since the major axis depends only on n, the orbits with the same major axis correspond to the same energy. Several possible orbits are schematically shown in Fig. 34.
CII. 3, § 20]

The result that the energy depends only on the principal quantum number n and not on the subsidiary quantum number is characteristic of the dynamical system where the force is inversely proportional to the square of the relative distance. The method of calculation explained above can be extended without appreciable alteration to the case



Fig. 34. Electron orbits of the hydrogen atom.

where the force has a different dependence on the distance. It may be shown that, in such cases also, the motion is characterized by giving two quantum numbers n and k but the energy now depends on both n and k. In fact, atoms other than hydrogen, the alkali atoms for example, give spectra quite similar to that of hydrogen and the only difference is that the spectral terms of these atoms are now labelled by giving both n and k; see § 24. The reason for this similarity is that, though these atoms are very complex dynamical systems of many electrons, the agent of light emission or absorption is the outermost electron, and hence they can be considered as hydrogen-like in the first approximation. However, in these atoms the force acting on this outermost electron is not only due to the nuclear charge but is modified by the screening effect of the inner electrons. The force therefore does not have the inverse square dependence on the distance and hence results the dependence of the spectral terms on k as well as on n. This is the reason why the spectral terms of atoms other than hydrogen have additional complexities.

Thus the electron orbits, or the corresponding spectral terms, are determined in general by two quantum numbers n and k. We use, therefore, the notation  $n_k$  for classifying or arranging these terms.

In spectroscopy, the terms k = 1, 2, 3, ... are called s-, p-, d-, ... terms, respectively, and the corresponding orbits are similarly called s-, p-, d-, ... orbits. The terminology 1s-, 2p-, electron implies that the electron is in the s- orbit of n = 1, or the p- orbit of n = 2, respectively.

# § 21. The Azimuthal Quantization

In the case of the hydrogen atom, orbits with the same n, but with different k, have the same energy. When, in this fashion, a number of different motions are possible for one energy, or, in other words, when a number of different motions belong to one and the same energy level, we say this energy level is degenerate. For example, to the *n*-level of the hydrogen atom belong n different motions with  $k = 1, 2, \ldots, n$  and the level is considered as *n*-fold degenerate. A three-dimensional treatment or the consideration of electron spin, however, increases this multiplicity as will be explained in the next section and in the next chapter, respectively.

The treatment in the last section is more general than the one in  $\S$  18 in that it took into account elliptic orbits, but cannot be considered as completely general since it assumed from the beginning that the motion would occur in one plane. Though this assumption is verified by the resulting solution of the equation of motion, the problem should in principle be treated in three-dimensional space.



Fig. 35. Polar coordinates of the electron.

We use the polar coordinates of an electron r,  $\theta$ , and  $\varphi$  with the nucleus as the origin; see Fig. 35. The Hamiltonian of the hydrogen atom in this coordinate system is then given by

$$H = \frac{1}{2\mu} \left( p_r^2 + \frac{p_{\theta}^2}{r^2} + \frac{p_{\varphi}^2}{r^2 \sin^2 \theta} \right) - \frac{e^2}{r}, \quad (21.1)$$

where the mass of electron is denoted by  $\mu$  for later convenience. One of the immediate consequences of this Hamiltonian is

$$p_{\varphi} = \text{const} = M_z \,. \tag{21.2}$$

The physical meaning of  $p_{\varphi}$  is apparent from another one of the equations of motion

$$\dot{\varphi} = \frac{\dot{p}_{\varphi}}{\mu r^2 \sin^2 \theta},$$

which can be rewritten as

$$p_{\varphi} = \mu r^2 \sin^2 \theta \cdot \dot{\varphi} = \mu r \sin \theta \cdot r \sin \theta \cdot \dot{\varphi} \,.$$

Namely, noting that  $r \sin \theta$  is the projection of the radius vector on the equatorial plane and that  $r \sin \theta \cdot \phi$  is the rotational velocity of the end point of this projected radius vector, we find that  $p_{\phi}$  is just the z-component of the angular momentum.

The next result derived by inspection of the equation of motion is

$$p_{\theta}^2 + \frac{p_{\varphi}^2}{\sin^2\theta} = \text{const} = M^2 , \qquad (21.3)$$

which, with the aid of the equation of motion, is transformed to

$$p_{\theta}^{2} + \frac{p_{\varphi}^{2}}{\sin^{2} \theta} = \mu^{2} r^{4} \dot{\theta}^{2} + \mu^{2} r^{4} \sin^{2} \theta \cdot \dot{\varphi}^{2}$$
$$= \mu^{2} r^{2} \left\{ (r\dot{\theta})^{2} + (r \sin \theta \cdot \dot{\varphi})^{2} \right\}$$
$$= \mu^{2} [r \wedge \dot{r}]^{2},$$

and this is simply the square of the angular momentum. As the third constant of motion we obviously have the energy E given by

$$\frac{1}{2\mu} \left( p_r^2 + \frac{p_{\theta}^2}{r^2} + \frac{p_{\varphi}^2}{r^2 \sin^2 \theta} \right) - \frac{e^2}{r} = E .$$
 (21.4)

Solving Eqs. (21.2), (21.3) and (21.4) for  $p_{a}$ ,  $p_{b}$  and  $p_{r}$ , we have

$$p_{\varphi} = M_z,$$

$$p_{\theta} = \pm \sqrt{M^2 - \frac{M_z^2}{\sin^2\theta}},$$

$$p_r = \pm \sqrt{2\mu E - \frac{M^2}{r^2} + \frac{2\mu e^2}{r}},$$
(21.5)

all of which arc of the form of Eq. (19.32). We now proceed to the calculation of the J's. First, we immediately obtain

$$J_{\varphi} = 2\pi M_z \,. \tag{21.6}$$

$$J_{\theta} = 2 \int_{\theta_1}^{\theta_2} \sqrt{M^2 - \frac{M_z^2}{\sin^2\theta}} \,\mathrm{d}\theta \,,$$

 $J_{\theta}$  is given by

where  $\theta_1$  and  $\theta_2$  satisfy

$$M^2 \sin^2 \theta - M_z^2 = 0 , \qquad (21.7)$$

or  $\sin \theta = |M_z| / M$ , which has two roots between 0 and  $\pi$ . In order for these angles to be real, we require

$$-M \le M_z \le M . \tag{21.8}$$

Since  $M^2$  has the meaning of the square of the total angular momentum, we take here M to be its magnitude and assume  $M \ge 0$ .

The calculation of  $J_{\theta}$  is carried out by elementary techniques and the result is

$$J_{\theta} = 2\pi \left( M - |M_{z}| \right). \tag{21.9}$$

The calculation of  $J_r$  is exactly the same as in the last section and

$$J_{r} = -2\pi \left( M - \frac{\mu e^{2}}{\sqrt{-2\mu E}} \right).$$
 (21.10)

All the J's being explicitly given, we can impose the quantum conditions on them:

$$J_{\varphi} = mh , \qquad -(k' + |m|) \le m \le k' + |m|$$

$$J_{\theta} = k'h , \qquad k' = 0, 1, 2, \dots \qquad (21.11)$$

$$J_{r} = n'h , \qquad n' = 0, 1, 2, \dots$$

Negative values of m, as well as positive ones, are to be allowed since m has the physical meaning of the z-component of the angular momentum which can naturally be negative as well as positive. The integer m is called the magnetic quantum number. The inequalities  $-(k' + |m|) \le m \le k' + |m|$  should hold to obtain Eq. (21.8) from Eq. (21.9). The cases of k' or n' zero, however, require special considerations. From the quantum conditions above, it is easily obtained that

$$E = -\frac{2\pi\mu e^2}{h^2} \frac{1}{\left(n' + k' + |m|\right)^2},$$
 (21.12)

which, by the introduction of k and n defined by k' + |m| = k and n' + k = n, respectively, becomes

$$E = -\frac{2\pi^2 \mu e^4}{h^2} \frac{1}{n^2}, \qquad (21.13)$$

while

$$M = \frac{h}{2\pi}k \tag{21.14}$$

and

$$M_z = \frac{h}{2\pi} m , \qquad (21.15)$$

where

$$-k \le m \le k . \tag{21.16}$$

Eqs. (21.13) and (21.14) were already obtained in the last section and, by the same argument as was given then, the cases n = 0 and k = 0 are to be excluded. As is shown by Eq. (21.13) the energy depends, as before, only on n but on neither k nor m.

We have determined the energy, the magnitude of the angular momentum, and its z-component in a given stationary state. We can further determine the orbit of the electron in this state but, since we already know that the electron will move along an ellipse in a plane, the actual determination will not be given here. The following result of the three-dimensional approach, however, is to be noticed. That is that the angular momentum vector cannot have an arbitrary direction in space because of the restrictions given by Eqs. (21.14) and (21.15).

Denoting the angle between the z-axis and the angular momentum by  $\Theta$ , this angle has to satisfy the relation

$$\cos \Theta = m/k$$
,  $-k \le m \le k$  (21.16')

where k is a positive integer and m a positive or negative integer (including zero).

If we take the case k = 1, for instance, the allowed values of the angle  $\Theta$  are given by

$$\cos \Theta = 1, 0, \text{ and } -1 \tag{21.16'}$$

and, hence,

$$\Theta = 0^{\circ}, 90^{\circ}, \text{ and } 180^{\circ}$$
. (21.16")

This is referred to as the azimuthal quantization. To help visualize the situation, the allowed directions of the angular momentum vector in the cases of k = 1, 2, and 3 are shown in Fig. 36.



Fig. 36. Azimuthal quantization of the angular-momentum vector.

When however we recall the fact that space is isotropic and hence that the direction of the z-axis can be chosen arbitrarily we must ask what physical meaning at all can we associate with the azimuthal quantization? Suppose we take the case of the allowed motion of  $\Theta = 0^{\circ}$ with respect to the chosen z-axis. We can now choose a new axis, z', which makes an angle  $\alpha$  with the original z-axis. The above allowed motion then is inclined by the angle  $\alpha$  with respect to this new z'-axis and, since  $\alpha$  can have any value, this motion can be in any direction with respect to the z'-axis. How can we get around this paradoxical situation?

Despite this paradox, we cannot regard azimuthal quantization as entirely outside reality. For example, consider the motion of an electron in an uniform magnetic field H in the direction of the z-axis. The interaction energy between the magnetic field and the magnetic moment of the atom, Eq. (18.15), is then given by

$$arDelta E=eta H\cos arOmeg$$
 ,

which, for the  $\Theta$  values of 0°, 90° and 180° of k = 1 state, becomes

$$\Delta E = \frac{eh}{4\pi\mu c}H, 0 \text{ , and } -\frac{eh}{4\pi\mu c}H, \qquad (21.17)$$

respectively. If, on the contrary, the direction is not quantized, E can assume any value between —  $ehH/4\pi\mu c$  and +  $ehH/4\pi\mu c$  continuously.

Since the energy values are changed by the amounts given by Eq. (21.17), the frequency of the emitted light is by Bohr's condition also shifted by  $\pm (e/4\pi\mu c)H$ , and this is exactly the situation observed in the Zeeman effect (the reason why the frequency shift of  $\pm 2(e/4\pi\mu c)H$  is not observed will be explained later). If, therefore, it were not for the azimuthal quantization, we should have observed a continuous broadening, instead of the actually observed splitting into three parts, of the emission line for an atom in a magnetic field. We hence have to conclude that, at least in the presence of a magnetic field, the azimuthal quantization is actually realized and that the direction of the magnetic field is to be taken as the direction of the z-axis.

In view of this, we can make the following statement concerning the azimuthal quantization. We first apply a magnetic field in a certain direction and take the z-axis in the same direction. When the strength of the field is gradually reduced, an atom in this decreasing magnetic field will be subject to azimuthal quantization as given by Eq. (21.16'), in the limit of zero field strength. If the magnetic field makes an angle  $\alpha$  from the z-axis, the atom is now quantized according to Eq. (21.16') with respect to the direction of this magnetic field.

A question now arises. When the magnetic field is originally in the direction of the z-axis, what should happen if we decrease this field gradually to zero and then switch it on, with slowly increasing intensity, in some other direction say the z'-axis? A sudden and discontinuous reorientation of the atom has to occur at the instant of this switching on of the magnetic field in the z'-direction. Namely, the exceptional situation explained at the end of (III), § 19, arises and hence the method of Ehrenfest's adiabatic invariant can no longer be applied. In fact, when the magnetic field is reduced to zero, one of the frequencies of our triply-periodic system, the Larmor frequency  $eH/4\pi\mu c$  in Eq. (16.4), becomes zero and the corresponding period becomes infinitely large, thus hindering the application of Ehrenfest's argument. This discontinuous change, therefore, cannot be discussed in terms of classical mechanics and, though no light is emitted, this is also a kind of quantum transition.

# § 22. Experimental Verifications of the Stationary States

Bohr's idea was thus gradually put into a more mathematically refined form and proved extremely useful in explaining the regularities in the emission spectra, from infra-red to X-ray, of various atoms and molecules. But the explanation of spectral regularities alone does not by itself provide a direct proof of the fact that atoms or molecules exist only in certain allowed and discrete states.

Direct experimental verifications in some way are needed of the fact that the atom can indeed not be allowed to take any energy values continuously. As typical cases of such experiments, we now explain the following two.

#### (i) THE FRANCK-HERTZ EXPERIMENT

Immediately after the publication in 1913 of Bohr's original idea, Franck and Hertz attempted to verify the existence of the discrete stationary states in Bohr's theory. Their line of argument went as follows. Consider an electron with a certain energy moving in a gas. According to Bohr's idea, these gas atoms can assume only certain discrete values of energy,  $W_1, W_2, \ldots$ . At room temperatures, they presumably all have the lowest energy  $W_1$ . If the traversing electron has a sufficient amount of energy, it is in principle possible for the electron, in its collisions with a gas atom, to impart a part of its energy to the target atom, which accordingly is excited to one of its higher energy levels,  $W_2, W_3, \ldots$  In fact, the very reason why Geisler tubes radiate is that these excited atoms subsequently make transitions  $W_2 \rightarrow W_1$ ,  $W_3 \rightarrow W_1, \ldots$ , thereby emitting light. If, however, the electron energy is smaller than the difference  $W_2 - W_1$ , it is energetically impossible for the atom to be excited to the  $W_2$  state by the collisions of this electron; in this case all the collisions are clastic scatterings in which the electron loses no energy.

When we gradually increase the electron energy the electrons injected into the gas do not suffer appreciable energy losses by collisions with gas atoms, until the electron energy reaches  $W_2 - W_1$ . When, however, the electron energy surpasses this value  $W_2 - W_1$ , we should expect that some electrons will appear at the other end of the container of gas with a smaller amount of energy, the difference having been spent to excite a target atoms with which they collided.

Franck and Hertz showed by the following experiment that this phenomenon actually occurs. Their experimental set-up is schematically shown in Fig. 37. In the figure, D is a heated filament used as the electron source and the space R is filled by a gas. A is the electron collector connected to a galvanometer G both of which are carthed, while the electron source D is kept at an electric potential of -x. In front of A is placed a metal grid, n in the figure, and this is kept at



Fig. 37. The Franck-Hertz experiment.

a potential about +0.5 volt. Electrons from D are then accelerated by the electric field between D and A as they pass through the gas, till they finally arive at A. The galvanometer G indicates how many electrons have arrived. When we increase the potential difference -x, we receive more and more electrons at A, until x is sufficiently large and it becomes eventually possible for some electrons to have energies larger than  $W_2 - W_1$ .

The situation explained above then occurs and some of the electrons arriving at A will have much smaller energies due to inelastic collisions. These electrons with energy below the value corresponding to the 0.5 volt potential difference are prevented by the grid n from arriving at A. When we increase x above this critical value  $W_2 - W_1$ , therefore, we observe a decrease in the electric current through G. A further increase of x will increase the current again, since then the energy of the collided electrons becomes large enough to overcome the grid potential of +0.5 volt. A still further increase of x above  $2(W_2 - W_1)$  gives rise to another decrease of the current, since in this case electrons lose their energy in two successive collisions and are not

allowed to arrive at A. It is further anticipated that the same thing will happen whenever the energy of the accelerated electron goes through integral multiples of  $W_2 - W_1$ . In Fig. 38 are shown the experimental results for the G-current versus x obtained by Franck and Hertz for mercury vapor. The figure clearly indicates that the above phenomenon does actually occur. The x-values at which the current undergoes abrupt decreases are seen to be 4.9, 9, 14, ... volts, which according to the above consideration should correspond to electron energies



Fig. 38. The Franck-Hertz experiment.

 $W_2 \rightarrow W_1$ ,  $2(W_2 \rightarrow W_1)$ ,  $3(W_2 \rightarrow W_1)$ , ..., respectively. In the meantime, the spectroscopic study of mercury can give an independent estimate of  $W_2 - W_1$ . The comparison of these two estimates will give a confirmation of our point of view. The results from the spectroscopic study, in fact, give a value of 4.9 volt for  $W_2 - W_1$ , in good agreement with the one found by Franck and Hertz. Our interpretation of the phenomenon is thus justified and Bohr's hypothesis of stationary states is now shown to be really true.

### (ii) THE STERN-GERLACH EXPERIMENT

In 1922 Stern in collaboration with Gerlach carried out an extremely important experiment which showed that azimuthal quantization actually occurs for atoms in a magnetic field.

Their experiment can be summarized as follows. A piece of silver is heated in vacuum in a small electric furnace. Silver atoms evaporated from this heated piece of silver will form a so-called atomic beam of



Fig. 39. The Stern-Gerlach experiment.

silver when they emerge through a small opening in the furnace. The atomic beam thus obtained is led through a small slit into the magnetic field produced by an electromagnet with pole-pieces shaped as shown in Fig. 39. The small slit is indicated by B in Fig. 39. Beyond the magnetic field is placed a glass plate on which the atomic beam of silver will be plated producing an image of the slit B.

When there is no current in the magnet coil, i.e. no magnetic field along the path of the beam, the atomic beam will naturally travel along a straight line and hence the image of the slit B will be a single line as shown in (a) of Phot. VII. When however the current is switched on in the magnet coil to produce a strong magnetic field along the path of the atomic beam, the image of the slit is observed to split into two lines as shown in (b) of Phot. VII. On the basis of the argument which follows, we can conclude that the azimuthal quantization of atoms in a magnetic field is verified by this result.

The magnetic field acting on the atomic beam is purposely made extremely non-uniform by the peculiar shape of the pole pieces S and N in Fig. 39. In other words, the gradient  $\partial H/\partial z$  is very large, where z is taken along the direction connecting the two poles of the magnet. The silver atom probably has a magnetic moment and, when a magnetic moment passes through a non-uniform magnetic field, it experiences a deflecting force parallel to the field. This deflecting force depends on the direction of the magnetic moment and is given by

$$F=rac{\partial H}{\partial z}eta_{z}$$
 ,

where  $\beta_z$  is the z-component of the magnetic moment. If now the atom is azimuthally quantized,  $\beta_z$  can have only certain discrete values, since the magnetic moment of an atom will be in the direction of its angular momentum. The above deflecting force then can have only certain discrete magnitudes and, accordingly, the atomic beam will be deflected in certain definite directions, producing a number of distinct images of the slit on the glass plate. If, on the contrary, there is no such effect as azimuthal quantization, the slit image will just be blurred to some extent. Therefore, their experimental result that the image is split into two lines can be considered as a proof of azimuthal quantization.

Though we have stated that the Stern-Gerlach experiment clearly shows the reality of azimuthal quantization, the theoretical prediction is not in complete agreement with their result. For according to the theory, the azimuthal quantization will result in three directions,  $0^{\circ}$ ,  $90^{\circ}$  and  $180^{\circ}$ , in the case of k = 1 and in 5 directions in the case of k = 2 and so forth. If so, we should observe the slit image splitting into three or five or still more lines as long as the silver atom has a non-zero magnetic moment. The experimental fact that the slit image splits into two lines cannot then be accounted for theoretically.

This unsatisfactory situation is due to the incompleteness of our quantum theory and it can be remedied by amending the theory together with the introduction of the spin of the electron. This will be done in § 29. It is, however, the remarkable merit of the Stern-Gerlach experiment that the apparently queer phenomenon of azimuthal quantization is shown to occur actually in nature.

#### § 23. Bohr's Correspondence Principle

Bohr's theory is thus supported by the experimental results but it cannot be considered as a completed theory. It does explain how the discontinuous elements of nature discovered by Planck can be incorporated into the laws governing the atomic world, but the essential point of the problem is left still untouched; for in this theory we do not yet know about the fundamental quantum phenomenon of "transitions". Without knowledge of the mechanism by which these transitions occur, Bohr's theory can determine only the frequencies but neither the intensity nor the polarization of light emitted by the atom. The same is true for the absorption of light by the atom. The classical theory was complete in this sense, if we put aside the problem of agreement with experimental results. For Newtonian mechanics together with Maxwell's electrodynamics allows us to calculate not only the frequency but also the intensity and the polarization of light emitted by an electron in motion. The interaction being small between the electron and the electromagnetic field of the light, the calculation can simply be carried out in the following manner. The motion of the electron is primarily governed by Newtonian mechanics and the weak interaction with the field of the light can be neglected without seriously affecting the motion. The electron then describes a multiply-periodic motion and the dipole moment of the electron system accordingly undergoes a multiply-periodic variation.

The dipole moment, denoted by P, is then as a function of time to be expressed as a multiple Fourier series such as

$$P(t) = \frac{1}{2} \sum_{\tau_1, \tau_2, \dots, \tau_k = -\infty}^{+\infty} C_{\tau_1, \tau_2, \dots, \tau_k} \exp \left\{ 2\pi i (\tau_1 \nu_1 + \dots + \tau_k \nu_k) t \right\}, \quad (23.1)$$

where in order to make P real the coefficients  $C_{\tau_1, \tau_2, ..., \tau_k}$  have to satisfy the relationship

$$C_{\tau_1,\tau_2,...,\tau_k} = C^*_{-\tau_1,-\tau_2,...,-\tau_k},$$

At this stage Maxwell's electrodynamics is introduced and it tells us that electromagnetic waves are emitted from this oscillating dipole moment. This electromagnetic wave then consists of waves of various frequencies corresponding to each term of Eq. (23.1); i.e. of frequencies given by

$$\nu_{\tau_1 \tau_2 \dots \tau_k} = | \tau_1 \nu_1 + \tau_2 \nu_2 + \dots + \tau_k \nu_k |,$$
(23.2)

where  $\tau_1, \tau_2, \ldots$ , and  $\tau_k$  can be zero, or positive or negative integers. When one of the  $\tau$ 's is equal to unity while all the others are zero, the wave corresponding to the fundamental mode appears, and when the  $\tau$ 's have any other set of values waves of higher harmonics are obtained. Each of these waves, according to Maxwell's theory, can be considered as being emitted independently by the appropriate terms of the dipole moment of (23.1). The wave with frequency  $\nu_{\tau_1, \tau_2,\ldots,\tau_k}$  of Eq. (23.2), for example, emitted by the dipole moment of Eq. (23.1) is equal to the wave emitted by the dipole moment of the corresponding harmonic oscillator

$$P_{\tau_{1},\tau_{2},...,\tau_{k}}(t) = \frac{1}{2} \left[ C_{\tau_{1},\tau_{2},...,\tau_{k}} \exp \left\{ 2\pi i \left( \tau_{1}\nu_{1} + \ldots + \tau_{k}\nu_{k} \right) t \right\} + C_{-\tau_{1},-\tau_{2},...,-\tau_{k}} \exp \left\{ -2\pi i \left( \tau_{1}\nu_{1} + \ldots + \tau_{k}\nu_{k} \right) t \right\}.$$
 (23.3)

## CH. 3, § 23] BOHR'S CORRESPONDENCE PRINCIPLE

Therefore, when one wants to obtain the intensity of this wave, one simply substitutes Eq. (23.3) into

$$\frac{\mathrm{d}E}{\mathrm{d}\ell} = \frac{2}{3c^3} \langle \dot{P}^2 \rangle , \qquad (23.4)$$

which gives the energy radiated per unit time in the form of waves. In Eq. (23.4), the symbol  $\langle \rangle$  means that the time average over one period of the enclosed quantity is to be taken. We hence obtain for the intensity of the wave with  $\tau_1, \tau_2, \ldots, \tau_k$  the expression

$$\frac{\mathrm{d}E_{\tau_1\tau_2...\tau_k}}{\mathrm{d}t} = \frac{(2\pi\nu_{\tau_1\tau_2...\tau_k})^4}{3c^2} |C_{\tau_1\tau_2...\tau_k}|^2$$
(23.5)

with  $v_{\tau_1 \tau_2 \dots \tau_k}$  given by Eq. (23.2).

The polarization of the emitted light depends on the polarization of the oscillating vector  $P_{\tau_1 \ \tau_2 \dots \tau_k}$ . The oscillation of  $P_{\tau_1 \ \tau_2 \dots \tau_k}$  in general is elliptic, that is, the successive positions of the end point of the vector  $P_{\tau_1 \ \tau_2 \dots \tau_k}$  describe an ellipse. The shape of this ellipse is determined by  $C_{\tau_1 \ \tau_2 \dots \tau_k}$ . More specifically, denoting the *x*-, *y*-, and *z*components of  $C_{\tau}$  by  $C_{\tau x}$ ,  $C_{\tau y}$  and  $C_{\tau z}$ , respectively, the shape is determined by the ratio  $C_{\tau x}$ :  $C_{\tau y}$ :  $C_{\tau z}$ . For instance,

$$C_{\rm rx} = C_{\rm ry} = 0$$
,  $C_{\rm rz} \neq 0$  (23.7)

gives an ellipse which actually is squeezed into a line in the direction of the z-axis as shown in Fig. 40(a). The vector  $P_{\tau_1...\tau_k}$  is then said to be linearly polarized in the z-direction. The case



is characterized by a clockwise circular motion of the vector end-point in the x-y plane as in Fig. 40(c), while the case

$$C_{\tau y} = -iC_{\tau x}$$
, and  $C_{\tau z} = 0$  for positive  $v_t$  (23.9)

141

corresponds to an anti-clockwise circular motion as in Fig. 40(b). In the case of Eq. (23.8), or of Eq. (23.9), the vector  $P_{\tau_1 \ \tau_2 \dots \tau_k}$  is said to have a left-handed, or right-handed, circular polarization, respectively.

Maxwell's theory, therefore, can determine completely the nature of light emitted by a dipole when its time variation is known.

In this fashion the old theory can in principle determine the frequency, the intensity and the polarization of light emitted by the atom and hence it can be considered as a complete theory of atomic spectra.

According to Bohr's concept, however, the atom emits light by a mechanism entirely different from the classical one. Since the light is emitted from the atom not by the orbital motion of an electron but by the transition, the frequency of emitted light has nothing to do with the frequency of the orbital motion as has been clearly shown in Fig. 28. The same would probably be true with the intensity or with the polarization.

In this sad state of affairs, a slight hope is given by the fact explained previously in (ii), § 18. Namely, although the classical theory failed in explaining correctly the behavior of an atomic electron, it describes with sufficient accuracy the emission of light by an electron outside the atomic region. Thus we might expect that the nature of the light emitted by an atom, which is eventually to be correctly described by the quantum theory, could, in the limit of very large n, be described asymptotically by the classical theory. Starting from this limiting case one can to some extent proceed to make conjectures concerning the light emission with not so large quantum numbers. This concept not only served as a temporary means of investigation before the discovery of the complete quantum theory, but it also gave important clues for anticipating the form of the complete theory. This concept is called Bohr's correspondence principle and was in fact the guiding principle which led Heisenberg to find the correct form of quantum mechanics. We will spend several of the following sections in explaining this principle.

# (i) THE MATHEMATICAL FOUNDATIONS OF THE CORRESPONDENCE PRINCIPLE

When we treated the case of the hydrogen atom in (ii) § 18, the above requirement, that the nature of light calculated by the quantum theory should, in the limit of large n, agree with the one calculated by the classical theory, was actually satisfied by taking the Rydberg

142

#### CH. 3, § 23] BOHR'S CORRESPONDENCE PRINCIPLE

constant as given by Eq. (18.9). Can this be achieved, in some way or other, for more general dynamical systems than the hydrogen atom? If this is not generally the case, i.e. if the above requirement which is the basis of the correspondence principle is not always satisfied, we have to conclude either that the theory we have developed so far is erroneous or that the correspondence principle itself is not generally founded. It is however possible, as will be shown in the following, to prove mathematically that this requirement can be satisfied for general dynamical systems. In order to avoid excessive mathematical complexities, we restrict ourselves to the simplest case of the one-dimensional periodic system as was treated in (iii), § 19. The essential feature of the problem can nevertheless be observed in this simplified case.

Denoting the energy of this dynamical system in its *n*-th stationary state by  $W_n$ , the frequencies of light emitted by the system in this stationary state are given by Bohr's frequency condition,

$$\nu_{n \to n-\tau} = \frac{W_n - W_{n-\tau}}{h}.$$
 (23.10)

The  $W_n$ 's, in turn, are determined by imposing the quantum condition

$$J = nh$$
,  $n =$ integers (23.11)

on the allowed motions of the system. More specifically, when the energy W is obtained as a function of J, the substitution of nh for J in this expression gives  $W_n$ , i.e.

$$W_n = W(nh) . \tag{23.12}$$

We consider a system having a Hamiltonian function of the form

$$H(p,q) = \frac{p^2}{2m} + V(q) . \qquad (23.13)$$

We note that, when the system is a periodic one, the coordinate as well as the momentum of the system can be expressed as Fourier series with respect to time,

$$\begin{pmatrix} p \\ q \end{pmatrix} = \sum_{\tau=-\infty}^{+\infty} \begin{pmatrix} P_{\tau} \\ Q_{\tau} \end{pmatrix} e^{2\pi i \tau v t}.$$
 (23.14)

Our problem is to obtain the energy W of the system as a function of J, but we first calculate the inverse function J(W).

Since the Hamiltonian function has the physical meaning of the energy, p and q have to satisfy, in the motion with energy W,

$$H(p, q) = W$$
. (23.15)

Eq. (23.15) describes a curve on the q-p plane which is the trajectory of the motion with energy W. The system being a periodic one, the trajectory is a closed curve. The area enclosed by this closed curve is equal to J.

In this fashion, we arrive, just as we did in (iii), § 19, at

$$J = 2 \int_{q_1}^{q_2} p(W, q) \, \mathrm{d}q \tag{23.16}$$

where p(W, q) is the momentum as a function of W and q obtained as the solutions of Eq. (23.15), while  $q_1$  and  $q_2$  are the solutions of p(W, q) = 0. The integration of Eq. (23.16) gives J as a function of W.

J(W) being thus obtained, we calculate, for later use, dJ(W)/dW from Eq. (23.16) as

$$\frac{\mathrm{d}J(W)}{\mathrm{d}W} = 2\left\{ p(W, q_2) \frac{\mathrm{d}q_2}{\mathrm{d}W} - p(W, q_1) \frac{\mathrm{d}q_1}{\mathrm{d}W} \right\} + 2 \int_{q_1}^{q_2} \frac{\partial p(W, q)}{\partial W} \mathrm{d}q.$$

The first term of this expression vanishes according to the definition of  $q_1$  and  $q_2$ , while the second term can be transformed to a simpler form. By the use of the relation,

$$\frac{\partial p(W, q)}{\partial W} = \frac{1}{\frac{\partial H(p, q)}{\partial p}}$$

which is a consequence of the fact that p(W, q) is the solution of H(p, q) = W, together with the equation of motion

$$\frac{\mathrm{d}q}{\mathrm{d}t} = \frac{\partial H}{\partial p} \,,$$

we finally arrive at

$$\frac{\mathrm{d}J(W)}{\mathrm{d}W} = 2 \int_{q_1}^{q_2} \frac{1}{\mathrm{d}q} \,\mathrm{d}q = \int_0^T \mathrm{d}t = T \,.$$

In this result, T is the time it takes for the system to make one revolution of the trajectory; that is, it is the period of the motion. But Tis the reciprocal of the frequency  $\nu$  appearing in the Fourier series of Eq. (23.14). Hence, we obtain the quite important relation

$$\mathbf{v} = \frac{1}{\frac{\mathrm{d}J(W)}{\mathrm{d}W}} = \frac{\mathrm{d}W(J)}{\mathrm{d}J}.$$
(23.18)

The above derivation of this relation is not very complicated mathematically but, in order to help our insight, the same problem is treated in Appendix VIII by a method which is more physical and more easily visualized.

Returning to Eq. (23.10), we can approximate the numerator in the limit of very large n and of negligibly small  $\tau$  compared with n, by

$$W_{n} - W_{n-\tau} = W(nh) - W((n-\tau)h)$$
$$= \tau h \left(\frac{\mathrm{d}W}{\mathrm{d}J}\right)_{J=nh},$$

and accordingly

$$\nu_{n \to n-\tau} = \tau \left(\frac{\mathrm{d}W}{\mathrm{d}J}\right)_{J=nh}$$
(23.19)  
=  $\tau \nu$ ,

where v is the fundamental frequency, in the sense of classical theory, in the state of motion of J = nh. We, thus, have arrived at the conclusion that, in the limit of large n, the frequencies  $v_{n \to n-1}$ ,  $v_{n \to n-2}$ ,  $v_{n \to n-3}$ ,... of the light emitted in the transitions,  $n \to n - 1$ ,  $n \to n - 2$ ,  $n \to n - 3 \dots$ , respectively, agree with the fundamental frequency v, the second harmonic 2v, the third harmonic 3v,..., of the light emitted by the system according to the classical theory in this state of motion. In other words, the quantum theoretical frequencies belonging to the transitions,  $n \to n - 1$ ,  $n \to n - 2$ ,  $n \to n - 3$ ,... can be said to agree, respectively, with the frequencies of the terms,  $\tau = 1$ ,  $\tau = 2$ ,  $\tau = 3$ ,... of the Fourier series, Eq. (23.14). We have thus succeeded in generalizing the result previously proved only for the hydrogen atom in (iii) of § 19 and now have a result of general validity.

The conclusion obtained above for a simple dynamical system of one dimension can be generalized also for the case of multiply-periodic systems. Namely, in the case of a multiply-periodic system of f degrees of freedom, it was previously stated that the coordinates  $q_s$  and the momenta  $p_s$  of the dynamical system generally have k-fold periods,

 $k \le f$ , and they can be expressed as Fourier series with respect to time of the form

$$p_{s} = \sum_{\tau_{1},\tau_{2},...,\tau_{k}=-\infty}^{+\infty} P_{\tau_{1},\tau_{2},...,\tau_{k}} \exp\left\{2\pi i(\tau_{1}\nu_{1}+\tau_{2}\nu_{2}+\ldots+\tau_{k}\nu_{k})t\right\}$$
(23.20)

$$q_s = \sum_{\tau_1,\tau_2,\ldots,\tau_k=-\infty}^{+\infty} Q_{\tau_1,\tau_2,\ldots,\tau_k} \exp\left\{2\pi i(\tau_1\nu_1 + \tau_2\nu_2 + \ldots + \tau_k\nu_k)t\right\}$$

where the frequencies  $v_1, v_2, \ldots$ , and  $v_k$  can be shown to be determined by the relation

$$\nu_s = \frac{\partial W(J_1, J_2, \dots, J_k)}{\partial J_s}$$
(23.21)

which corresponds to Eq. (23.19) of the one-dimensional case. On the other hand,  $J_1, J_2, \ldots$  and  $J_k$  are to be quantized according to the quantum conditions

$$J_1 = n_1 h$$
,  $J_2 = n_2 h$ ,...,  $J_k = n_k h$  (23.22)

with  $n_1, n_2, \ldots, n_k$  integers.

The frequency of the light emitted in the transition from the stationary state of quantum numbers,  $(n_1, n_2, \ldots, n_k)$ , to that of quantum numbers  $(n_1 - \tau_1, n_2 - \tau_2, \ldots, n_k - \tau_k)$  is given by

which, for sufficiently large n's and for sufficiently small  $\tau$ 's, reduces approximately to

$$= \sum_{s} \tau_{s} \frac{\partial W(J_{1}, J_{2}, \dots, J_{k})}{\partial J_{s}}$$
  
=  $\sum_{s} \tau_{s} \nu_{s}$ . (23.23)

This is exactly the same as the frequency of the term in the Fourier series of Eq. (23.20) which corresponds to the higher harmonics with  $\tau_1, \tau_2, \ldots$  and  $\tau_k$ .

#### CH. 3, § 23] BOHR'S CORRESPONDENCE PRINCIPLE

The agreement thus observed of the quantum theoretical frequency with the classical one suggests a method of calculating the intensity of the light emitted by quantum transitions. Namely, we assume with Bohr that, in the case of very large n, the light emitted by transitions can be correctly described by the classical theory with respect, not only to its frequency, but also to its strength and polarization. In other words, we assume that the frequency, intensity and polarization of the light emitted in certain quantum transitions agree, for large n, with the frequency, intensity and polarization of the corresponding light in the classical theory. The meaning of "the corresponding light in the classical theory" is apparent from the consideration given above; i.e., to the transition from  $n_1, n_2, \ldots, n_k$  to  $n_1 - \tau_1$ ,  $n_2 - \tau_2$ , ...,  $n_k - \tau_k$  correspond the higher harmonics of  $\tau_1, \tau_2, \ldots$ ,  $\tau_k$  of the light classically emitted in the state of motion  $J_1 = n_1 h$ ,  $J_2 = n_2 h$ ,...,  $J_k = n_k h$ . Some of the  $\tau$ 's can take negative values provided that the resulting frequency  $v_{t_1, t_2}$  remains positive.

We may also state this correspondence as follows; to each of the quantum mechanical transitions corresponds a certain Fourier component of the classical motion and, in the case of large n, the nature of the light emitted in a quantum transition agrees with that of the light emitted classically by the corresponding Fourier component.

When we want to calculate the intensity of the light emitted in a certain quantum transition, we make use of this correspondence relation. That is, we first find the Fourier component of the motion in the classical theory which in the above sense corresponds to the quantum transition under consideration. Using the dipole moment, given by Eq. (23.3), corresponding to this Fourier component, we calculate the intensity of the emitted light by the formula of Eq. (23.5). When n is large, the intensity thus calculated will agree with that of the actually emitted light. When n is not so large, we naturally cannot expect a complete agreement but even in this case the calculated value will give an approximate measure of the actual intensity. This is the correspondence principle.

#### (ii) THE TRANSITION PROBABILITY

In the last section it was stated that, in the limit of large quantum number, the frequency, intensity, and polarization of the quantum theoretically emitted light agree with the corresponding quantities for the light emitted according to the classical theory. As has been repeatedly remarked, this does not at all imply that the emission mechanisms in the two theories are the same. In the classical theory, the atomic electrons emit the light in their process of orbital motion and they emit both the fundamental mode and the higher harmonics simultaneously. In the quantum theory, on the contrary, the light is emitted at the instant of the transition and the emitted light is a monochromatic one with the frequency given by Eq. (18.3). Therefore, when only a single transition of an atom is involved, one obtains, in a vague sense, only an instant flash of one of the spectral lines of the atom and the energy emitted thereby is exactly equal to  $h\nu$ . Under these circumstances, we cannot define the intensity of a spectral line. Since the intensity means the amount of energy emitted per unit time, knowledge of the rate of occurrence of such a transition is essential for the definition of intensity.

It is, then, necessarily concluded that we should be referring to the overall effect of a large number of atoms and not to a single transition of an atom when we say that the correspondence principle relates the quantum theoretical frequency or intensity to the respective counterpart in the classical theory. Physically the context of the correspondence principle primarily claims that when the intensity of a spectral line of an atom calculated in the classical theory is strong, or weak, the corresponding quantum transition should occur frequently, or rarely, in the aggregate of a large number of atoms of the kind under consideration. Namely, the correspondence can be realized only through the introduction of statistical considerations.

What mechanism then makes the transition occur? And what determines the rate of occurrence when some of the transitions occur more frequently and some others less frequently? Or, tracing the problem still further back, what is the reason why an atom cannot indefinitely stay in one of its stationary states but instead makes a transition to some other state? What factors then influence the atom in choosing a particular one of the many possible final states? Suppose that a number of atoms of the same kind are initially in the same stationary state; some of them will make transitions at time  $t_A$  to a state A while some others at time  $t_B$  to a state B. What difference in the two atoms of the same kind make them behave differently? We have, for the time being, no answer at all to these questions.

Under these circumstances, it seems advisable that with Einstein, we give up any inquiry into the causal structure of the transition and proceed with the assumption that they are governed by a certain probability law. That is, we assume that an atom in a stationary state has a probability of falling spontaneously in a given time interval to another state of lower energy. This probability may and will be determined by yet unknown laws of interaction between the atom and radiation, but here we just assume that this probability, like the one governing the decay of radioactive elements, is proportional to the time interval and that the proportionality constant, or the probability of occurence per unit time of the transition, is characteristic of the transition process under consideration. This constant is called the transition probability.

By  $A_{n \to n-\tau}$ , we denote the transition probability for an atom to make the transition from a state n to a state  $n - \tau$ . The frequency of the light emitted in this transition is given by Eq. (18.3) and hence the amount of energy released by the atom is

$$h\nu = W_n - W_{n-\tau}.$$

Accordingly, when an ensemble of a large number of atoms is considered, the intensity of the light emitted from this ensemble per one atom is given by

$$\frac{\mathrm{d}E_{\tau}}{\mathrm{d}t} = h\nu_{\tau} A_{n \to n-\tau} \,. \tag{23.24}$$

Although the laws of interaction between the atom and radiation are not known, Bohr's correspondence principle furnishes a method of calculating the transition probability. Namely, the principle requires that the quantum theoretical intensity given by Eq. (23.24)should agree with the one obtained from classical theory, Eq. (23.5). By comparing these two equations we find that the transition probability is to be given by

$$A_{n \to n-\tau} = \frac{(2\pi)^4 \nu_{\tau}^3}{3c^3 h} | C_{\tau} |^2 . \qquad (23.25)$$

As has been repeatedly stated, this result is correct only for very large n but would give an approximate value also for smaller n.

This relationship, Eq. (23.25), was applied to a number of related problems and its validity was verified by comparing with experimental results. It is true that the agreement with experiments was not very good in the case of not large n, but even there the relationship proved to be useful in giving the over-all feature of the phenomena. In particular, the usefulness of this relationship becomes quite apparent if one observes that, even in some cases of small quantum numbers, Eq. (23.25) holds correctly. These are the cases where, according to the classical approach, the amplitude  $C_{\tau}$  of the corresponding Fourier component becomes zero and, hence, the corresponding transition probability  $A_{n\to n-\tau}$  of Eq. (23.25) becomes zero implying that the transition  $n \to n - \tau$  does not actually occur; i.e., this transition is "forbidden". It was verified experimentally that an otherwise possible transition does not actually occur when it is forbidden by the correspondence principle and that this holds irrespective of the magnitude of the quantum number n.

#### (iii) THE CASE OF THE HARMONIC OSCILLATOR

As a simple illustration of the correspondence principle, we take, in this section, the case of Planck's oscillator.

The motion of the mass point in the case of the harmonic oscillator is quite simple and the Fourier series consists only of two terms, namely,

$$x = X_1 e^{2\pi i v t} + X_{-1} e^{-2\pi i v t}, \qquad (23.26)$$

where the x-axis is taken in the direction of the oscillation. When the point mass has an electric charge e, the dipole moment of this harmonic oscillator is given by

$$P = \frac{1}{2} \left( C_1 e^{2\pi i \nu t} + C_{-1} e^{-2\pi i \nu t} \right)$$
  
=  $-e \left( X_1 e^{2\pi i \nu t} + X_{-1} e^{-2\pi i \nu t} \right).$  (23.27)

Since the Fourier series of the dipole moment has only terms with  $\tau = \pm 1$ , all but the transitions  $n = n \pm 1$  are forbidden and, in the emission of light, only the transition  $n \to n - 1$  is allowed (since we have  $W_n > W_{n-1}$ ). If the transition  $n \to n - \tau$  were possible, the frequency of the light emitted by the oscillator would be given by

$$\nu_{\tau} = \frac{W_n - W_{n-\tau}}{h}$$
  
=  $\frac{1}{h} \{ hvn - hv(n-\tau) \} = \tau v$ ,  $\tau = 1, 2, 3, ...$ 

But since, as shown above, all except the transition with  $\tau = 1$  are forbidden, only light of frequency corresponding to  $\tau = 1$ , i.e.

$$\boldsymbol{\nu}_1 = \boldsymbol{\nu} \tag{23.28}$$

is actually emitted. In the case of the harmonic oscillator, therefore, the quantum theory and the classical theory give identical results for the frequency of the emitted light.

The transition probability in this case is according to the formula of Eq. (23.25)

$$A_{n \to n-1} = \frac{(2\pi)^4 \nu^3}{3c^3 h} |C_1|^2$$
$$= \frac{4(2\pi)^4 \nu^3 c^2}{3c^3 h} |X_1|^2,$$

or, it can also be written as

$$A_{n \to n-1} = \frac{(2\pi)^4 \nu^3 e^2}{3c^3 h} a^2 \tag{23.29}$$

where a stands for the amplitude of the oscillation as given in Eq. (23.26); for if we write

$$x = a \cos(2\pi v t + \delta)$$
,

 $a = 2 |X_1|$ . Besides this the amplitude of oscillation in the state of quantum number *n* is given by Eq. (19.46), i.e.,

$$a^2 = \frac{n}{\pi (2\pi\nu)m} h \tag{23.30}$$

and the transition probability is written as

$$A_{n \to n-1} = \frac{8\pi^2}{3} \frac{e^2 \nu^2}{mc^3} n . \qquad (23.31)$$

(iv) THE SELECTION RULE

As we have seen above in the example of the harmonic oscillator, among a great number of transitions that are *a priori* permissible, very few can actually happen, most being forbidden. Therefore, in atomic spectra, not all the possible combinations of two spectral terms give rise to frequencies of the actually observed light. We must then inquire between which terms is the transition actually allowed and between which is it forbidden, or, in other words, which terms combine with each other and which do not. It is possible to derive a formula of rather general applicability which we call a selection rule. In this section, we derive the selection rule for one of the simplest cases of an electron moving in the field of a central force. The result however can be extended to cover more general cases. For the sake of simplicity, we give up the three-dimensional approach and satisfy ourselves with the two-dimensional case.

Treating, in polar coordinates in two dimensions, the problem of a point mass in the field of a central force, we have the Hamiltonian function given by

$$H = \frac{1}{2m} \left( p_r^2 + \frac{p_{\varphi}^2}{r^2} \right) + V(r), \qquad (23.32)$$

where the potential V(r) of the central force can be of any functional form as long as the point mass can perform some doubly-periodic motion.

Readily derived is the conservation of angular momentum,

$$p_{\sigma} = \text{const.} = M , \qquad (23.33)$$

which, by the introduction of  $J_{\varphi}$  as usual, can be written as

$$M = \frac{J_{\varphi}}{2\pi}.$$
 (23.34)

Substituting Eqs. (23.33) and (23.34) into Eq. (23.32), we obtain

$$H = \frac{1}{2m} \left\{ p_r^2 + \left( \frac{J_\varphi}{2\pi} \right)^2 \cdot \frac{1}{r^2} \right\} + V(r) . \qquad (23.32')$$

For the given value, (23.34), of the angular momentum, the motion of the system with respect to r is easily seen to be the same as the motion of a one dimensional system with the Hamiltonian function of Eq. (23.32'). Hence, the determination of the period of this motion can be carried out in the same way as in Eq. (23.13) through Eq. (23.18). Namely, putting

$$\frac{1}{2m}\left\{p_r^2 + \left(\frac{J_\varphi}{2\pi}\right)^2 \cdot \frac{1}{r^2}\right\} + V(r) = W, \qquad (23.35)$$

which is solved to give

$$p_r = p_r (r, W, J_{\varphi}),$$
 (23.36)

we calculate

$$J_{r} = 2 \int_{r_{1}}^{r_{2}} p_{r} (r, W, J_{\varphi}) dr \qquad (23.37)$$

and obtain  $J_r$  as a function of W and  $J_{\varphi}$ , i.e.

$$J_{\mathbf{r}} = J_{\mathbf{r}} \left( W, J_{\varphi} \right) \,. \tag{23.38}$$

W is then obtained as a function of  $J_r$  and  $J_{\varphi}$ , i.e.,

$$W = W \left( J_{r}, J_{\varphi} \right) , \qquad (23.39)$$

which by differentiating with respect to  $J_r$  gives the fundamental frequency  $v_r$  of the motion in the r coordinate, namely,

$$\mathbf{v}_{\mathbf{r}} = \frac{\partial W \left( J_{\mathbf{r}}, J_{\boldsymbol{\varphi}} \right)}{\partial J_{\mathbf{r}}}.$$
(23.40)

In terms of this frequency the radius r is expressed in the form of Fourier series as

$$r = \sum_{\tau_r = -\infty}^{+\infty} R_{\tau_r} e^{2\pi i \tau_r \nu_r t}. \qquad (23.41)$$

The reason why we have here an ordinary Fourier series with a single fundamental frequency  $v_r$  instead of a doubly periodic one is that the motion in the r direction can be regarded as the motion of a one-dimensional system as mentioned above. The actual determination of the coefficients  $R_{\tau_r}$  needs a more detailed solution of the equation of motions and, hence, the specification of V(r) to start with. For our present purpose we do not have to go into such details.

We instead calculate the quantity  $dW(J_r, J_{\varphi}) / dJ_{\varphi}$  which is, according to the general theory, to be identified as the fundamental frequency  $v_{\varphi}$  of motion in the  $\varphi$  coordinate, i.e.,

$$\boldsymbol{v}_{\boldsymbol{\varphi}} = \frac{\partial W\left(\boldsymbol{J}_{\boldsymbol{r}}, \, \boldsymbol{J}_{\boldsymbol{\varphi}}\right)}{\partial \boldsymbol{J}_{\boldsymbol{\varphi}}}.$$
(23.42)

A remark may be needed as to the meaning of the frequency  $\nu_{\varphi}$  in the  $\varphi$ -direction, since  $\varphi$  is an angular variable increasing indefinitely with

153

time and, hence, does not show a periodic behavior. What  $\nu_{\varphi}$  means will be seen in the following manner.

Let us investigate the nature of the motion in the  $\varphi$ -direction. From the Hamiltonian function of Eq. (23.32), we readily obtain

$$\dot{\varphi}=\frac{p_{\varphi}}{mr^2},$$

which, by the use of Eqs. (23.33) and (23.34), becomes

$$\dot{q} = \frac{1}{mr^2} \frac{J_o}{2\pi}.$$
 (23.43)

We have already seen that the coordinate r is a periodic function of time as given by Eq. (23.41). The factor  $1/mr^2$  in Eq. (23.43) should then be expressed as a Fourier series similar to Eq. (23.41), the constant term of which, i.e., the term  $\tau_r = 0$ , is obviously  $\langle 1/mr^2 \rangle$ , the average of  $1/mr^2$  over one period of r. Hence, Eq. (23.43) can be rewritten as

$$\dot{\varphi} = \left\langle \frac{1}{mr^2} \right\rangle \frac{J_{\varphi}}{2\pi} + F , \qquad (23.44)$$

where F stands for a periodic function of time which has the fundamental frequency  $v_r$  but lacks the constant term in the Fourier expansion. Integrating Eq. (23.44) with respect to t, we get

$$\varphi = \left\langle \frac{1}{mr^2} \right\rangle \frac{J_{\varphi}}{2\pi} t + G , \qquad (23.45)$$

where G is also a periodic function of time with  $v_r$  as the fundamental frequency since it is an integral of F which has these characteristics. Thus we see that the motion in the  $\varphi$ -direction consists of a uniform rotation with a periodic oscillation superposed on it. The angular frequency of the former is  $\langle 1/mr^2 \rangle J_{\varphi}/(2\pi)^2$  while the fundamental frequency of the latter is  $v_r$ . We shall now prove that the frequency  $v_{\varphi}$  defined by (23.42) is simply this angular frequency of the uniform rotation.

The calculation of  $r_{\varphi}$  can be carried out first, noting that

$$\frac{\partial W(J, J_{\varphi})}{\partial J_{\varphi}} = -\frac{\frac{\partial J_{r}(W, J_{\varphi})}{\partial J_{\varphi}}}{\frac{\partial J_{\varphi}}{\partial W}},$$

which results from Eq. (23.38). By the use of Eq. (23.37), it is further transformed to

$$= \frac{\oint \frac{\partial p_r(r, W, J_{\varphi})}{\partial J_{\varphi}} \, \mathrm{d}r}{\oint \frac{\partial p_r(r, W, J_{\varphi})}{\partial W} \, \mathrm{d}r},$$

Furthermore by the use of the relations

$$\frac{\partial p_r(r, W, J_{\varphi})}{\partial W} = \frac{1}{\frac{\partial W(r, p_r, J_{\varphi})}{\partial p_r}} = \frac{1}{\frac{\mathrm{d}r}{\mathrm{d}t}} \quad \text{and}$$

$$\frac{\partial p_r(r, W, J_{\varphi})}{\partial J_{\varphi}} = -\frac{\frac{\partial W(r, p_r, J_{\varphi})}{\partial J_{\varphi}}}{\frac{\partial W(r, p_r, J_{\varphi})}{\partial p_r}} = -\frac{\frac{1}{\frac{\mathrm{d}r^2}{\mathrm{d}t}} \frac{J_{\varphi}}{\mathrm{d}t}}{\frac{\mathrm{d}r}{\mathrm{d}t}}$$

which are obtained from the equation of motion together with the fact that  $p_r$  is the solution of Eq. (23.35), we finally obtain

$$\boldsymbol{v}_{\varphi} = \frac{\partial W(J_r, J_{\varphi})}{\partial J_{\varphi}} = \frac{\frac{J_{\varphi}}{(2\pi)^2} \int_{0}^{T_r} \frac{1}{mr^2} dt}{\int_{0}^{T_r} dt}$$
$$= \left\langle \frac{1}{mr^2} \right\rangle \frac{J_{\varphi}}{(2\pi)^2}, \qquad (23.46)$$

where  $T_r$  is the period of motion in the *r*-direction so that the multiplying factor of  $J_{\varphi}/(2\pi)^2$  is the average of  $1/mr^2$ . Eq. (23.46) is the first required result. By the use of Eq. (23.43), Eq. (23.45) can be written as

$$\varphi = 2\pi v_{\varphi} t + G$$

$$\nu_{\varphi} = \frac{\partial W(J_r, J_{\varphi})}{\partial J_{\varphi}}.$$
(23.47)

The nature of the variations of both r and  $\varphi$  being thus known, the time variation of the x and y coordinates and hence the dipole moment

of the system can also be inferred. If the point mass has an electric charge -e, its dipole moment is given by

$$P_{x} = -ex = -er \cos \varphi$$

$$P_{y} = -ey = -er \sin \varphi .$$
(23.48)

From Eq. (23.47) we have

$$\cos \varphi = \frac{1}{2} \left( e^{i\varphi} + e^{-i\varphi} \right)$$
$$= \frac{1}{2} \left( e^{iG} e^{2\pi i v \varphi t} + e^{-iG} e^{-2\pi i v \varphi t} \right),$$

which together with Eq. (23.41) in Eq. (23.48) leads to

$$P_{x} = \frac{1}{2} \sum_{\tau_{r}=-\infty}^{+\infty} \left\{ C_{\tau_{r}} e^{2\pi i (\tau_{r} v_{r} + v_{\varphi})t} + C_{-\tau_{r}}^{*} e^{2\pi i (\tau_{r} v_{r} - v_{\varphi})t} \right\}$$

$$P_{y} = \frac{1}{2} \sum_{\tau_{r}=-\infty}^{+\infty} \left\{ -iC_{\tau_{r}} e^{2\pi i (\tau_{r} v_{r} + v_{\varphi})t} + iC_{-\tau_{r}}^{*} e^{2\pi i (\tau_{r} v_{r} - v_{\varphi})t} \right\}$$
(23.49)

where use was made of the fact that both G and r are periodic functions with fundamental frequency  $v_r$  and so also are  $re^{-iG}$  and  $re^{-iG}$ . In the above expressions for  $P_x$  and  $P_y$  as double Fourier series with fundamental frequencies  $v_r$  and  $v_{\varphi}$ , it should be noticed that terms of higher harmonics with respect to  $v_{\varphi}$  are not involved; the only terms which enter are those with  $\tau_{\varphi} = \pm 1$ .

In quantum theory, the quantum numbers n' and k are defined through the J's as

$$J_r = n'h \quad \text{and} \quad J_\varphi = kh \,. \tag{23.50}$$

According to the general rules of the correspondence principle, we can conclude the following from the result of Eq. (23.49). That is, in quantum transitions, the quantum number k either increases by one unit or decreases by one unit and other variations of k are forbidden. Recalling the rules for polarization given previously, Eq. (23.8) and Eq. (23.9), we can state that the light emitted by the transition  $k \rightarrow k-1$  (or  $k \rightarrow k+1$ ) has the same polarization as light from a dipole moment which is rotating in an anti-clockwise (or clockwise) direction in the xy plane.

We thus have derived the rule that the subsidiary quantum number k can change only by an amount  $\pm 1$ . This is the selection rule for the subsidiary quantum number. On the other hand, there is no such selection rule for the quantum number n', since Eq. (23.49) contains all  $\tau_r$  values. Therefore, n', and hence n, can vary by any amount. In fact, we can take any combination of values for the two principal quantum numbers in the Balmer formula.

The treatment of the problem in three dimensions yields the selection rule for the magnetic quantum number m, which says that in quantum transitions the magnetic quantum number changes by  $\pm 1$  or remains unchanged. Furthermore, it is seen that in the transition  $m \rightarrow m$  the dipole moment is oscillating in the direction of the magnetic field, while in the transition  $m \rightarrow m - 1$ , or  $m \rightarrow m + 1$ , it has, when viewed in the direction of the magnetic field, a clockwise, or anticlockwise, rotation in a plane normal to the field. The selection rule for k is the same as in the two dimensional case and allows only the transitions  $k \rightarrow k - 1$ , and  $k \rightarrow k + 1$ . A definite polarization in each transition of the orbit is not fixed.

In short, possible transitions for such an atom are

$$(n', k - 1, m - 1)$$

$$(n', k - 1, m)$$

$$(n, k, m) \rightarrow \frac{(n', k - 1, m + 1)}{(n', k + 1, m - 1)}$$

$$(n', k + 1, m)$$

$$(n', k + 1, m + 1).$$

$$(23.51)$$

The selection rule for the magnetic quantum number explains why we observe the splitting of a spectral line into only three lines in the Zeeman effect. The explanation goes as follows. The magnetic quantum number m of an atom can assume (2k + 1) values from -k to k. Therefore, when the atom makes a transition from the initial state of subsidiary quantum number k to the final one of k', the number of all the combinations of initial and final terms of various m values is equal to (2k + 1) (2k' + 1). This, however, is not the number of spectral lines to be observed, since, these terms in the presence of a magnetic field H being spaced equidistantly with the separation of  $\Delta E =$ 

 $(eh/4\pi\mu c)H$ , most of the above combinations give the same frequency. Neverthcless, if the above combinations are all allowed, we should observe (2k + 2k' + 1) spectral lines. It is the selection rule for m, i.e.  $m \rightarrow m$  or  $m \rightarrow m \pm 1$ , that gives the further reduction of this number to three in conformity with the experimental result. In Fig. 41, the case of k = 3 and k' = 2 is schematically illustrated. The arrows in the figure show the allowed transitions and the symbols +, 0, and signify that the transition is from m to m + 1, m, and m - 1, respectively. Due to the fact that these energy levels are spaced at equal intervals and that the level with m = 0 is not shifted by the effect of the magnetic field, we see that the frequency of the emitted light in the presence of the magnetic field H is increased, or decreased, as compared to the frequency without the magnetic field, by the same amount  $(eh/4\pi \mu c)H$  in all the transitions  $m \to m - 1$ , or  $m \to m + 1$ and unchanged in all the transitions  $m \rightarrow m$ . We, therefore, observe only three lines. Furthermore, the rules for polarization show that the



Fig. 41. Transitions in the normal Zeeman effect.

light of unchanged frequency has the same polarization as the one emitted from a dipole oscillation in the direction of the magnetic field while the light of increased, or decreased, frequency has the same polarization as the one from a dextrogyrating, or levogyrating, dipole, respectively, in the plane normal to the field direction. In other words, when viewed from the direction of the magnetic field the light of increased, or decreased, frequency shows right-handed, or left-handed, circular polarization while the one of unchanged frequency is missing. Observation from the direction normal to the magnetic field gives us light of unchanged frequency having linear polarization in the direction of the field together with the ones of shifted frequencies, both linearly polarized in the direction normal to the field. The situation is shown schematically in Fig. 42. These results are all veri-

fied experimentally.

We have in the above explained the selection rule for the cases where one electron is moving in the field of a certain central force. Also for more general atoms, we can derive the selection rules for the changes of the angular momentum or of its z-component but we do not go into the details of this derivation here.

# (v) THE CORRESPONDENCE PRINCIPLE AS A GUIDING PRINCIPLE

The correspondence principle cannot only be applied, as in the preceding section, to the case of spontaneous light emission by an atom but also to cases where an atom absorbs the incoming light, or makes

a forced emission of light (caused by the influence of the incident light), or where an atom scatters the incident light. The applications of the principle to these cases can be made in almost the same way as has been done for the spontaneous emission. Summarizing the procedure, we first solve the given problem by the classical approach and then reinterpret the results by replacing the concept of the frequency of orbital motion (or of the Fourier component) with the corresponding concept of the quantum frequency (or the probability) of the transition; i.e., the classical frequency  $v_{t+1}$  is replaced by the quantum theoretical one of the transition  $n_1$  $\rightarrow n_1 - \tau_1, n_2 \rightarrow n_2 - \tau_2, \ldots, n_k \rightarrow n_k - \tau_k$  and the square of the Fourier component,  $|C_{\tau_1 \tau_2 \dots \tau_k}|^2$ , of the dipole moment by the corresponding tran-

sition probability  $A_{n_1n_2...n_k \rightarrow n_1 - \tau_1, n_2 - \tau_2,..., n_k - \tau_k}$ . In this way Bohr's theory is supplemented and a useful method for calculating various processes is obtained. Nevertheless, this procedure is to be considered as merely a convenient recipe for calculating quantum theoretical quantities such as transition probabilities, since we still do not know the cause of a quantum transition. When we arrive at the true theory, the first thing to be clarified is the mechanism through which only a certain discrete set of states can occur in nature and then to understand what determines why some atoms jump from A to B at time  $t_B$  while some others jump from A to B' at  $t_{B'}$  and so on. We would then be able to calculate the frequency of occurences of such transitions and presumably would find that it actually is in harmony with the correspondence principle result in the limit of large n.

It is very natural to anticipate such a future for the quantum theory. However, in applying the correspondence principle to various problems, a group of physicists, with Bohr as leader, began to think differently. Namely, they began to realize that the nature of the discontinuities or of the transitions should be sought in the correspondence prin-



Fig. 42. Polarizations of the triplet lines in the normal Zeeman effect. The upper figures are for the direction parallel to the magnetic field, the lower for the perpendicular direction. Arrows above the spectral lines show the polarization. ciple itself and that there are no *ad hoc* fundamental laws which have no correspondence to the classical theory. According to the viewpoint of common sense, a hidden mechanism is to exist to make the states discontinuous, and there should be laws of a more fundamental nature which describe the course of a transition, but this viewpoint should be abandoned. The correspondence principle so far described is still too vague in its formulation, but, following Bohr, it is anticipated that the correct laws of the quantum world should be obtained not by introducing certain additional laws for the transition mechanism but instead by a revised form, expressed mathematically in a clear-cut way, of the correspondence principle itself.

That the correspondence principle described so far in this section is still incomplete is clear from the way it is applied in solving a problem. In fact, we first treat the problem in terms of classical ideas. The obtained results are then reinterpreted from the quantum mechanical viewpoint by replacing certain quantities with the corresponding quantum mechanical ones. This has been the procedure of applying the correspondence principle. The true theory cannot have such an unsatisfactory form but must be expressed, from the very beginning, in terms of only quantum mechanical quantities obeying certain laws. These laws will be similar in form to those of classical mechanics, or electrodynamics, governing the corresponding classical quantities and in the limit of very large quantum numbers the conclusions of the two will coincide with each other.

Thus, if one is to proceed without expecting laws which have no correspondence to the classical theory, the question, for instance, of why the states of a periodic dynamical system become discrete should be answered as follows. Namely, the motion of a periodic system is described in the classical theory in terms of a Fourier series, i.e., the motion contains, besides the fundamental mode, discrete modes of higher harmonics corresponding to  $\tau = 1, 2, 3, \ldots$ . The correspondence principle, translating this result into quantum theoretical language, concludes that the energy of the system in those states to which it makes the transitions has discrete values. This is the answer and one should not look for any unknown mechanism behind this discontinuity. Furthermore, the causal description of the transition processes, particularly the factor which determines why some atoms jump from A to B at  $t_B$  while some others jump from A to B' at  $t_{B'}$  is also to be given up, since the correspondence principle does not say anything

about it. What the correspondence principle tells us is that the jump is due to the interaction between the atom and radiation, corresponding to the classical fact that the orbit of the electron shrinks more and more by emitting light due to this interaction, and that each transition corresponds to a Fourier component of the classical orbital motion, the probability of the transition being given by something which corresponds to the square of the Fourier amplitude. There is no mechanism for the transition process, independent of the shrinking of orbital motion presupposed in the structure of the correspondence principle. Accordingly, an important role would be played in the true theory by a certain quantity which, when squared, determines the probability of the transition and would take over the concept of the orbital motion in conventional mechanics. Furthermore, this quantity in the true theory would oscillate with the quantum theoretically correct frequency instead of the classical one  $\nu_r$ . Thus this "certain quantity", describing the statistical properties of the atom, would play in the true theory a fundamental role similar to that of the orbital motion in the classical theory. What we have called the "motion" in the classical sense would fade away in the new theory and the time variation of this "certain quantity" which gives rise to the statistical nature of various transitions would constitute the main subject of the theory. Corresponding to the circumstance in the classical theory that the emission of light is due to the interaction between the orbital motion of the electrons and the electromagnetic field, the transition would then be caused by the interaction between the "certain quantity" and the electromagnetic field. The temporal variation of this "certain quantity" would be determined in the same way as the Fourier component of the orbital motion in the classical theory, although the causal description of the transition itself would have to be given up. Thus the transition law in the new theory would not be obtained by introducing an additional mechanical entity into the classical theory and giving a new law to govern it, but instead, would be achieved without increasing the number of basic clements of the theory. In other words, all the elements of the new theory should have corresponding elements in the classical theory and vice versa. The correspondence principle in this interpretation then becomes the guiding principle for suggesting what form the future theory should take, and is not merely a convenient procedure for remedying the defects of Bohr's theory.

According to this point of view, the previously introduced Hypotheses III and IV are to be regarded as incomplete and tentative in nature and are to be reformulated in accordance with the correspondence principle. We have assumed that the motion of a system in a stationary state, or in an adiabatic process, obeys classical mechanics. In the new theory, however, the concept of orbital motion, for instance, is not permissible. Even Bohr's quantum condition I = nh is also to be reformulated in accordance with the new interpretation of the correspondence principle, for it is still expressed as an ad hoc condition on the orbital motion; that is, as mentioned before, the discrete nature of the stationary states should appear as something which corresponds to the discreteness of  $\tau$  instead of to such an *ad hoc* condition as Bohr's. In fact, the straight-forward application to complex dynamical systems, of Hypotheses III and IV is known to sometimes lead to inconsistent results. Take, for instance, a dynamical system containing two electrons such as a hydrogen molecule. According to Hypothesis III, two electrons in stationary states of the system will interact with each other in accordance with classical mechanics and hence the effect of one electron on the other is to be considered as due to some periodic force with a period obtained by combining those of orbital motions of the two electrons. On the other hand, when the two atoms of the hydrogen molecule are separated so far that the two electrons belong to different atoms, the interaction between the electrons is affected by the absorption, by one atom, of light emitted by the other atom. The period of the force acting on one of the electrons has then nothing to do with that of its orbital motion. As a matter of fact, not only the hydrogen molecule but even the helium atom could not be handled successfully by the old quantum theory. As is shown here in a simple example, the old quantum theory contained many difficulties and their final solution had to wait till a young German physicist, Heisenberg, made a giant step forward in 1925.

## § 24. The Quantization of Rotational and Translational Motion

The main topics of the present chapter have been concluded in the preceding section and this section is meant to give some supplementary notes on the quantization of rotational and translational motion. The former is related to the problem of the specific heat of diatomic molecules while the latter is needed in the later development of the theory.

#### (i) THE QUANTIZATION OF ROTATIONAL MOTION

It was stated previously in § 2(ii) that the specific heat of a diatomic gas behaves at low temperatures like that of a monoatomic gas and we conjectured there that this was due to the dying out of rotational motion at low temperatures. In order to see if this conjecture is correct, we have to estimate the magnitude of the energy quantum for rotational motion and compare it with kT to find the temperature at which kTbecomes smaller than the quantum of energy or, in other words, at which the dying out of rotational motion begins. Comparison of the temperature thus found with the experimental one will show the validity of our interpretation. Let the moment of inertia of the diatomic molecule under consideration and its angular momentum vector be I and M, respectively. Its energy E is then given by

$$E = \frac{\mathsf{M}^2}{2I} \,. \tag{24.1}$$

But as we have seen in § 20, the angular momentum is quantized and its magnitude M is given by

$$M = \frac{h}{2\pi}n$$
,  $n = 0, 1, 2, ...$  (24.2)

Consequently, the energy E is also quantized and

$$E = \frac{h^2}{8\pi^2 I} n^2 , \qquad n = 0, 1, 2, \dots .$$
 (24.3)

Eq. (24.3) gives as the energy difference between the lowest energy state and the next higher one

$$\Delta E = \frac{h^2}{8\pi^2 I} \,. \tag{24.4}$$

It is then conjectured that, at temperatures at which kT becomes smaller than  $\Delta E$  of Eq. (24.4), the rotational motion dies out; thus the specific heat of a diatomic gas behaves like that of a monatomic gas at low temperatures T satisfying

$$T \lesssim \frac{h^2}{8\pi^2 kI}.$$
 (24.5)

This relationship can be compared with experiment by taking the case of hydrogen gas as an example. A hydrogen molecule consists of two protons separated by a distance of about  $10^{-8}$  cm. The mass of a proton being  $1.66 \times 10^{-24}$  gram, the moment of inertia I becomes

$$I = 2 \times 1.66 \times 10^{-24} \times (\frac{1}{2} \times 10^{-8})^2$$
  

$$\approx 0.83 \times 10^{-40} \,\mathrm{g \cdot cm^2} \,.$$

Introducing this value into Eq. (24.5), we obtain

$$T \lesssim 50^{\circ} \text{ K}$$
 .

In the meantime, inspection of Fig. 2 of § 2 tells us that the experimentally measured specific heat of the hydrogen gas begins to approach that of a monatomic gas at around this temperature. The agreement with experiment is thus satisfactory and our conjecture is thereby confirmed. A later modification of Bohr's theory leads to

$$E = \frac{h^2}{8\pi^2 I} n(n+1) , \qquad (24.3')$$

instead of Eq. (24.3), and the use of Eq. (24.3') together with statistical mechanics reproduces the experimental specific heat versus temperature curve of Fig. 2.

## (ii) THE QUANTIZATION OF TRANSLATIONAL MOTION

We consider a number of electrons contained in a box the walls of which totally reflect the electrons. Furthermore, for simplicity, the electrons are assumed to move independently of each other as point masses and the box is assumed to be a cube of side L. Obviously such a dynamical system does not exist in nature but under a crude approximation the electrons in metals can be described by this simple model. For inside a metal there exists a number of "conduction electrons" and they move almost freely in the metal so that this model
#### CH. 3, § 24] ROTATIONAL AND TRANSLATIONAL MOTION

of the "electron gas" can in fact be used in discussing the electrical conductivity or other characteristics of a metal.

Now, since in our dynamical system each electron moves freely as a point mass, the Hamiltonian of the system is

$$H = \sum_{l=1}^{N} H_{l}$$

$$H_{l} = \frac{1}{2m} \{ p_{l_{x}}^{2} + p_{l_{y}}^{2} + p_{l_{z}}^{2} \}$$
(24.6)

where N denotes the total number of electrons.

This Hamiltonian is expressed as a sum of the Hamiltonian functions for each degree of freedom just as in Eq. (19.26) of § 19 and consequently the quantum condition can be applied separately for each degree of freedom. They are

$$\oint p_{l_x} dx = n_{l_x} h,$$

$$\oint p_{l_y} dy = n_{l_y} h, \quad n_{l_x}, n_{l_y}, n_{l_z} = 0, 1, 2, \dots, l = 1, 2, \dots, N.$$

$$\oint p_{l_z} dz = n_{l_z} h,$$
(24.7)

Under our simplifying assumption of totally reflecting walls, the x, y and z coordinates of each electron oscillate in a periodic manner between 0 and L and hence the above integrals can be taken as twice the integral from zero to L. Since  $|p_{l_x}|, |p_{l_y}|$  and  $|p_{l_x}|$  remain constant, the quantum conditions for our dynamical system give

$$|p_{l_x}| = \frac{h}{2L} n_{l_x}, \quad |p_{l_y}| = \frac{h}{2L} n_{l_y}, \quad |p_{l_z}| = \frac{h}{2L} n_{l_z},$$
  
$$n_{l_x}, n_{l_y}, n_{l_z} = 0, 1, 2, \dots, \quad l = 1, 2, \dots, N.$$
(24.8)

Thus we can say that the translational motions of the electrons of our dynamical system are quantized and the x, y and z components of their momenta can only be integral multiples, positive as well as negative, of h/2L.

Since the values of the moment of the electrons are given by Eq. (24.8), the total energy of the system can assume only the values given by

$$\frac{h^2}{8mL^2} \sum_{l=1}^{N} \left( n_{l_x}^2 + n_{l_y}^2 + n_{l_z}^2 \right)$$

$$n_{l_x}, n_{l_y}, n_{l_z} = 0, 1, 2, \dots$$
(24.9)

165

This shows that the energy of the translational motion is also quantized.

In Eq. (24.9), the total energy of the system is obtained by first finding the value which the momentum of each individual electron takes of the values allowed by Eq. (24.8), and thereby the  $n_{l_x}$ ,  $n_{l_y}$  and  $n_{l_x}$  values with l = 0, 1, 2, ... N. However, the energy can be obtained also in the following manner. Namely, since the momentum of each electron is given, according to Eq. (24.8), by

$$|p_{x}| = \frac{h}{2L} n_{x}, \quad |p_{y}| = \frac{h}{2L} n_{y}, \quad |p_{z}| = \frac{h}{2L} n_{z}$$

$$n_{x}, \quad n_{y}, \quad n_{z} = 0, 1, 2, \dots$$
(24.10)

we can count the number of the electrons with momenta given by a particular set of these values. Denoting this number by  $N_{n_{x_i}, n_{y_i}, n_{z_i}}$  corresponding to the values  $n_x$ ,  $n_y$ , and  $n_z$ , we have as the total energy of the system

$$H = \frac{h^2}{8mL^2} \sum_{n_x, n_y, n_x} N_{n_x, n_y, n_x} \left( n_x^2 + n_y^2 + n_z^2 \right).$$
(24.11)

Mathematically, Eq. (24.11) is obtained by a mere rearrangement of the summands of Eq. (24.9). That is, among the summands on the right hand side of Eq. (24.9), we first pick up those *l*'s for which  $n_{l_x} = n_x$ ,  $n_{l_y} = n_y$  and  $n_{l_x} = n_z$ . The number of such *l*'s being denoted by  $N_{n_x, n_y, n_z}$ , the partial sum of *H* becomes

$$\frac{h^2}{2mL}N_{n_x,n_y,n_z}(n_x^2+n_y^2+n_z^2)$$

and the total H is given by summing this up for all combinations of  $n_x$ ,  $n_y$  and  $n_z$ . Obviously the condition

$$\sum_{n_x, n_y, n_x} N_{n_x, n_y, n_x} = N .$$
 (24.12)

must be satisfied.

#### CHAPTER 4

## THE SHELL STRUCTURE OF THE ATOM

## § 25. The Theory of Optical Spectra

As was stated at the end of §23, the old quantum theory had trouble even with the helium problem and we cannot expect the theory to give detailed conclusions concerning the structure of atoms more complex than hydrogen. Nevertheless, it is clear from its success in the case of the hydrogen atom that the gross features of the theory must be correct. In fact, by the aid of this theory, one can discuss, though semi-empirically, various problems connected with the structure of atoms and a number of important discoveries were thus made. The most important of all is that concerning the shell structure of the atom, which will be described in the present chapter. The periodicity of the elements discovered by Mendelejeff is closely related to this shell structure of the atom and, accordingly, the discussion of this structure is extremely important since its implications are not restricted to physics, but extend also to the field of chemistry. Moreover, through these discussions, important discoveries such as the electron spin and Pauli's principle were made.

The discussion of the shell structure has very much to do with the spectroscopy of atomic emission lines. For it is by analysing the complex terms of the experimentally observed atomic spectra and by comparing them with the theory that the structure of the atom is inferred.

The method of analysing the experimentally observed complex spectra is called spectroscopy and constitutes by itself a vast branch of physics; hence we are obliged to restrict ourselves only to its general outlines without going into details.

As Rydberg observed, the frequencies of the atomic emission spectra of various elements can be expressed by the formula

$$\frac{v}{c} = \frac{R}{(n'+a')^2} - \frac{R}{(n+a)^2}, \quad n', n = \text{integers}. \quad (25.1)$$
167

For instance, Rydberg found for the spectra of alkali metals, or more specifically for Na, three series corresponding to the formulae;

the principal series, 
$$\frac{v}{c} = \frac{R}{(3+s)^2} - \frac{R}{(n+p)^2}$$
,  $n = 3, 4, \ldots$ 

the 1st subordinate series,  $\frac{v}{c} = \frac{R}{(3+p)^2} - \frac{R}{(n+d)^2}$ ,  $n = 3, 4, \ldots$ 

the 2nd subordinate series, 
$$\frac{\nu}{c} = \frac{R}{(3+p)^2} - \frac{R}{(n+s)^2}$$
,  $n = 3, 4, ...$ 
(25.2)

In addition to these, there is another one found by Bergman; for Na, for example, we have

the Bergman series, 
$$\frac{v}{c} = \frac{R}{(3+d)^2} - \frac{R}{(n+f)^2}$$
,  $n = 4, 5, ...$  (25.2')

In these expression, s, p, d, or f stand for particular values of a or a' in Eq. (28.1) and they give the deviation of each series from the corresponding one of the hydrogen spectra. They are called Rydberg's corrections and are known to depend on neither n nor n'. Among these corrections, f is especially small and is almost zero.

Eq. (25.2) and Eq. (25.2') imply, according to Bohr's interpretation, that these atoms have energy levels of the form

$$E_n = -\frac{Rhc}{\left(n + a\right)^2} \tag{25.3}$$

and that the spectral line of frequency given by Eq. (25.1) is emitted by the transition of the atom from the initial state of energy  $E_n = -\frac{Rhc}{(n + a)^2}$  to the final one of energy  $E_{n'} = -\frac{Rhc}{(n' + a')^2}$ .

The relationship (25.3), is somewhat more complicated than the corresponding one of the hydrogen atom,

$$E_n = -\frac{Rhc}{n^2}, \qquad (25.3')$$

but it is still simple enough. Especially in the Bergman series of alkali metals Rydberg's correction f is almost zero, and this shows that the

approximate energies of the initial states of this series are given by the values for the hydrogen atom, Eq. (25.3').

This circumstance is seemingly suggestive of the idea that although the actual atom of many electrons may be enormously complicated, the emission of the optical spectrum is due to a mechanism not much different from that in the hydrogen atom. Namely, only one of many electrons in the atom, the outermost electron, is considered to be responsible for emitting light. Following this idea, Bohr singled out this active electron among many others and reduced the complicated problem of the many electron system to a simple one body problem of this "optical electron". This particular electron is influenced not only by the Coulomb field of the nucleus but also by the electric field of the other electrons, the "core electrons"; in other words, it moves in the field of a central force different from the simple Coulomb force. When the orbit of this electron is extended far out from the nucleus, i.e., when the electron does not come near to the core of the atom, the motion of this electron will be quite the same as that in the case of the hydrogen atom and Eq. (25.3') can be applied. On the other hand, when the electron penetrates through the core part of the atom, the motion will be substantially different from the simple hydrogen case and, hence, Rydberg's correction in this case will become appreciable.

In the case of an arbitrary central force, the motion is still classified according to the values of the two quantum numbers n and k just as in § 20. The only difference is in the fact that here the energies of stationary states depend not only on n but on both n and k. The spectral terms are then represented by specifying two integers n and k as

$$\begin{array}{c}
 1_1, 2_1, 3_1, 4_1, 5_1, \dots \\
 2_2, 3_2, 4_2, 5_2, \dots \\
 3_3, 4_3, 5_3, \dots \\
 4_4, 5_4, \dots \\
 5_5, \dots \end{array}$$

where the large numbers and the small subscripts stand for n and k, respectively, and we also note that  $n \ge k$ . If we use the convention

of denoting k = 1, 2, 3, 4, ... by the symbols s, p, d, f, ..., respectively, we have

1s, 2s, 3s, 4s, 5s, ...
2p, 3p, 4p, 5p, ...
3d, 4d, 5d, ...
4f, 5f, ...

and we accordingly call the k = 1, 2, 3, 4, ... terms the s-, p-, d-, f-, ... terms, respectively. By a similar argument as in § 20 together with the assumption of a certain appropriate non-Coulombic central force, it can be shown that the terms can in fact be put approximately into the form of Eq. (25.3) and that Rydberg's correction *a* depends on *k* but only slightly on *n*. We designate these corrections for k = 1, 2, 3, 4, ... terms by s, p, d, f, ..., respectively.

The application of the selection rules resulting from the correspondence principle reveals that the transitions are forbidden for changes of k other than by  $\pm 1$  and, hence, that the s-term can be combined only with p-terms; and the p-terms, or the d-terms, only with s- and d-terms, or p- and f-terms, respectively. We thus obtain the series of spectral lines as given by Eq. (25.2) and Eq. (25.2').

It has been stated before that the Rydberg correction for the f-term is almost zero. This implies that in the motion corresponding to the f-term the electron hardly approaches the core part of the atom. The distance from the center of the lower apsis of the elliptic orbit in the case of the hydrogen atom is in fact given by

$$r_{\min} = \frac{h}{4\pi^2 m e^2} \left(n - \sqrt{n^2 - k^2}\right),$$

which, though small for small k-values, becomes appreciably large for large k-values of, say, 4 or more and the electron is then kept away from the core part.

In order to have a more realistic discussion, there is shown as an example in Fig. 43(a) the diagram indicating the s-, p-, d-, and f-terms of Na. On the vertical axis are plotted the values of the terms in units of  $\nu/c$  and for the sake of comparison the terms of hydrogen are also given in the figure. One can easily see the parallelism between Na and H

in the case of the d- and f-terms. The well-known D-line of Na, by the way, is emitted in the transition from the lowest p-term to the lowest s-term.

In assigning *n*-values to the experimentally observed spectral terms, certain precautions are necessary. Namely, the comparison of the experimental values with Eq. (25.2) allows the estimate of only the sum n + a and neither *n* nor *a* separately. If the Rydberg correction *a* can be calculated exactly by the theory, we can naturally subtract this from the (n + a) value to obtain the *n*-value. However, the theory itself is not of such an accurate nature and cannot be quantitatively trusted.

It is nevertheless certain from a rather general argument of the theory that Rydberg's corrections are always positive and that for large n they can be neglected. With these qualitative results in mind,



Fig. 43. (a) Energy levels of H and Na atoms. Numbers attached to each energy level are the values of the principal quantum number n.



one compares, in the large n region, the spectral terms under consideration with that of hydrogen, determines the n value of a certain term, and then extends the assignment to smaller n-values. Or, in some cases, one estimates the magnitude of Rydberg's correction from the size of the core, as already known from some other experiments. The comparison of the terms with those of the neighboring atom in the periodic table is also used advantageously to determine the *n*-value. The *n*-values thus estimated have been included in Fig. 43(a). Similar diagrams are shown for Li, K, Lb, Cs in Fig. 43(b), for Mg, Ca, Ba in

 TABLE 3

 pers of the optical electron in the normal state of value

guantum numbe		01	unc	υp	tical ci	cet	i on	in the	110	mai ste		01 (	ano	us a	toma
Atomic number	1	2	3	4	510	11	12	1218	19	2029	30	21	36	37	38
Atom	Н	He	e Li	Be	BNe	Na	. Mg	A1A	к	Ca…Cu	Zn	Ga	Kr	RЪ	Sr
k of the optical electron in the normal state	1	1	1	1	22	1	1	22	1	1 1	1	2	1	1	1
n of the optical electron in the normal state	1	1	2	2	22	3	3	33	4	4 4	4	4	5	5	5

Fig. 44, and for Al, Ga, In, Ti in Fig. 45. We now have determined the n- and k- values for the "optical electron" of various atoms. Of special importance for our later discussions are the n- and k- values of the



lowest energy state. These values of n and k determine the orbit of the "optical electron", which is the outermost electron. In Table 3 are shown the values of n and k for the ground states of various atoms.

There is one more important fact. Namely, as is clear from Figs. 43, 44, and 45, together with Table 3, there is a close relation between the periodicity of the elements and their atomic spectra. For instance, the alkali elements whose spectral terms are shown in Fig. 43(b) have some characteristics in common among themselves; the lowest term is the s-term and the next is the p- term, etc. The same is true also for the



Fig. 45. Energy levels of the elements of the third group.

elements of the second group (the alkali earths), and for those of the third group, in Fig. 44, and in Fig. 45, respectively. In fact, we notice that in the former group they all have two series of s, p, d, f,... terms and that in the latter group the lowest term is the p-term and the next one is the s-term. On the other hand, the structure of the term arrangement is quite different in the two groups. In conclusion, we can state that the chemical properties of an atom and its optical properties have their origin in common since the elements of the same group in the periodic table have

similar chemical properties as well as similar atomic spectra, while the elements of different groups do not have such similarities. Or, in other words, the atomic spectra of the elements follow the same periodic regularity as the chemical properties.

## § 26. The X-Ray Spectrum

In generating X-rays, we let cathode rays hit the anticathode in a X-ray tube. The anti-cathode then emits two kinds of X-rays, one which increases its penetrability with increasing cathode ray energy and the other which has the same wave length, characteristic of the anti-cathode material, and hence has the same penetrability irrespective of the cathode ray energy. The X-rays of the first kind have a continuous spectrum in a certain range of wave lengths and its enhanced penetrability with increased large cathode ray energy is due to the shortening of its average wave length. The latter kind, on the other hand, has a line spectrum and these are called the characteristic X-rays

of the material used as the anti-cathode. The X-rays which interest us in this section are of the latter kind, the characteristic X-rays.

When one studies the characteristic X-rays of various elements, one readily finds that the patterns of the spectral line configurations are the same while the wave lengths of the spectral lines becomes shorter with increasing atomic number of the elements. Thus the characteristic X-rays generally consist of a number of groups of line spectra, a K-group of the shortest wave length, an L-group of longer wave length, an M-group of still longer wave length, and so forth. For the elements of smaller atomic numbers, however, these spectra are of longer wave length and are not observed in the X-ray region. In fact, in elements lighter than Na the characteristic X-rays are not observable at all and it is first at Na that the K-group comes into the X-ray wave length region and is actually observed. With increasing atomic numbers, from Na on, the K-group shifts to a shorter wave length region and, at an atomic number of about 30, the L group comes into the X-ray region. The M- and N-groups come into play at still larger atomic numbers. In

Phot. VIII is shown the gradual shifting toward shorter wave lengths of the K-lines for elements from Z = 33 (As) to Z = 45 (Rh), while in Fig. 46 is given a graph showing the relationship between the frequency,  $\nu$ , of the K-group and the atomic number, Z, of various elements. It is to be noted that in Fig. 46 the square root of  $\nu$ , instead of  $\nu$  itself, is plotted on the ordinate and that the square



Fig. 46. Frequency of the KX-ray and the atomic numbers.

root of the frequency of the KX-ray is exactly proportional to the atomic number of the element. This relationship was obtained experimentally by Moseley.

This relationship between  $\nu$  and Z compels us to recognize the large difference between the X-ray spectrum of an element and its optical spectrum. Namely, the X-ray spectrum does not show any periodic variation but instead varies monotonically with the atomic number Z. The more important experimental fact to be noted is that the characteristic X-ray of an element is always the same irrespective of the chemical state of the element, i.e., in what compound the element is to be found. These experimental facts seem to suggest that the electron responsible for the emission of the characteristic X-rays is not the one which plays a role in the optical emission and that it must reside in the inner-core part of the atom which is scarcely affected by a change of chemical state. In fact, as we have seen just above, the quantum number n of the optical electron does not start with I except in the cases of hydrogen and helium. This can be explained by considering that the electrons in the core have already occupied these orbits, n = 1, 2, ..., near the center of the atom and that the "optical electron" cannot go into these orbits of smaller n's. If this is the case, the electrons in these inner orbits of the atom can hardly be affected by the other atom with which it combines chemically; their behavior would not change with the chemical state of the element. Furthermore, the electron in an inner orbit will feel the full strength of the nuclear charge Z and is hydrogen-like, however large Z may be, while the optical electron in the outermost orbit will feel the nuclear Coulomb field which is shielded by the presence of the core electrons. The frequency of the orbital motion of such an inner electron would then become quite large for large Z. Therefore, if we take the point of view that the characteristic X-rays are emitted by such inner electrons, the experimental facts that these X-rays have frequencies about 1000 times as large as those of optical emission lines and that the frequency of the X-ray increases with increasing Z can be understood quite naturally.

Before going into a quantitative discussion of this viewpoint, we need further qualitative considerations. According to Bohr's theory, the emission of a characteristic X-ray is also to be considered as due to a transition between two states. Denoting the energies of the initial and final states by  $W_1$  and  $W_2$ , respectively, the frequency of the emitted X-ray is given by

$$h \mathbf{v} = W_1 - W_2$$

so that the concept of "terms" needs to be introduced here also.

Due to the fact that the atom is an extremely complicated dynamical system of many electrons, it is impossible to calculate these terms accurately. The simplicity, however, of the law observed in experiment seems to suggest that, as in the case of optical emission, the number of electrons responsible for the X-ray emission is only one.

In investigating the nature of the transitions corresponding to X-ray emission, the strange fact is to be noticed that these characte-

ristic X-ray lines manifest themselves only in the emission spectra but never in the absorption spectra. The lines in the ordinary optical emission spectra generally appear also in the absorption spectra. For instance, when one observes spectroscopically the light of an arc through Na vapour, dark lines appear with frequencies corresponding to the D-lines of Na. This is simply because the light of the same frequencies as those of the D-lines have been absorbed by the Na atom from the continuous emission spectrum of the arc.

If a spectral line is to be due to a transition from some level of higher energy to the ground state, this line has to be observed as an absorption line in the continuous spectrum of light which passes through an aggregate of atoms in the ground state, since, if the emission line is due to the transition  $W_n \to W_1$ , the light of the same frequency in the continuous spectrum can obviously be absorbed leading to the inverse transition  $W_1 \to W_n$  in one of the atoms. While the optical lines, like the D-lines of Na, do have this property, the characteristic X-rays do not.

In order to understand this peculiar property of the characteristic X-rays, Kossel considered the following mechanism for the emission of these lines. That is, when the cathode ray hits the anti-cathode it will eventually knock off an electron from an atom in the anti-cathode, sometimes an electron in an outer orbit and sometimes one in an inner orbit in the core of the atom. This would give rise to an empty orbit in the atom and if this vacancy is produced in the core it is possible that an electron in some outer orbit makes a transition to this vacancy. This was the mechanism proposed by Kossel for the emission of characteristic X-rays. We here note that in this idea of Kossel the following assumption is tacitly made. Namely, there cannot exist an arbitrary number of electrons in an orbit and indeed there is an upper limit to this number for each orbit. The previously mentioned fact that the principal quantum number n of the optical electron generally does not start from one can be considered as supporting this assumption. For example, the fact, as revealed in Fig. 43, that the *n*-value of Na starts from 3 can be understood on the basis that the orbits of n = 1 and 2 are completely occupied up to the maximum possible numbers of electrons, thereby forcing the optical electron to be out in the orbit n = 3. We will come back to this point again in later sections.

In the normal state of an atom, all the orbits in the core are filled to their maximum capacity. When a vacancy is produced by the above mechanism in one of these orbits, it is then possible for an electron to make a transition from one of the outer orbits to this vacancy thereby emitting a characteristic X-ray line. We denote the inner orbit, where the vacancy was originally produced, and the outer orbit, from which the electron makes the transition, by 1 and 2, and their energies by  $W_1$  and  $W_2$ , respectively; the frequency of the emitted X-ray line is then given by

$$\nu = \frac{W_2 - W_1}{h}$$

With this mechanism, it is quite naturally understood why a line of the same frequency does not appear in the absorption spectrum. In order for such a line to appear in the absorption spectrum, the transition  $W_1 \rightarrow W_2$ , i.e. the inverse of the above  $W_2 \rightarrow W_1$ , has to occur in the atom. This, however, is impossible since the orbit 2, unless it is a far outer one, is in the normal state of the atom already occupied by the maximum allowed number of electrons and leaves no space for an extra electron, thereby preventing the occurence of the transition.

If we accept the assumption that each orbit of an atom has an upper limit for the number of electrons it can accomodate, the structure of an atom will be pictured as follows. Out of Z electrons associated with an atom a certain number will fill the innermost orbit, n = 1 and k = 1. When this orbit is filled to its maximum capacity, which happens to be two, the next electrons will have to be assigned to the orbit of next higher energy, i.e. the one of n = 2 and k = 1. When this second orbit also is filled up, the maximum number here being also two, the electrons have to go into the orbit of the third lowest energy, n = 2 and k = 2. The orbits are thus one after another filled in the order of increasing energy. The lower the energy, the nearer is the orbit to the center, and the higher, the further away from the center. Hence the electrons in the core part of an atom will form a number of shells each consisting of orbits filled to their maximum capacity. Such shells, totally filled with electrons and unable to accomodate any more, are called closed shells. A number of optical electrons will move outside these closed shells. This outermost orbit does not have to be filled to its maximum capacity. We shall come back later to this shell structure of the atom for more detailed discussions.

Coming back to the quantitative considerations we had meant to make in this section, we consider the problem of the characteristic X-ray

emission simplified to a one electron problem just as we did in dealing with the emission of the optical spectrum. The electron under consideration, being in an orbit well inside the core, will suffer the full strength of the nuclear charge, as we remarked earlier. The energy corresponding to this orbit is then given by changing the nuclear charge e to Ze in the calculation of § 20 and § 21. That is,

$$W_n = -\frac{Rhc\,Z^2}{n^2}$$

and, in particular, corresponding to the innermost orbit

$$W_1 = -\frac{RhcZ^2}{1^2}$$
(26.1)

and to the next orbit

$$W_2 = -\frac{Rhc\,Z^2}{2^2}\,.$$
 (26.2)

Since these energies are so normalized that the electron will have zero energy at infinity, their absolute values represent physically the amount of work necessary to remove from the atom the electron in the orbit.

If, following Kossel's picture, one of the core electrons, say one in the innermost orbit, is removed from an atom in the anti-cathode by the impact of the cathode rays, the atom will be left with an energy higher by an amount given by the absolute value of Eq. (26.1). If an electron in the next orbit makes a transition to this vacancy in the innermost orbit leaving thus a vacancy in the second orbit, the energy of the atom is now lowered by the difference of Eq. (26.1) and Eq. (26.2). The frequency of the X-ray emitted by this transition will, then, be given by

$$\frac{\nu}{c} = RZ^2 \left( \frac{1}{1^2} - \frac{1}{2^2} \right) = \frac{3}{4}RZ^2 .$$
 (26.3)

Should this formula be correct, the square root of the frequency,  $\sqrt{\nu}$ , would be proportional to the atomic number Z. This, in fact, is the relation discovered experimentally by Moseley and clearly scen in Fig. 46. The explanation of the two lines seen there, or in Fig. 45, will be given shortly. Eq. (26.3) is numerically correct for the X-rays of the K-group.

Similar arguments can be made also for the L-, M-, or N-groups; the results for the L-group are

$$\frac{\nu}{c} = RZ^2 \left( \frac{1}{2^2} - \frac{1}{3^2} \right) \tag{26.4}$$

and for the M-, and the N-group

$$\frac{\nu}{c} = RZ^2 \left(\frac{1}{3^2} - \frac{1}{4^2}\right),$$
 (26.5)

and

$$\frac{v}{c} = RZ^2 \left( \frac{1}{4^2} - \frac{1}{5^2} \right)$$
,

respectively.

Just as in the case of optical spectra, the discussion can be made much easier if we talk about the terms rather than about the spectral lines themselves. We hence call  $RZ^2/1^2$ ,  $RZ^2/2^2$ ,  $RZ^2/3^2$ , and  $RZ^2/4^2$ , the K-, L-, M- and N-term respectively. The K-term is then the energy of the atom when there is a vacancy in the orbit n=1 and k=1 and the L-term is the energy when it has a vacancy in the orbit n=2 and k=1, or n=2 and k=2. Consequently, there are two L-terms but only a single K-term. If the force acting on the electron is perfectly Coulombic, the double L-term will remain unresolved. Since, however, this is not actually the case, the term appears as a doublet. On the same basis, we can conclude that the M-term (n=3, k=3; n=3,k=2; n=3, k=1) and the N-term (n=4, k=4; n=4, k=3;n=4, k=2; n=4, k=1) are to appear as a triplet and a quartet, respectively.

When the terms are determined experimentally from the analysis of the X-ray spectra, serious deviations are noticed contrary to our expectation. Namely, while the K-term is a singlet as expected, the L-, M-, and N-terms appear as triplet, quintet, and septet, respectively. This is a very important fact suggesting that the quantum numbers nand k by themselves are not sufficient to specify the orbit of an electron. As will be discussed in a later section, the same situation arises also for the optical terms of the atom and with this fact as a clue one can arrive at the conclusion that an electron is not merely a simple point mass with three degrees of freedom.

Except for the necessity of introducing additional degrees of freedom, the regularity in the characteristic X-ray spectra can be well explained, quantitatively as well as qualitatively, by Kossel's idea. This in turn implies that the concept of the shell structure of the atom is reflecting the truth. In the next section, we will show that this shell structure of the atom is closely related to the periodicity of the elements. For convenience in such discussions, the shells are called K-, L-, M-, N-shells and so on in order, starting with the innermost shell, i.e. in the order n = 1, 2, 3, 4, and so on.

# § 27. The Shell Structure of the Atom and Periodicity

It was Kossel who first found and discussed in detail, in 1916, the relationship between the electronic shell structure of the atom, and its periodicity discovered earlier by Mendelejeff. He presupposes, in a vague sense, a shell-like structure of atomic electrons and assumes that each shell becomes particularly stable when it is occupied by a certain number of electrons making the shell closed. He further considers that electrons in the outermost unclosed shell are responsible for chemical binding. Then the periodicity of chemical properties suggest that similar configuration of the outermost shell appears repeatedly when we go from an element to another. He then used the following lines of argument. Namely, he first started with a group of inert gases, He, Ne, A, Kr, and X which lie at the extreme right of the periodic table as shown in Table 4. It is well known that these gases do not make any compounds and are hard to ionize. These facts suggest that the atoms of these gases have no electrons outside the stable closed shell. We next consider the halogens, F, Cl, Br, and I, which come, in the periodic table, just before the inert gases. These elements are known to easily form negative ions. Kossel considered this tendency as due to the following. The atoms of these elements which have electrons less in number by one than the following inert gas atoms, lack one electron to complete the last shell and they are consequently apt to pick up an additional electron in order to form stable shell configurations, similar to those of the corresponding inert gas, thereby becoming negative ions. On the contrary, the elements Li, Na, K, Rb, and Cs, coming next after the inert gases, easily form positive ions; i.c., they have an excess electron as compared with the neighbouring incrt gases and this electron outside the inert cores can easily be taken away leaving the atom as a positive ion.

Quite analogous arguments can be given for the valences of the other

The	The	The	The	The	The	The	The				T	ABLE 4			
first	group	group	fourth	fifth	sixth group	seventh	eighth			P	eriodic ta	ble of the	elemen	ts	
I H		8F	[ 8- ° 4P	Broup	8.04P	8.0 up	2 He	Ne tun	93 ep- ium	94 Pluto- nium	95 Ameri- cum	96 Curium	97 Berke	7 elium	98 Califor- nium
3 Li	4 Be	5 B	6 C	7 N	8 0	9 F	10 Ne	1	Np	Pu	Am	Cm	Bk	C .	Cf
11 Na	12 Mg	13 A1	14 Si	15 P	16 S	17 Cl	18 A								
19 K	20 Ca							21 Sc	22 Ti	23 V	24 Cr	25 Mn	26 Fe	27 Co	28 Ni
29 Cu	30 Zn	31 Ga	32 Ge	33 As	34 Se	35 Br	36 Kr								
37 КЪ	38 Sr							39 Y	40 Zr	41 Nb	42 Mo	43 Tc	44 Ru	45 Rh	46 Pd
47 Ag	48 Cd	49 In	50 Sn	51 Sb	52 Te	53 I	54 Xe								
55 Cs	56 Ba							57 La			The ra	re earth g	roup*		
									72 Hf	73 Ta	74 W	75 Re	76 Os	77 Ir	78 Pt
79 Au	80 Hg	81 T1	82 Pb	83 Bi	84 Po	85 At	86 Rn								
87 Fr	88 Ra							89 Ac	90 90	91 Pa	92 U				

					* T	he rare e	arth gro	up					
58	59	60	61	62	63	64	65	66	57	68	69	70	71
Ce	Pr	Nd	Pm	Sm	Eu	Gd	Tb	Dy	Ho	Er	Tm	Yb	Lu

сн. 4, § 27]

181

groups. The Be, Mg, Ca, St, and Ba, coming in the second column of the periodic table, for instance, have a tendency to become positive ions of valence two and this is because they each have two excess electrons outside the inert cores which are similar to those of the neighbouring inert gases.

One can readily notice that these lines of argument are quite the same as those given previously about the shell structure of the atom. In the terminology of shell structure, we can state the above results as follows. The atom of an inert gas has all its electrons in the form of closed shells. The alkali atom has an excess electron which has to be placed in a new orbit outside the closed shells. An atom of the next group of elements, the alkali carths, have one more excess electron and altogether two electrons move in this new orbit.

With increasing atomic number, the number of excess electrons in the outer orbit also increases and eventually this orbit, filled with electrons to its maximum capacity, becomes a closed shell and we here have another inert gas atom.

Investigating the periodic table from this point of view, we can determine the capacities of each shell. The first is hydrogen which has one electron. Hydrogen is apt to become a positive ion of unit valence and this is so because its electron does not complete a closed shell and can easily be taken away. The next element is helium the atom of which has two electrons. These two electrons must complete a closed shell since helium is one of the inert gases. In other words, the first shell is filled by two electrons. The atom of the next element, lithium, has three electrons, two of which fill up the shell just mentioned and the third of which goes into the next shell. This third electron can hence be easily removed from the atom therely making it a positive ion. The next atom beryllium has two excess electrons, apart from the two which have completed the first closed shell, and these can be removed rather easily. Beryllium has thus a tendency to become a positive ion of valence two. From Li, through Be, B, C, N, O, and F, the number of excess electrons in the outer shell increases steadily one by one till it becomes eight at neon which is the first inert gas since helium. This implies that the second shell becomes full with eight electrons. The capacity of the second shell being eight, the atom of Na, the element next to Ne, has to locate its last electron in the third shell and accordingly it tends to become a positive ion of unit valence by rejecting this excess electron. Furthermore we note that this atom, Na, is similar

to H and to Li in that they all have one electron in their outermost shell. The next atom, Mg, has two electrons in the third shell and, hence, is similar to Be. The number of electrons in the third shell increases through Al, Si, P, S, and Cl, till it becomes eight at A which is again an inert gas. The maximum capacity of the third shell is therefore eight.

The next period from A through K, Ca, Sc, Ti,... Fe, Co, Ni, Cu, Zn,..., to Kr is longer than the above two periods, He to Ne and Ne to A. The first inert gas after A is Kr and there are seventeen intervening elements between the two. This implies that the fourth shell becomes full with eighteen electrons. It should be noticed however, that Cu, and Zn, respectively, have properties somewhat similar to that of an alkali and alkali earth, suggesting that they have one, and two electrons outside a closed-shell-like configuration which is comparatively stable. The inert gas coming next is X and we can conclude that the capacity of the fifth shell is also eighteen. Just as in the case of Cu above, Ag also has properties similar to those of an alkali and we conclude that a relatively stable sub-shell is closed just before Ag.

These are the structures of atoms as pictured by Kossel. Combining the theory of optical spectra with these considerations of Kossel, Bohr was able further to clarify the shell structure of the atom.

According to Bohr's theory, the electron responsible for the emission of the optical spectrum is one of the rather loosely bound electrons in the outer orbit of the atom and the other electrons are considered to form a core part comparatively tightly bound around the center of the atom. In the meantime, we already know from spectroscopic studies, at least for atoms with simple spectra, the quantum numbers of the orbit in which this optical electron is bound. Therefore, by analysing those experimental data, we can determine the quantum numbers of the orbit corresponding to each shell in the atom. We proceed with simple examples.

H (Z = 1): In the case of the hydrogen atom, there is no problem since there is only one electron. In the ground state of the atom, this electron is in the orbit of the lowest possible energy, i.e. in the 1s orbit, n = 1, k = 1.

He (Z = 2): The helium atom contains two electrons and from spectroscopic studies we know that the term of the lowest energy is 1s, (see Table 3). This indicates that the optical electron is in the 1s orbit. The other electron, of energy not exceeding that of the optical electron, is believed to be also in this same 1s orbit since for hydrogen-like atoms the ls orbit is the one of the lowest energy. In view of the fact that helium is an inert gas, we conclude that the ls orbit will be closed by these two electrons. This means that the K-shell is closed by two ls electrons.

Li (Z = 3): In the case of lithium, as can be seen from Fig. 43(b) and Table 3, the optical electron in the ground state is in the 2s orbit. This fact supports the statement above that the 1s orbit is closed by two electrons: the first two electrons of the lithium atom being in the 1s orbit, the third electron cannot enter this orbit and has to go to the orbit of the next lowest energy, the 2s orbit. The existence of a single electron which is loosely bound outside the tightly bound closed K-shell, explains why the optical spectrum of lithium resembles that of hydrogen.

Be (Z = 4): From the analysis of its optical spectrum, it is shown that this atom has two 2s electrons outside the closed K-shell. That the last electron is in the 2s orbit is inferred from the fact that the lowest energy term in the spectrum of beryllium is a 2s-term; that the other of the last two electrons is believed to be also in the 2s orbit comes from the fact that in the spectrum of the positive beryllium ion, beryllium atom minus one electron, the lowest energy term is also 2s. The spectrum of the positive beryllium ion, by the way, can be obtained from a spark between beryllium electrodes since the existing strong electric field produces many ions, while in an arc with beryllium electrodes the spectrum of the beryllium ion does not show up strongly.

B (Z = 5): At the time of Bohr's theory, the spectrum of boron had not yet been analysed but studies of the spectrum of the carbon ion, having the same number of electrons as the neutral atom of boron, had yielded the conclusion that the lowest energy term is 2p, i.e. that the fifth electron of the boron atom is in the 2p orbit. This fact can be understood by considering that outside the innermost K-shell another closed shell is formed by the two electrons in the 2s orbit and hence the last electron has to go to the 2p orbit. Recalling the fact that the 1s orbit was closed also by two electrons at helium, one is tempted to expect that the s-orbits are closed always by two electrons and it will later be shown that this is actually the case. Despite the fact that the boron atom has one electron outside the inner closed shells, it does not have the properties of the alkalis. The reason is that there is not much difference in energy between 2s and 2p orbits, though the angular momentum and hence the shape of the orbits differ substantially from each other. In fact, for a central force field of precisely Coulomb type the two orbits have exactly the same energy. Therefore, there is no reason why the fifth electron in the 2p orbit should be more unstable, or in other words more easily removable from the atom, than the fourth, or the third electron in the 2s orbit. Boron is, in fact, known to have the property of easily becoming an ion of valence three.

The spectral terms of carbon, (Z = 6), nitrogen (Z = 7), oxygen (Z = 8), and fluorine (Z = 9), are so complicated that a more advanced theory is needed for their analysis. We simply remark here that the number of electrons in the 2p orbit increases steadily from two for carbon to six for neon (Z = 10). In view of the fact that neon is an inert gas, we conclude that the 2p orbit is closed by six electrons; thus all the orbits of n = 2, 2s and 2p, are closed at neon and the closed L-shell is formed by eight electrons, two in 2s and six in 2p.

Na (Z = 11): It is known from spectroscopy that the lowest energy term in the spectrum of sodium is 3s, (see Fig. 48 as well as Table 3). That is, the sodium atom has one 3s electron outside the closed K- and L- shells and hence shows a spectrum similar to that of the hydrogen atom.

Mg (Z = 12): Mg has two 3s electrons outside the closed K- and L-shells.

Al (Z = 13): Two of the three electrons outside the closed K- and L- shells fill the 3s orbit and the last electron goes to the 3p orbit. The argument for this conclusion is similar to that given in the case of boron. The 3s orbit is closed also by two electrons as was anticipated.

Through Si (Z = 14), P (Z = 15), S (Z = 16), Cl (Z = 17), the number of electrons in the 3p orbit increases and becomes six at A (Z = 18).

The fact that argon is an inert gas implies that a shell is completed here. Just as the 2p orbit closed with six electrons at neon, the 3p orbit also closes with six electrons at argon. We can thus anticipate that the p orbit in general is closed by six electrons.

K (Z = 19): From the analysis of its spectrum, it is known that the optical electron is in the 4s orbit; i.e. the last electron outside the closed 3p orbit goes to the 4s orbit.

This situation calls for a comment. Namely, why is it so that the last electron of the potassium atom goes to the 4s instead of the 3d orbit? If the force field were exactly of Coulomb type, the energy of the 3d orbit would be the same as that of the 3s or 3p orbit and smaller than

that of the 4s orbit. Hence, the experimental fact that the 4s orbit, instead of the 3d orbit, is chosen by the last electron should be interpreted as due to the non-Coulombic nature of the force acting on the optical electron, this reducing the energy of the 4s orbit below that of the 3d orbit.

Ca (Z = 20); The analysis of the spectrum indicates that Ca has two 4s electrons outside the 2p orbit which was closed at argon.

For elements from Sc (Z = 21), through Ti (Z = 22), V (Z = 23)..., to Ni (Z = 28), a more elaborate theory is necessary for the analysis of their complicated spectra. These elements do not show the chemical properties of the 3rd, 4th,..., groups corresponding to the number of electrons outside the closed shells. It is then concluded that the same argument given for the elements from carbon to neon, or from aluminium to argon, cannot be applied to these elements. We will not go into the details of the structures of these atoms, but merely quote the conclusions of Ladenburg, inferred from the fact that these elements have characteristic colors as well as magnetism. The essence of his conclusions is as follows.

At scandium, the last electron goes, instead of to the 4p orbit, to the 3d orbit, which has been neglected since potassium and at titanium one more electron goes to the 3d orbit. In Table 5 are shown the rather complicated behaviors of the last electrons of these elements.

Cu (Z = 29): The spectrum of copper is simple again. The optical electron is in the 4s orbit as it was in the case of potassium. Since the the number of electrons of the copper atom is larger by ten than that of the potassium atom, it is natural to conclude that the 3d orbit has been closed by these ten electrons. The atom of copper, then, has one electron in an s orbit outside the closed M-shell, n = 3, and thus behaves like an alkali atom.

Zn (Z = 30): Zn has two electrons in the 4s orbit as one expects in going from the first group to the second group of the periodic table.

Ga (Z = 31): The optical electron is in the 4p orbit, the 4s orbit being closed at Zn. This is also expected for the third group.

From Ge (Z = 32), through As (Z = 33), Se (Z = 34), Br (Z = 35), to Kr (Z = 36), the last electrons behave as in the corresponding cases between Si and A; i.e., the number of electrons in the 4p orbit steadily increases till the orbit is closed by six electrons at Kr, an inert gas. Here again the p orbit is closed by six electrons. We further note that the number of electrons in the 3d orbit remained at ten as we proceeded

#### TABLE 5

Number of the electrons in various orbits of the atoms

	1 s	s	2 P	s	3 p	d	s	p	4 d	f	s	5 P	d	s	6 p	d	7 s
1 H 2 He	1 2																
3 Li 4 Be 5 B 6 C 7 N 8 O 9 F 10 Ne	2 2 2 2 2 2 2 2 2 2 2 2	1 2 2 2 2 2 2 2	1 2 3 4 5 6														
<ul> <li>11 Na</li> <li>12 Mg</li> <li>13 Al</li> <li>14 Si</li> <li>15 P</li> <li>16 S</li> <li>17 Cl</li> <li>18 A</li> </ul>	2 2 2 2 2 2 2 2 2 2 2 2 2	2 2 2 2 2 2 2 2 2 2 2 2	6 6 6 6 6 6 6	1 2 2 2 2 2 2 2 2 2 2	1 2 3 4 5 6												
19 K 20 Ca 21 Sc 22 Ti 23 V 24 Cr 25 Mn 26 Fe 27 Co 28 Ni 29 Cu 30 Zn 31 Ga 32 Ge 33 As 34 Se 35 Br 36 Kr	2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2	2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2	6 6 6 6 6 6 6 6 6 6 6 6 6 6 6 6 6 6 6	2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2	$\begin{array}{c} 6 \\ 6 \\ 6 \\ 6 \\ 6 \\ 6 \\ 6 \\ 6 \\ 6 \\ 6 $	1 2 3 5 5 6 7 8 10 10 10 10 10 10 10 10	$ \begin{array}{c} 1\\2\\2\\2\\1\\2\\2\\2\\2\\2\\2\\2\\2\\2\\2\\2\\2\\2\\2\\2$	1 2 3 4 5 6									
37 Rb 38 Sr	2 2	2 2	6 6	2 2	6 6	10 10	2 2	6 6			1 2						

	1		2		3				4			5			6		7
	s	s	р	S	р	d	S	p	d	f	S	p	d	S	р	d	S
<ul> <li>39 Y</li> <li>40 Zr</li> <li>41 Nb</li> <li>42 Mo</li> <li>43 Tc</li> <li>44 Ru</li> <li>45 Rh</li> <li>46 Pd</li> <li>47 Ag</li> <li>48 Cd</li> <li>49 In</li> <li>50 Sn</li> <li>51 Sb</li> <li>52 Te</li> <li>53 I</li> <li>54 X</li> </ul>	2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2	2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2	6 6 6 6 6 6 6 6 6 6 6 6 6 6 6 6 6 6 6	3 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2	$\begin{array}{c} 6 \\ 6 \\ 6 \\ 6 \\ 6 \\ 6 \\ 6 \\ 6 \\ 6 \\ 6 $	10 10 10 10 10 10 10 10 10 10 10 10 10 1	2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2	6 6 6 6 6 6 6 6 6 6 6 6 6 6 6 6 6 6 6	1 2 3 4 5 7 8 10 10 10 10 10 10 10 10 10		2 2 2 2 2 2 2 2 1 1 1 1 2 2 2 2 2 2 2 2	1 2 3 4 5 6	2				
55 Cs 56 Ba 57 La 58 Ce  72 Hf 73 Ta 74 W 75 Re 76 Os 77 Ir 78 Pt 79 Au 80 Hg 81 Tl 82 Pb 83 Bi 84 Po 85 At 86 Rn	2222. $2222222222222222222222222222222$	2222 · · 222222222222222222222222222222	66666666666666666666666666	$ \begin{array}{c} 2 \\ 2 \\ 2 \\ 2 \\ 2 \\ 2 \\ 2 \\ 2 \\ 2 \\ 2 \\$	6666 · · 6666666666666666666	10 10 10 10 10 10 10 10 10 10 10 10 10 1	2222 · · 222222222222222222222222222222	6666···6666666666666666666666666666666	10 10 10 10 10 10 10 10 10 10 10 10 10 1	$1 \\ 14 \\ 14 \\ 14 \\ 14 \\ 14 \\ 14 \\ 14 \\ $	2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2	6 6 6 6 · · 6 6 6 6 6 6 6 6 6 6 6 6 6 6	1 1 2 3 4 5 6 7 8 10 10 10 10 10 10 10	$1 \\ 2 \\ 2 \\ 2 \\ 2 \\ 2 \\ 2 \\ 2 \\ 2 \\ 2 \\ $	1 2 3 4 5 6		
87 Fr 88 Ra 89 Ac 90 Th 91 Pa 92 U	2 2 2 2 2 2	2 2 2 2 2 2 2 2	6 6 6 6 6	2 2 2 2 2 2 2	6 6 6 6	10 10 10 10 10 10	2 2 2 2 2 2 2	6 6 6 6 6	10 10 10 10 10 10	14 14 14 14 14 14 14	2 2 2 2 2 2 2	6 6 6 6 6	10 10 10 10 10 10	2 2 2 2 2 2 2 2 2	6 6 6 6 6	1 2 3 5	1 2 2 2 2 1

from copper to krypton and, hence, that the 3d orbit can be considered as being closed by ten electrons.

For other elements we do not go into the details of the analysis but only give the results in Table 5, in which the number of electrons in a specified orbit is given for each atom. One of the conclusions drawn from this table which has an important bearing on later discussions is the maximum capacity, for electrons, of the various orbits. From the table, one can see that the maximum capacities for electrons of s p d and f orbits are 2, 6, 10, and 14, respectively. These figures seem to suggest the relation.

$$N = 4k - 2 \tag{27.1}$$

between the capacity, N, and the quantum number, k, of an orbit. The essential meaning of this relation should not at present be clear but will become so in later discussions.

The knowledge of the capacities of various orbits tells us in turn the capacities of the various shells which enter in the theory of X-rays. The K-shell consists of the orbit n = 1 and since k can only be 1 for n = 1 the capacity of the K-shell is 2. The L shell however, consisting of the orbits of n = 2, takes k-values of both 1 and 2 and, hence, its capacity is the sum of the capacities of s and p orbits, and is 2 + 6 = 8. By a similar consideration, one finds the capacity of the M-shell, and the N-shell, to be equal to 2 + 6 + 10 = 18, and 2 + 6 + 10 + 14 = 32, respectively. The periods of 2, 8, 18, and 32, appearing in the periodic table of the elements are closely related to the capacities of the shells given above.

# § 28. The Multiplet Structure of Spectral Terms and the Inner Quantum Number

The use of a high dispersion spectrometer reveals that the apparently single yellow line of the Na-spectrum, the well-known D-line, actually consists of two close lines. Similar precise investigations of the other lines of the Na spectrum further disclose the fact that the other lines are also of the same nature. These experimental results indicate that the spectral terms themselves are not single levels but that instead each of them consists of two closely adjacent levels. The s-terms are an exception to this rule and each consists of a single level. The situation is schematically shown in Fig. 47. The same is true also for the spectra of other alkali elements. Because of this property, all the spectral terms of the alkali-atoms are called doublets. Among these doublets however the s-terms show the single level structure as stated above. The reason why we call s-terms also "doublets" despite of the fact that they are in



Fig. 47. The

doublet structure

of the Na term.

reality all single will be seen later. For the sake of convenience in later discussions, we call the actually observed number of levels the "level number" and denote it by r. In this terminology, the above results can be stated as follows. Namely, the terms of the alkali atoms form doublets and among them the s-term has level number 1 while all the others, p-, d-, f-,..., terms have level number 2.

Turning now to the spectrum of magnesium, we find that there are two kinds of lines, one consisting of a single line and the other of three closely spaced lines. Analyzing the observed results by terms, we see that, while the s-terms are always

single, the p-, d-, f-,..., terms each consists of two kinds, one with a single level and the other with three close levels. It is further seen that single p, d, f... levels and triple p, d, f... levels never com-

bine, and thus we have non-combining groups of singlet and triplet terms. As for s levels, some of them combine only with singlet groups and some others only with triplet groups. Thus we have also two groups of s levels which may be called "singlet" and "triplet" groups respectively although they are all single. Then all the spectral terms are classified into two groups, singlet and triplet, where the singlet and triplet terms combine among themselves only and no combination occurs between the two groups.

More precisely, the last statement is true only in first approximation and a term of one kind does, under certain conditions, combine with that of the other kind with



Fig. 48. The singlet and triplet structures of the Mg term.

the emission of a very faint line, i.e., with a very small transition probability.

The situation that the level number of each singlet term is unity while that of a triplet term is one for the s-terms and three for the rest, is schematically shown in Fig. 48. The fact discussed in § 25, that the elements of the seond group in the periodic table have each two series of s-, p-, d-, f-,..., terms, corresponds to the existence of these two types of terms, the singlet terms and the triplet terms.

As an additional example we take aluminium. The spectral terms are here again divided into two groups. The first group consists of the sterms of level number one and the p-, d-, f-,..., terms of level number two, while the second group consists of the s-terms of level number one, the p-terms of three, the d-, f-,... terms of four. Transitions occur between the terms belonging to the same group and seldom between those belonging to different groups. We call the first group the doublet terms and the second the quartet terms.

From the above three examples, we find that the spectral terms are divided into separate groups which do not combine each other, that each of these groups has a respective upper limit for the level number, and furthermore that in general the highest level number is one for s-terms and is three for p-terms. Investigation of further examples tells us that the highest level number is five for d-terms, seven for f-terms, and so forth. Generalizing these results we conclude that the highest level number starts from 1 for s-terms and increases by 2 for each increase of the k-value, i.e., as we proceed in the order of s-, p-, d-, f-,... terms. We will then use the following terminology in classifying these groups of terms. Namely, we call the group of terms which has a highest level number u the u-fold terms and call this number u the multiplicity of this group. In Table 6 are shown the level numbers of s-, p-, d-,..., terms belonging to the groups of various multiplicities. The multiplicities up to 8 have actually been observed.

The fact that each spectral term is not a single term but instead, as shown above, consists of a number of close levels forces us to conclude that the principal quantum number n and the subsidiary quantum number k, or the classification into s-, p-, d-,..., terms, are not sufficient for a complete specification of terms. Therefore, we introduce a fourth quantum number j for the purpose of labelling each of the close-lying levels of the same n and k. The third quantum number, by the way, is the previously introduced magnetic quantum number m. The levels are then denoted, for instance, by  $5s_3$  or  $5p_4$ , which means that the level has the principal quantum number n = 5, the subsidiary quantum number, k = 1, and the fourth quantum number j = 3, or n = 5, k = 2, and j = 4, respectively. The actual convention of assigning *j*-quantum numbers will be explained in the next section. In the

#### TABLE 6

Number of the levels in the term s, p, d, ... for term except the svarious multiplicities

Multiplicit	y u =	1	2	3	4	5	6	7	8	
		Nı	umb	er c	of th	ie le	vels	; r =	_	
Term	s	1	ł	1	1	1	1	1	1	
	р	1	2	3	3	3	3	3	3	
	đ	1	2	3	4	5	5	5	5	
	f	1	2	3	4	5	6	7	7	
-	g	1	2	3	4	5	6	7	8	
	:	:	:	:	:	:	:	:	:	

above example of the sodium spectrum, each term except the sterms consisted of two close levels and hence the spectral lines resulting from the combinations of these terms are expected to consist of four closely spaced lines. The actually observed spectral lines

however, consist of at most two close lines. This implies that certain transitions are forbidden. It has been verified experimentally that the selection rule yielding this observed fact is that the change of the j-quantum number (when its numerical values are properly assigned) can only be zero, or  $\pm 1$ . We call this fourth quantum number j the "inner quantum number".

# § 29. The Spin of the Electron and the Intrinsic Magnetic Moment

In the last section, we have explained the experimental facts concerning the multiplet structure of spectral terms. We now ask what it is which brings about such multiplet structures. This question was finally answered by introducing the hypothesis that the electron is not a mere mass point having a certain mass value and an electric charge but that it also has an intrinsic angular momentum and an intrinsic magnetic moment. At the same time some flaws in Bohr's theory were brought to light and had to be remedied. Before arriving at such an ultimate conclusion, various explanations had been proposed by a number of physicists. One of the principal ones is due to Sommerfeld. He sought an explanation of the new experimental facts in the theory of relativity. In the theory described so far, the relativistic increase of the electron mass with its velocity has not been taken into account and Sommerfeld at first conjectured that the introduction of this relativistic effect might yield the multiplet structure as explained in the last section. This however did not turn out to be so and Sommerfeld and Landé proposed the following mechanism as a second alternative.

As stated before, an atom consists of an inert core and the so-called optical electrons which are responsible for the emission of the optical spectrum of the atom. We now assume that this inert core has an angular momentum which is denoted by s in units of  $h/2\pi$ . Furthermore, the orbital angular momentum of the optical electron, also in units of  $h/2\pi$ , is denoted by *I*. When the core of the atom has the angular momentum s, it will probably have a magnetic moment parallel to this angular momentum. This magnetic moment would then produce an axially symmetric magnetic field with its symmetry axis in the direction of the angular momentum. If this is actually the case, the optical electron moving in such an axially symmetric magnetic field would have its angular momentum I quantized in the direction of this symmetry axis; i.e., only a set of discrete angles would be allowed for the direction of I with respect to s. The interaction energy between the magnetic field due to the magnetic moment of the core on the one hand and the magnetic moment due to the orbital motion of the optical electron on the other would then assume a different value for each different direction. It would then be possible for the orbits of the same I to have slightly different energy values depending on the angles they have with respect to the direction of s, thus giving rise to the possibility of having a number of close levels for one and the same I value. This is simply the multiplet structure. In short, the idea is to make use of a kind of internal Zeeman effect in the magnetic field due to the atomic core for the explanation of the multiplet structure.

The angular momentum s of the atomic core is in this picture combined vectorially with the angular momentum l of the optical electron to give the total angular momentum of the atom as a whole, j; j = s + l in units of  $h/2\pi$ . The manner in which the direction of l is quantized with respect to that of s is to be such as to give discrete values to the magnitude j of j with a separation of unity in accordance with the general properties of angular momentum. These values of j can thus be used to label each of the close levels of a multiplet term. In fact, when, with the help of the correspondence principle, the selection rule is derived on the basis of this idea, it can be shown that in the allowed transitions the change of the j-value can only be 0, or  $\pm 1$ . Thus j is the fourth quantum number introduced before.

Let us now determine the number of possible *j*-values for a given s and

1. The maximum value of j is naturally the one which it has when l is aligned in the same direction as s and the minimum value corresponds to the l direction opposite to that of s. That is, the maximum and the minimum values of j are l + s and |l - s|, respectively, where l = |l| and s = |s|:

$$|l - s| \le j \le l + s . \tag{29.1}$$

The *j*-values are spaced with unit spacing between these two limits.

The number of j-values satisfying this condition is simply the level number r of the terms with given l and s. Accordingly, r is given by

$$r = 2l + 1 \text{ when } l \leq s,$$

$$r = 2s + 1 \text{ when } l \geq s.$$
(29.2)

From Eq. (29.2), we see that when l increases, the level number r also increases as long as  $l \leq s$ . For instance, for l = 0, 1, 2, ..., r is 1, 3, 5,... respectively, and r reaches the maximum value of (2s + 1) for l = s. When r reaches this maximum value a further increase in l does not lead to larger values of r.

We now compare these results with the experimental facts explained in the last section. According to the conclusions there, the *r*-value is 1, 3, 5,... for s-, p-, d-,... terms and when it reaches the value u it does not increase any further. Therefore, in order for both results to agree we have to require, in the first place,

$$l = k - 1 \tag{29.3}$$

and furthermore,

$$u = 2s + 1$$
. (29.4)

These two conclusions are remarkable in the following sense. Although the experimental facts can apparently be explained on the basis of the Landé-Sommerfeld interpretation together with the requirements of Eq. (29.3) and Eq (29.4), this explanation, if true, forces us to conclude: (1) that, from Eq. (29.3), the angular momentum of the optical electron corresponding to s-, p-, d-,... terms is 0, 1, 2,..., in units of  $h/2\pi$ , respectively; and (2) that the angular momentum of the atomic core assumes a half-odd integral value, i.e. an integer plus  $\frac{1}{2}$ , for an even value of u while it has an integral value for an odd value of u.

According to Bohr's theory, however, the magnitude of the angular

momentum of an electron's orbital motion is, in units of  $h/2\pi$ , the value of the subsidiary quantum number k itself and hence should be 1, 2, 3,... for s-, p-, d-,..., terms, respectively. The above conclusions, therefore, disagree by one unit with the results of Bohr's theory for the assignment of the angular momentum of the optical electron. Furthermore, the appearance of half-odd-integral values is, within the framework of Bohr's theory, entirely beyond comprehension even for the angular momentum of the atomic core.

In this connection, we recall the Stern-Gerlach experiment of § 22, which showed the azimuthal quantization into two directions of silver atoms in a magnetic field. Now, the ground state of a silver atom is known from spectroscopy and is an s-term belonging to a doublet group, i.e. u = 2. The angular momentum, s, of the atomic core is thus, from Eq. (29.4), equal to  $\frac{1}{2}$ . Since it is an s-term, we have also |I| = 0. The total angular momentum, |j| or |I + s|, of the atom is hence  $\frac{1}{2}$ . Therefore, on introducing this atom in a magnetic field the total angular momentum will have a component of either  $\frac{1}{2}$  or  $-\frac{1}{2}$  along the direction of the magnetic field; in other words, the silver atom in the magnetic field is azimuthally quantized in two directions. The result of the Stern-Gerlach experiment is quite naturally anticipated from this point of view.

On the contrary, Bohr's theory assigns a unit value of the angular momentum for s-terms and hence the silver atom is to be azimuthally quantized into three directions in the magnetic field, even if we assume zero angular momentum for the atomic core. This is in contradiction with the experimental finding. The assumption of a non-zero angular momentum for the atomic core leads to a level number of the ground state term larger than unity. This is also in disagreement with the known experimental result. By these arguments, we come to the conclusion, one which is in serious contradiction with the original version of Bohr's theory, that the appearance of half-odd-integral angular momentum and the assignment l, instead of k, to the orbital angular momentum of the optical electron are necessary for explaining the experimental facts.

Furthermore, once this idea is accepted, proper use of the correspondence principle on this basis can explain successfully the regularities of the spectral lines of almost all the atoms, while with the original Bohr theory there are disagreements with various experimental facts.

However, there still exist some unsatisfactory aspects of the Landé-

Sommerfeld interpretation; if this interpretation is correct, the singlet term and the triplet term of an alkali atom, for instance, correspond to different states of the atomic core. Namely, u = 1, and hence s = 0, for the former and u = 3, and hence s = 1, for the latter terms. If the state of the core is actually different as in this example, the energy of the core would also be different for the singlet group and the triplet group. The separation energy of the optical electron, or in other words, the value of the term in the limit of infinitely large quantum number n since this is simply the energy of the core itself, would be different for these two groups.

Contrary to this expectation experiment shows that the values, for  $n = \infty$ , are the same for the two groups, as shown in Fig. 48. Whether or not the term limits of the singlet and the triplet group are the same can be determined by a study of the faint spectral lines emitted by combinations between singlet and triplet terms.

This seems to suggest that it is inadequate to interpret the angular momentum s as the angular momentum of the atomic core. Uhlenbeck and Goudsmit remarked that the difficulty could be avoided by considering the electron not as a simple mass point, as has so far been assumed, but instead as having a certain intrinsic angular momentum. According to them, taking the simplest case of an alkali atom for instance, the atomic core does not have any angular momentum and the necessary angular momentum s with magnitude  $\frac{1}{2}$  is to be ascribed, as an intrinsic property, to the optical electron moving outside the atomic core. Since the fact that the optical electron is moving around the core can be interpreted, when observed from the electron, as meaning that the core is moving around the electron, the rotation of the electric charge of the core exerts on the electron a magnetic field in the direction of the angular momentum *l*. The intrinsic angular momentum s of the electron is then azimuthally quantized in the direction of this magnetic field thus yielding the doublet structure of the term with a given I. A similar line of argument can successfully be made in explaining the multiplet structure of the spectral terms of more complex atoms provided that the sum of the intrinsic angular momenta of the electrons outside the core is taken to be s.

Goudsmit and Uhlenbeck went further and interpreted this intrinsic angular momentum as due to the self-rotation of the electron, that is to a rotation about an internal axis. This self-rotation is called the spin of the electron.

The self-rotation of an electrically charged object necessarily gives rise to a magnetic moment. The magnitude of the intrinsic magnetic moment of the electron is one of the factors which determine the magnitude of the level spacing within a multiplet group. It furthermore plays a decisive role in the Zeeman effect observed when the atom is placed in a magnetic field. In the simplest case of an alkali atom in a magnetic field, the change of the s-terms will be as follows. Namely, since the angular momentum *l* of the electron's orbital motion is zero for the s-terms and since the atomic core does not carry any angular momentum the total angular momentum of the atom is the intrinsic angular momentum itself, i.e., the spin, of the electron. Since s in this case is  $\frac{1}{2}$  in magnitude, this angular momentum in the magnetic field is azimuthally quantized in two directions; one is the direction of the field and the other is the opposite direction. The fact that even an angular momentum of a half-odd-integral value has, when azimuthally quantized, its z-components separated from one to the next by one unit has already been confirmed by the Stern-Gerlach experiment.

If we denote the intrinsic magnetic moment of the electron by  $\mu_e$  the energy of the former state is  $\mu_e H$  and that of the latter state is  $-\mu_e H$ , where H is the strength of the applied magnetic field. This result indicates that the experimental determination of these line shifts due to the known magnetic field gives the value of  $\mu_e$  and, in fact, it was shown by this method that

$$\mu_e = \frac{eh}{4\pi mc} \,. \tag{29.5}$$

Recalling that the magnitude of the intrinsic magnetic moment of electron is  $\frac{1}{2} h/2\pi$ , we obtain for the ratio of the magnetic moment to the angular momentum

$$\frac{\mu_{\rm c}}{\frac{h}{2\pi} \cdot |{\bf s}|} = \frac{e}{mc} \,. \tag{29.6}$$

On the other hand, as was stated in § 18, the magnetic moment resulting from the orbital motion of electron is

$$\mu = \beta |I| = \frac{eh}{4\pi mc} |I|$$

and, hence, its ratio to the angular momentum

$$\frac{\mu}{\frac{h}{2\pi} \cdot |I|} = \frac{e}{2mc}.$$
(29.7)

Due to the difference between Eq. (29.6) and Eq. (29.7) the Zeeman effect of an alkali metal, or more generally with many of the other elements, differs from the prediction of the earlier theory. On the other hand, the singlet groups of the alkali earths, do show the Zeeman effect expected from the simple theory since in this case one may take s = 0 so that the magnetic moment is purely of orbital origin. In the former case, in which the earlier theory does not hold, we have the "anomalous Zeeman effect" while in the latter, where the earlier theory is valid, we have the "normal Zeeman effect". It is to be remembered that the source of the anomalous Zeeman effect is the fact that the ratio of the magnetic moment to the angular momentum is different, as shown in Eq. (29.6) and Eq. (29.7), for the intrinsic spin of electron and its orbital motion.

We do not intend to go into the detailed description of the anomalous Zeeman effect since it would be beyond the scope of this volume, but we satisfy ourselves by simply showing into how many levels each of the terms of an alkali atom splits under the influence of a magnetic field.

We begin with the simplest case of an s-term in which the orbital angular momentum l is zero. The total angular momentum, j, is then entirely due to the intrinsic angular momentum s of the electron, so that

$$|j| = \frac{1}{2}$$
,

implying that the s-term is split into two levels.

For p-terms |I| = 1, and **s** is azimuthally quantized with respect to this *I* so that it either orients itself parallel or anti-parallel to *I* i.e., |j| = |I| + |s| or |j| = |I| - |s|. The two previously mentioned closely adjacent terms of the doublet group are due to this effect. In the former, the parallel case

$$|j| = \frac{3}{2}$$

and in the latter, the anti-parallel case

 $|j| = \frac{1}{2}.$ 

Therefore, in the notation introduced in § 28, the terms of the former case should be denoted by

P3,

and those of the latter case by

 $p_{\frac{1}{2}}$ .

So far the effect of the magnetic field has not been considered. When a magnetic field in the direction of the z-axis is introduced, the total angular momentum j is azimuthally quantized with respect to this direction. Accordingly, the z-component of  $|j| = \frac{3}{2}$  assumes one of the four values of  $j_z = \frac{3}{2}, \frac{1}{2}, \dots, \frac{1}{2}$  and  $\dots, \frac{3}{2}$ , and the term  $p_{\frac{1}{2}}$  splits into four levels in the presence of the magnetic field. In the same way, the term  $p_{\frac{1}{2}}$  splits into two levels. We have in this argument used the fact, confirmed by the Stern-Gerlach experiment, that even the angular momenta of half-odd-integral values have their z-components, when azimuthally quantized, equally spaced by one unit. In the next case of the d-term we have |l| = 2, and, hence, |j| is either  $\frac{5}{2}$  or  $\frac{3}{2}$ . The doublet, therefore, consists of

and

The application of a magnetic field leads to the splitting of  $d_{\frac{1}{2}}$  into 6 levels and  $d_{\frac{1}{2}}$  into 4 levels.

We summarize the results in Table 7 the content of which, by the

TABLE 7

Number of the levels of alkali in a magnetic field

Т	erm	Num   the l	ber of levels
s	Sł	2	2
р	{ <del>P</del> <del>1</del>   <del>P</del> <del>1</del>	24	} 6
d	(d3	4	10
f		6	14

way, has been thoroughly verified by experiments. The figures in the last column of Table 7, i.e., 2, 6, 10, and 14, are the number of different orbital motions and spin directions involved in the s-, p-, d-, and f-terms, respectively, each of which was roughly considered previously as a single term.

The concept of the spinning electron, though in a sense attractive, seems somewhat arbitrary and not free from criticism. As a matter of fact, it is said that Goudsmit and Uhlenbeck sent a telegram to the editor of "Nature" asking to withdraw their paper, because their

idea had caused a good deal of criticism; if the magnetic moment of the electron is due to its self-rotation the radius of the electron should be at

least as large as  $h/mc = 10^{-10}$  cm in order that the magnetic moment  $e/4\pi mc$  be produced by the rotation of the angular momentum  $\frac{1}{2}h/2\pi$ . Such a large size for the electron (about the size of the K-shell of an atom of intermediate Z) would certainly destroy all the good agreements which we have seen between theory and experiment. Unfortunately or fortunately, the telegram was delivered too late, after the paper had been already printed. One had to wait for some years for the true explanation of this peculiar nature of the electron; the phenomenon of the electron spin is not something like an ordinary self-rotation but instead it is deeply rooted in the fundamental structure of nature. The discrepancies between theory and experiment observed at the end of § 26 in the multiplicities of the L-, M-, and N-terms of the electron spin, together with the assignment of l instead of k to the angular momentum.

Another remark is appropriate. Namely, the fact that the angular momentum of the electron's orbital motion is not k but instead l = k - 1 clearly indicates that we must not take literally the content of Bohr's Hypothesis III. The fact that the s-orbit has an angular momentum zero is required by the experimental results for the multiplet terms and is also supported by the Stern-Gerlach experiment. But, if the electron were to obey classical mechanics in its atomic motions, as Hypothesis III says it should, the state of zero angular momentum would be impossible since in this state the electron would collide with the atomic nucleus.

## § 30. Pauli's Principle

In § 27 we discussed the shell structure of the atom and the periodic table. One of the important conclusions there was that the s orbit of the atom was completely filled by two electrons and the p, d, f orbits, by 6, 10, and 14 electrons, respectively. Comparing these results with the content of Table 7 we come to a remarkable conclusion. Namely, the maximum capacities of s, p, d, f,..., orbits are exactly the same as the numbers of levels, in the presence of a magnetic field, of the s-, p-, d-, f-,..., terms respectively, of an alkali atom. This fact was first pointed out by Stoner in 1926. As has been stated in the discussion of Table 7, the numbers 2, 6, 10, and 14 in the Table are the respective numbers of different orbital motions and spin directions involved in the
s-, p-, d-, and f-terms of an alkali atom. The observed fact that these numbers coincide with the maximum capacities of the s, p, d, and f orbits in the atom seems to suggest the following. That is, when we start from the rough classification in terms of s, p, d, f,..., orbits and sub-classify them into 2, 6, 10, 14, ... different orbits respectively, only one electron is allowed for each of these sub-divided orbits. This sub-division, by the way, is not merely a conceptual one but is actually observed experimentally by the application of a magnetic field as was explained in the last section. According to the discussion there, each of these sub-classified orbits is labelled by giving the four quantum numbers: the principal quantum number n, the subsidiary quantum number  $j_z$ . The quantity j here, as in the Sommerfeld-Landé theory, is the magnitude of the total angular momentum j and  $j_z$  is its z-component.

In general,

$$j = \frac{1}{2}$$
 for  $l = 0$  , (30.1)

$$= l + \frac{1}{2}$$
 or  $j = l - \frac{1}{2}$  for  $l > 0$ , (30.1,)

and

from which one can conclude that the number, N, of realizable levels for a given 
$$l$$
 is given by

$$N = 2(l + \frac{1}{2}) + 1 + 2(l - \frac{1}{2}) + 1$$
  
= 4l + 2 (30.3)  
= 4k - 2.

This is identical to Eq. (27.1)

i

Therefore, the shell structure of the atom seems to be quite naturally explained if we make the hypothesis that "only one electron is allowed for each of the sub-classified orbits". Furthermore, experiments have shown that this hypothesis is indispensable for understanding the spectra of more complex atoms. The hypothesis turns out to be applicable not only to the electrons in the atom but to more general systems of electrons in which quantum theoretical orbits can be considered. The generalization of Stoner's finding and its application to various problems was carried out by Pauli in 1925 and the name Pauli's principle has been given to it. What then prevents the existence of more than one electron in one and the same orbit? Could it be the electric repulsion between the electrons? The answer is no. It is not due to such a dynamical effect but instead, like the spin of the electron, it is a phenomenon resulting from the fundamental structure of nature.

Complete understanding of the implications of this phenomenon, however, came much later.

### CHAPTER 5

## THE BIRTH OF MATRIX MECHANICS

## § 31. The Clue for Solving the Difficulties

In the preceding chapter we have seen how the quantum theory was developed to yield fruitful results despite its inherent inconsistencies. The guiding role in this was played by the correspondence principle of Bohr. As has been stressed in § 23, this principle not only served as a tentative working theory but also played the role of indicating to Bohr and to the physicists of his school the correct direction of advancement. It was, in fact, by the aid of this principle that these people could eventually figure out, from the rather distorted picture given by the old quantum theory, the correct form of the physical world. It is certainly a remarkable fact that the results conjectured by them were so close to the truth that most of the results could remain true without any revision even after the correct new theory was finally discovered. However, the approach used in obtaining sensible conclusions from the available distorted picture was still of such an intuitive nature and of such an intricacy that it did not follow any certain definite methodology as it should in a precise science. This unsatisfactory state of affairs came to an end when a young German physicist Heisenberg discovered in 1925 the recipe for converting in a systematic way the distorted picture into the correct one. By the aid of this recipe, Heisenberg constructed from the old quantum theory the new mathematical formalism of the so called matrix mechanics and by this theory he obtained a number of important results. For the thorough understanding of the physical implications of this mathematical formalism, however, one had to wait for aid given by a different method of approach. This came from the so called wave mechanics.

Outside the current in which Bohr and other physicists of his school were progressing led by the correspondence principle, a French physicist de Broglie was quietly going his own way and carefully studying the dual nature, particle and wave, observed in light. Through a far-reaching insight he noticed that an inherent inter-relation does exist between this duality of light and Bohr's quantum condition, and he developed the ambitious idea that this duality must exist not only in light but also in matter such as an electron and that, hence, an electron which was generally believed to be of a particle nature must have also a wave nature. He then proposed that the discontinuity appearing in the quantum theory might be due to a kind of interference of this wave. This happened in the year 1923.

This idea of de Broglie, though unnoticed for a while, was developed mathematically by Schroedinger in Switzerland. He found the fundamental wave equation for the electron wave and applied it to various real problems, and thus successfully obtained results in very good agreement with experiment. Thus he could derive the Balmer formula from the theory of the hydrogen atom based on this point of view. This was in the year 1925. The wave nature of the electron was also verified experimentally by Davison and Germer of the United States.

These two lines of theoretical approach, Heisenberg's theory based on Bohr's correspondence principle on the one hand and Schroedinger's theory of wave mechanics originating from the wave nature of electron proposed by de Broglie on the other, were apparently quite different in their basic elements. Despite the numerous conflicting arguments concerning theoretical interpretation aroused frequently by the apparent difference between the two theories, they gave identical answers to many real problems. It was later proved that, except for the difference in the interpretation, the two theories were in fact mathematically equivalent.

Thus we could grasp, at least mathematically, a fragment of the correct quantum mechanics and the clue for solving the various difficulties mentioned in the last chapter of the old quantum theory, although there still existed a number of ambiguities concerning its interpretation. Leaving the problem of interpreting the theory for later chapters, we shall in this chapter explain the physical ideas and the mathematical structure of Heisenberg's theory.

# § 32. Heisenberg's Discovery

According to the correspondence principle, there is a close correspondence between the quantum transitions of an atomic system on the one hand and the Fourier components of the electron's orbital motion in the sense of classical mechanics on the other. That is, for instance, to the transition  $n \rightarrow n - \tau$  of the quantum theory there corresponds the  $\tau$ -th higher harmonic of the motion of the electron in the state n. When the quantum number n is very large, the light resulting from the transition  $n \rightarrow n - \tau$  has the same frequency as that of this higher harmonic and, furthermore, they have the same intensity as well as the same polarization.

Although a number of considerations were successfully guided by this principle, the principle itself gives correct answers only in these cases of very large values of the quantum number n. Should it be possible at all to obtain a theory which can give correct answers also in the cases of small n-values? In the cases of very large n, this correspondence principle could act as the code-book for translating the results obtained from the classical considerations such as the electron's orbital motion and the emission of light waves in accordance with Maxwell's equations, into their quantum theoretical counterparts, the transition between the stationary states, the emission of light by the transition, and the transition probability. But this code-book is still unsatisfactory and often gives wrong, or ambiguous, conclusions in the cases of small n. The problem now is to find the code-book which can be used also in cases of small n.

Heisenberg proceeded like an expert decoder who reads a cryptogram. His approach was as follows. He started with the expectation that there must be a "certain quantity" of quantum theoretical nature which corresponds to the Fourier component of an electron motion in the classical theory and the absolute square of which should play a fundamental role in obtaining the transition probabilities. Since the correspondence between the classical quantities and the quantum theoretical ones is given through the appropriate terms of Fourier series, the first thing to be done is to express the problems, for instance, of handling the equation of motion or of introducing the quantum condition, as relationships between Fourier terms. We then have a cryptogram written in the language of Fourier terms and this is to be translated into the correct quantum language. If in this cryptogram appears the Fourier component of the coordinate  $x_i$ 

$$X(n, \tau) \exp \{ 2\pi i \nu(n, \tau) t \}, \qquad (32.1)$$

pertaining to the *n*-th stationary state and the  $\tau$ -th harmonic, we should expect that, corresponding to this Fourier component we shall

have a certain quantity associated with the transition  $n \rightarrow n - \tau$ . This quantity may be called the "transition component" of the quantum theoretical coordinate. Assume now the existence of this transition component and let it be denoted by

$$X(n; n - \tau) \exp \{ 2\pi i \nu(n; n - \tau) t \}.$$
 (32.1')

We replace the Fourier component by this hypothetical quantity. This is the first fundamental key to decode the cryptogram. We note here that, for the classical frequencies,

$$\nu(n,\tau) = -\nu(n,-\tau) \qquad (32.2)$$

due to the fact that  $v(n, \tau) = \tau v(n, 1)$ , while for the quantum theoretical frequencies, that

$$\nu(n; n-\tau) = -\nu(n-\tau; n) \qquad (32.2')$$

because of the relation  $\nu(n; n - \tau) = (W_n - W_{n-\tau})/h$ . This fact implies that the correct translation of Eq. (32.2) into the quantum theoretical language is not a simple change of sign of  $\tau$ , such as  $\nu(n; n - \tau) = -\nu(n; n + \tau)$ , but rather the interchange of the initial and final states as is apparent in Eq. (32.2'). Furthermore, for the coordinate x to be real, it is required in the classical theory that in Eq. (32.1)

$$X(n,\tau) = X^{\bullet}(n,-\tau) \qquad (32.3)$$

should hold; the correct translation of this relationship is again the interchange of the initial and final states:

$$X(n; n-\tau) = X^{\bullet}(n-\tau; n) . \qquad (32.3')$$

The reason why  $X^*(n - \tau; n)$ , instead of the simple change of sign of  $\tau$  (i.e.  $X^*(n; n + \tau)$ ) appears in Eq. (32.3'), is naturally due to the fact that Eq. (32.2') holds for the quantum theoretical frequencies and, hence, exp {  $2\pi i \nu(n - \tau; n)t$  } rather than exp {  $2\pi i \nu(n; n + \tau)t$  } is the complex conjugate of exp {  $2\pi i \nu(n; n - \tau)t$  }.

We now proceed to the next step of the decoding. We have shown previously, in Eq. (23.19), that the classical frequency  $\nu(n, \tau)$  in (32.1) satisfies

$$\mathbf{r}(n,\tau) = \frac{\tau}{h} \frac{\partial W(n)}{\partial n} \,. \tag{32.4}$$

206

сн. 5, § 32]

However, we should use

$$\nu(n; n-\tau) = \frac{1}{h} \left( W(n) - W(n-\tau) \right) \tag{32.4'}$$

for the quantum frequency in (32.1'). We now set the rule that whenever  $\nu(n, \tau)$  appears in the cryptogram we replace it by  $\nu(n; n - \tau)$ . This rule of replacing the classical frequency,  $\nu(n, \tau)$ , by the quantum theoretical one,  $\nu(n; n - \tau)$ , can also be considered as replacing the differential

$$\tau \frac{\partial W(n)}{\partial n} \tag{32.5}$$

by the difference

$$W(n) - W(n - \tau)$$
. (32.5')

Or, generalizing this rule to the case of an arbitrary function F(n) defined for each stationary state, the differential

$$\tau \, \frac{\partial F(n)}{\partial n} \tag{32.6}$$

is to be replaced by the difference

$$F(n) - F(n - \tau) . \qquad (32.6')$$

Further, we often encounter in the classical theory a differential of the type

$$\tau \, \frac{\partial \nu(n,\tau)}{\partial n} ; \qquad (32.7)$$

then we interpret this as the second differential

$$\frac{1}{h}\left\{\tau\frac{\partial}{\partial n}\left(\tau\frac{\partial W(n)}{\partial n}\right)\right\}$$

and replace this in turn by the second difference

$$\frac{1}{h} \left[ \left\{ W(n + \tau) - W(n) \right\} - \left\{ W(n) - W(n - \tau) \right\} \right] \\ = \nu(n + \tau; n) - \nu(n; n - \tau)$$

in conformity with the definition of second differentiation

$$\frac{d^2 f}{dx^2} = \lim_{\Delta x \to 0} \frac{f(x + \Delta x) - 2f(x) + f(x - \Delta x)}{(\Delta x)^2}$$
$$= \lim_{\Delta x \to 0} \frac{\{f(x + \Delta x) - f(x)\} - \{f(x) - f(x - \Delta x)\}}{(\Delta x)^2}.$$

This replacement can also be generalized to the case of an arbitrary function  $F(n, \tau)$  which, like  $v(n, \tau)$ , depends both on the integer n specifying the stationary state, and on  $\tau$  indicating the order of the harmonic, and which accordingly can be associated with each quantum theoretical transition  $n \to n - \tau$ . Namely, the differential

$$\tau \, \frac{\partial F(n,\tau)}{\partial n} \tag{32.8}$$

in the cryptogram is to be replaced by the difference

 $F(n + \tau; n) - F(n; n - \tau)$ . (32.8')

An example of  $F(n, \tau)$  would be the Fourier component  $X(n, \tau) \cdot \exp \{2\pi i \nu(n, \tau)t\}$  of the coordinate x. A straightforward application of the rule by which we replaced (32.5) by (32.5') together with the correspondence of  $F(n; n - \tau)$  to  $F(n, \tau)$  may suggest the replacement of (32.8) by  $F(n; n - \tau) - F(n - \tau; (n - \tau) - \tau)$  instead of (32.8'). But this possibility is excluded because by this replacement a certain symmetry possessed by the original quantity (32.8) with respect to a change of sign of  $\tau$  would be lost. For example, in the case where  $F(n, \tau)$  is  $\nu(n, \tau)$ , the quantity  $\tau \partial \nu(n, \tau)/\partial n$  is invariant under the change of sign of  $\tau$ . This symmetry is preserved in  $\nu(n + \tau; n) - \nu(n; n - \tau)$ but not in  $\nu(n; n - \tau) - \nu(n - \tau; (n - \tau) - \tau)$ . We encounter a similar situation also when  $F(n, \tau)$  is, for example,  $X(n, \tau)$ .

In the classical theory a quantity like the one in (32.1) was obtained by expanding in Fourier series the coordinate x(t), for instance, as a function of time. The original quantity x(t), the sum of the terms like (32.1), naturally was an important physical quantity in the classical theory. However, in its quantum theoretical translation, (32.1'), it is doubtful that the quantity as given by the sum of the terms like (32.1') would also play an important role. The reason is that, since each of the summands, (32.1'), is associated with a different transition, the function of time as given by their sum cannot possibly be considered as having an important meaning. Viewed from the quantum theoretical stand-point the essential significance of the fact that in the classical theory the quantity (32.1) is given as the Fourier component of the coordinate x(t) is that it allows the calculation of the Fourier components of  $x^2, x^3, \ldots$ . For instance, if the coordinate x is expressed as

$$x = \sum_{\tau=-\infty}^{+\infty} X(n,\tau) \cdot \exp\{2\pi i \nu(n,\tau)t\},\$$

its square  $x^2$  is given as a series of the same structure, that is a series with the same frequencies, namely

$$\begin{aligned} x^{2} &= \sum_{\tau=-\infty}^{+\infty} \left[ \sum_{\tau'=-\infty}^{+\infty} X(n,\tau') \cdot \exp\left\{ 2\pi i \nu(n,\tau') t \right\} \right] \\ &\quad \cdot X(n,\tau-\tau') \cdot \exp\left\{ 2\pi i \nu(n,\tau-\tau') t \right\} \\ &= \sum_{\tau=-\infty}^{+\infty} \left\{ \sum_{\tau'=-\infty}^{+\infty} X(n,\tau') X(n,\tau-\tau') \right\} \cdot \exp\left\{ 2\pi i \nu(n,\tau) t \right\}. \end{aligned}$$

Or, in accordance with this, the  $\tau$ -th Fourier component of  $x^2$  is given by

$$\begin{cases} \sum_{\tau'=-\infty}^{+\infty} X(n,\tau') \cdot \exp\left\{2\pi i\nu(n,\tau')t\right\} \\ \cdot X(n,\tau-\tau') \cdot \exp\left\{2\pi i\nu(n,\tau-\tau')t\right\} \end{cases}$$
$$= \begin{cases} \sum_{\tau'=-\infty}^{+\infty} X(n,\tau')X(n,\tau-\tau') \\ \cdot \exp\left\{2\pi i\nu(n,\tau)\right\}. \end{cases} (32.9)$$

It is an essential point that the Fourier series of  $x^2$  contains no other frequencies than those already appearing in the Fourier series of x. This very essential point obviously comes from the fact that the classical frequencies  $\nu(n, \tau)$  consist of the fundamental frequency v(n, 1) and its integral multiples and that, consequently, the relationship

$$\nu(n, \tau) + \nu(n, \tau') = \nu(n, \tau + \tau')$$
, (32.10)

or, the same relationship

$$\nu(n,\tau') + \nu(n,\tau-\tau') = \nu(n,\tau)$$

holds. By a similar calculation it is shown that the  $\tau$ -th term in the Fourier series of  $x^3$  is given by

$$\left\{\sum_{\tau'=-\infty}^{+\infty}\sum_{\tau''=-\infty}^{+\infty}X(n,\tau')X(n,\tau'')X(n,\tau-\tau'-\tau'')\right\}\cdot\exp\left\{2\pi i\nu(n,\tau)t\right\}.$$
(32.11)

The relationships of Eq. (32.10) do not hold for the quantum theoretical frequencies and, instead, we have the Rydberg-Ritz combination law,

$$\nu(n; n-\tau) + \nu(n-\tau; n-\tau-\tau') = \nu(n; n-\tau-\tau'),$$
  
or  
$$\nu(n; n-\tau') + \nu(n-\tau'; n-\tau) = \nu(n; n-\tau).$$
(32.10')

Therefore, if we follow the procedure of considering the quantity given by (32.1') as one of the terms of a series, derive the quantity x(t) as the sum of this series, and calculate  $x^2$ ,  $x^3$ , ..., from the obtained x(t), then these quantities contain components with frequencies which did not exist in the original expression for x. This is the very reason why the function of time as given by the sum of (32.1') can not be expected to play an important role. Therefore the replacement of the function x(t) appearing in the cryptogram by a certain quantum theoretical counterpart as a function of time can hardly be considered the correct procedure for decoding. Since, however, what we are looking for are quantities like the intensity or the polarization of the light emitted in each transitions, we already have sufficient information for our purpose when quantities of the form of (32.1') are known separately and it is of secondary importance whether they can be summed up into the form of a function of time or not. That is, we should be content with the result that the quantum theoretical quantity corresponding

to the coordinate x is simply the aggregate of the quantities of the form of (32.1'). This aggregate is called a "matrix" for the reason given in the next section (§ 33).

In the meantime, however, it should be possible for us to calculate the quantum theoretical transition components of  $x^2$ ,  $x^3$ ,..., from those of x as given by (32.1') just as (32.9) and (32.11) express the Fourier components of  $x^2$ ,  $x^3$ ,..., in terms of X, (32.1), in the classical theory. We require with Heisenberg, that no other frequencies should appear in the transition components of  $x^2$ ,  $x^3$ ,..., than those already existing in the transition components of x. This requirement fixes the method for calculating the transition components of  $x^2$ ,  $x^3$ ,..., from those of x in the following manner.

We first note that, to the classical Fourier components  $X(n, \tau) \cdot \exp\{2\pi i \nu(n, \tau)t\}$  and  $X(n, \tau - \tau') \cdot \exp\{2\pi i \nu(n, \tau - \tau')t\}$ , correspond the transition components  $X(n; n - \tau) \cdot \exp\{2\pi i \nu(n; n - \tau)t\}$  and  $X(n; n - \tau + \tau') \cdot \exp\{2\pi i \nu(n; n - \tau + \tau')t\}$ , respectively. The straightforward introduction of these replacements into Eq. (32.9) i.e. the interpretation of

$$\left[\sum_{\tau'=-\infty}^{+\infty} X(n; n-\tau') \cdot \exp\left\{2\pi i\nu(n; n-\tau')t\right\} \\ \cdot X(n; n-\tau+\tau') \cdot \exp\left\{2\pi i\nu(n; n-\tau+\tau')t\right\}\right]$$

as the  $\tau$ -th transition components of  $x^2$  does not satisfy our requirement since the quantum theoretical frequency does not obey Eq. (32.10) and, hence,  $\nu(n; n - \tau') + \nu(n; n - \tau + \tau')$  is not equal to  $\nu(n; n - \tau)$ . Recalling the fact that the quantum theoretical frequencies satisfy Eq. (32.10') instead of Eq. (32.10), and that

$$\nu(n; n-\tau') + \nu(n-\tau'; n-\tau) = \nu(n; n-\tau) ,$$

we arrive at the conclusion that the correct assignment as the quantum theoretical transition component of  $x^2$  must be

$$\begin{bmatrix} \sum_{\tau'=-\infty}^{+\infty} X(n; n-\tau') \exp \{ 2\pi i \nu(n; n-\tau')t \} \\ \cdot X(n-\tau'; n-\tau) \cdot \exp \{ 2\pi i \nu(n-\tau'; n-\tau)t \} \end{bmatrix}$$

$$= \left\{ \sum_{\tau'=-\infty}^{+\infty} X(n; n-\tau') X(n-\tau'; n-\tau) \right\} \cdot \exp \{ 2\pi i \nu(n; n-\tau)t \}.$$
(32.9)

This expression with the correct frequency is then something which replaces Eq. (32.9) in the cryptogram.

By a similar consideration we should take, as the quantum theoretical transition component of  $x^3$ ,

$$\sum_{\tau'=-\infty}^{+\infty} \sum_{\tau''=-\infty}^{+\infty} X(n; n - \tau') \cdot \exp \left\{ 2\pi i \nu(n; n - \tau') t \right\}$$

$$\cdot X(n - \tau'; n - \tau'') \cdot \exp \left\{ 2\pi i \nu(n - \tau'; n - \tau'') t \right\}$$

$$\cdot X(n - \tau''; n - \tau) \cdot \exp \left\{ 2\pi i \nu(n - \tau''; n - \tau) t \right\},$$
(32.11')

which has the correct frequency  $\nu(n; n - \tau)$  due to the combination law

$$\nu(n; n-\tau') + \nu(n-\tau'; n-\tau'') + \nu(n-\tau''; n-\tau) = \nu(n; n-\tau) .$$

We have thus obtained the key for finding the quantity which is to replace the square and the cube of the coordinate x appearing in the cryptogram. The essential difference of these quantum theoretical multiplication rules from those of the classical theory is as follows. According to the classical theory, the Fourier component of the product corresponding to the *n*-th state of motion depends naturally only on those Fourier components of the factors corresponding to the same n-th state of motion. On the contrary, in the quantum theory the transition component of the product from the *n*-th stationary state involves also the transition components of the factors from stationary states other than the *n*-th. For instance, in Eq. (32.9) the Fourier component of  $x^2$  in the *n*-th state is expressed in terms of  $X(n, \tau - \tau')$  and  $X(n, \tau')$ . i.e., in terms of the Fourier components in the same n-th state, while in Eq. (32.9') the transition component of  $x^2$  from the *n*-th state depends not only on those from the same *n*-th state,  $X(n; n - \tau')$ , but also on those from different stationary states,  $X(n - \tau'; n - \tau)$ . In view of this, it is now clear that the transition component is to be replaced by something entirely different from the electron motion in a certain state though there is in a sense some correspondence between the two. This fact had already been vaguely anticipated in Bohr's correspondence principle but it is the essential contribution of Heisenberg's finding that he could express this fact precisely in the form of his multiplication rule. The important key point which led him to this finding was the recognition that the Rydberg-Ritz combination law,

Eq. (32.10'), was the one to replace the classical frequency relation,  $v(n, \tau) = \tau v(n, 1)$ , or Eq. (32.10).

This multiplication rule can easily be generalized. For instance, when we have in the classical theory two quantities x and y, (which need not necessarily mean x and y coordinates but may be any two dynamical quantities) and when we have

the Fourier component of 
$$x = X(n, \tau) \cdot \exp \{ 2\pi i \nu(n, \tau) t \}$$
,  
the Fourier component of  $y = Y(n, \tau) \cdot \exp \{ 2\pi i \nu(n, \tau) t \}$ ,  
(32.12)

and, as their product,

the Fourier component of xy

$$=\left\{\sum_{\tau'=-\infty}^{+\infty} X(n,\tau')Y(n,\tau-\tau')\right\} \cdot \exp\left\{2\pi i\nu(n,\tau)t\right\}, \quad (32.13)$$

we should replace them by

the quantum theoretical transition component of x

$$= X(n; n-\tau) \cdot \exp \left\{ 2\pi i \nu(n; n-\tau) t \right\},$$

the quantum theoretical transition component of y

 $= Y(n; n-\tau) \cdot \exp \left\{ 2\pi i \nu(n; n-\tau) t \right\}$ 

and, as their product,

the quantum theoretical transition component of xy

$$=\left\{\sum_{\tau'=-\infty}^{+\infty} X(n; n-\tau) Y(n-\tau; n-\tau')\right\} \cdot \exp\left\{2\pi i \nu(n; n-\tau)t\right\},$$
(32.13')

respectively.

A number of keys for decoding the cryptogram have been obtained in this way. A point worthy of special attention in connection with Eq. (32.13') is that, according to this quantum theoretical multiplication rule, the product xy is generally different from the product yx since

$$\sum_{\tau'} X(n; n-\tau') Y(n-\tau'; n-\tau) = \sum_{\tau'} Y(n; n-\tau') X(n-\tau'; n-\tau)$$

holds only in exceptional cases.

213

(32.12')

Before we apply Heisenberg's key to some of the standard simple cases to arrive at meaningful results, we shall apply it to Bohr's quantum condition. Since the quantum theoretical quantity in the Heisenberg theory is considered as an aggregate of the transition components, the correspondence between the quantum theoretical calculations and those of the old theory becomes possible only when all the calculations in the old theory have been performed in the form of Fourier series and when all the results have been expressed in the form of relationships between the Fourier components. Now the quantum condition in the old theory,

$$\oint p \mathrm{d}x = J = nh , \qquad (32.14)$$

does not in this form show explicitly a similarity to the corresponding quantum mechanical equation. For the decoding of this quantum condition it is necessary to rewrite it as a relationship between the Fourier components of x and p. We now carry through the actual decoding of the quantum condition.

In order to avoid excessive complexities which may hinder the reader from understanding the essential reasoning, we consider the simple one-dimensional case. The quantum condition can be written as

$$\oint p \mathrm{d}x = \oint p \dot{x} \mathrm{d}t = nh \, .$$

The substitution, for p and x in this relation, of their Fourier expansions,

$$p = \sum_{\tau=-\infty}^{+\infty} P(n,\tau) \cdot \exp \{ 2\pi i \nu(n,\tau) t \},$$

and

$$x = \sum_{\tau=-\infty}^{+\infty} X(n,\tau) \cdot \exp \{ 2\pi i \nu(n,\tau) t \},$$

gives

$$nh = 2\pi i \sum_{\tau=-\infty}^{+\infty} \sum_{\tau'=-\infty}^{+\infty} \int P(n,\tau) X(n,\tau') \nu(n,\tau') \\ \cdot \exp\left\{ 2\pi i \left\{ \nu(n,\tau) + \nu(n,\tau') \right\} t \right\} dt,$$

where the integration is over one complete period of motion, i.e., from t = 0 to  $1/\nu(n, 1)$ . Upon performing this integration, the right hand side of the above equation becomes

$$-2\pi i \sum_{\tau=-\infty}^{+\infty} P(n,\tau) X^{\bullet}(n,\tau) \frac{\nu(n,\tau)}{\nu(n,1)} = -2\pi i \sum_{\tau=-\infty}^{+\infty} P(n,\tau) X^{\bullet}(n,\tau) \tau .$$

214

This is the quantum condition expressed in terms of Fourier components, and this is now to be translated into quantum language.

Before entering into this problem we must first make clear the true meaning of the relationship (32.14). In Bohr's theory this relationship was introduced as the quantum condition, i.e., the condition which selects the quantum-mechanically allowed states: Bohr's stationary states were determined by the requirement that n on the right-hand side should be an integer. But, according to our philosophy described in § 23 (v), the discrete nature of the stationary states should appear as something which corresponds to the discreteness of  $\tau$  in the course of the correct translation of the cryptogram, instead of resulting from such an ad hoc conditions. Thus we should allow any positive values for n in the cryptogram, and the relation (32.14) should be merely used as the equation by which one fixes the so far undetermined value of I or n when the amplitude of p and x are specified; or conversely, one fixes the undetermined amplitudes when the value of J or n is specified. Bearing this in mind we differentiate both sides of the equation with respect to n to obtain

$$\frac{h}{2\pi i} = -\sum_{\tau=-\infty}^{+\infty} \tau \frac{\partial}{\partial n} \left\{ P(n,\tau) X^*(n,\tau) \right\}.$$
(32.15)

This differentiation was performed because in the resultant form we can utilize the keys (32.8) and (32.8').

The immediate use of the key, Eqs. (32.8) and (32.8'), in Eq. (32.15) gives then

$$\frac{h}{2\pi i} = -\left\{\sum_{\tau=-\infty}^{+\infty} P(n+\tau;n)X^*(n+\tau;n) - \sum_{\tau=-\infty}^{+\infty} P(n;n-\tau)X^*(n;n-\tau)\right\}.$$

Making use of the relation (32.3') which states that  $X(n; n + \tau)$  and  $X(n + \tau; n)$  are complex conjugates of each other, we can further rewrite this relationship in the form

$$\frac{h}{2\pi i} = -\left\{\sum_{\tau=-\infty}^{+\infty} P(n+\tau;n)X(n;n+\tau) - \sum_{\tau=-\infty}^{+\infty} P(n;n-\tau)X(n-\tau;n)\right\}.$$

Or, changing the sign of  $\tau$  in the first series which does not affect the result, we finally obtain

$$\frac{h}{2\pi i} = \left\{ \sum_{\tau=-\infty}^{+\infty} P(n; n-\tau) X(n-\tau; n) - \sum_{\tau=-\infty}^{+\infty} X(n; n-\tau) P(n-\tau; n) \right\}.$$
(32.15')

This is then the quantum mechanical relationship which replaces Eq. (32.15). We have thus completed the decoding of Eq. (32.15). In case the momentum  $\phi$  is given by  $m\dot{x}$ , Eq. (32.15') can also be written as

$$h = - 8\pi^2 m \sum_{\tau = -\infty}^{+\infty} |X(n; n - \tau)|^2 v(n; n - \tau) . \quad (32.15'')$$

We shall next show how the decoding procedure of Heisenberg works by taking as an example a practical dynamical system. The simplest case is that of Planck's oscillator. The point mass, the coordinate of which is x, suffers in this case a force proportional to x. The problem therefore would not require the use of Heisenberg's multiplication rule (32.9') and, since our present aim is to show the actual application of this rule, we add an additional force proportional to  $x^2$ . The equation of motion then becomes

$$\ddot{x} + (2\pi v)^2 x + \lambda x^2 = 0$$
, (32.16)

where  $(2\pi\nu)^2$  is written for  $\varkappa/m$ .

As mentioned above, the first thing we have to do is to express Eq. (32.16) as relationships between Fourier components of x. In classical mechanics, the coordinate x is expanded in the Fourier series with respect to time, the general term of which is

$$X(n, \tau) \cdot \exp\{2\pi i \nu(n, \tau)t\},$$
 (32.17)

where  $v(n, \tau) = \tau v(n, 1)$ ,  $v(n, \tau) = -v(n, -\tau)$  and, in addition,  $X(n, \tau) = X^*(n, -\tau)$  in order to make the coordinate x real.

The Fourier component of x being given by (32.17), that of  $\ddot{x}$  becomes

$$- \{ 2\pi \nu(n, \tau) \}^2 X(n, \tau) \cdot \exp \{ 2\pi i \nu(n, \tau) t \}, \qquad (32.18)$$

and that of  $x^2$  becomes

$$\sum_{\tau'=-\infty}^{+\infty} X(n,\tau')X(n,\tau-\tau') \cdot \exp\left\{2\pi i\nu(n,\tau)t\right\}.$$
 (32.19)

Introducing these expressions into the left hand side of the equation of motion, Eq. (32.16), and putting it equal to zero for each component, we obtain

$$(2\pi)^{2} \{ \nu^{2} - \nu^{2}(n,\tau) \} X(n,\tau) + \lambda \sum_{\tau'=-\infty}^{+\infty} X(n,\tau') X(n,\tau-\tau') = 0.$$
(32.20)

These are the simultaneous equations for  $X(n, \tau)$ . Then the solution of these simultaneous equations gives all the Fourier components of the coordinate x, and at the same time the frequency v(n, 1) is determined from the condition that these equations should have a solution.

These are the procedures of the classical theory. In order to translate them into quantum mechanical language, we proceed, in accordance with the general patterns described above, as follows. Namely, in quantum mechanics we start, instead of (32.17), with the quantum mechanical transition component

$$X(n; n-\tau) \cdot \exp \left\{ 2\pi i \nu(n; n-\tau) t \right\}, \qquad (32.17')$$

where  $v(n; n - \tau) = -v(n - \tau; n)$  and  $X(n; n - \tau) = X^*(n - \tau; n)$  to make X a real-number-like quantity.

The transition component of  $\ddot{x}$  is obtained by differentiating Eq. (32.17') twice with respect to t, and is

$$- \{ 2\pi\nu(n; n-\tau) \}^2 X(n; n-\tau) \cdot \exp \{ 2\pi i\nu(n; n-\tau)t \}, \qquad (32.18')$$

and that of  $x^2$  is given, by the use of our multiplication rule (32.9'), by

$$\sum_{\tau'} X(n; n-\tau') X(n-\tau'; n-\tau) \cdot \exp\left\{ 2\pi i \nu(n; n-\tau) t \right\}.$$
(32.19')

Upon substitution of these expressions into the left hand side of Eq. (32.16), we obtain

$$(2\pi)^{2} \{ \nu^{2} - \nu^{2}(n; n - \tau) \} X(n; n - \tau) + \lambda \sum_{\tau'} X(n; n - \tau') X(n - \tau'; n) = 0. \quad (32.20')$$

This is the equation for the  $X(n; n - \tau)$ , the quantum mechanical transition components of x.

Since the actual solution either of Eq. (32.20) or of Eq. (32.20') is difficult, we treat in what follows only the case  $\lambda = 0$ . Eq. (32.20) for  $\lambda = 0$  can readily be solved to give

$$X(n, \tau) = 0 \text{ unless } \tau = \pm 1,$$
  

$$X(n, + 1) = X^{*}(n, -1) \neq 0,$$
 (32.21)  

$$v(n, 1) = v = -v(n, -1),$$
  
for all values of  $n$ .

Although there exist, besides this solution, a number of apparently different ones, for instance  $X(n, \tau) = 0$  unless  $\tau = \pm 2$ ,  $X(n, \pm 2) \neq 0$  and  $\nu(n, 2) = \nu$ , their use would not give rise to different final conclusions. Similarly, we obtain from Eq. (32.20') for  $\lambda = 0$  the solutions.

$$X(n; n - \tau) = 0 \text{ unless } \tau = \pm 1,$$
  

$$X(n; n - 1) = X^{\bullet}(n - 1; n) \neq 0,$$
 (32.21')  

$$v(n; n - 1) = v = -v(n - 1; n),$$
  
for all values of  $n$ .

The yet undetermined amplitudes X(n, +1) and X(n; n-1) in Eqs. (32.21) and (32.21'), respectively, can then be determined from the conditions (32.15) or (32.15'), respectively. Making use of p = mx, we obtain from Eq. (32.15)

$$h = 8\pi^2 m \nu \frac{\partial}{\partial n} |X(n, 1)|^2, \qquad (32.22)$$

which upon integration gives

$$|X(n, 1)|^2 = \frac{1}{8\pi^2 vm} (n + \text{const.}) h.$$

Analogously, we have from Eq. (32.15')

$$h = 8\pi^2 m \nu \{ | X(n + 1, n) |^2 - | X(n, n - 1) |^2 \}, \quad (32.22')$$

СН. 5, § 32]

which upon "integration" becomes

$$|X(n; n-1)|^2 = \frac{1}{8\pi^2 \nu m} (n + \text{const.})h$$
.

The integration constant in each case can be determined as follows. The oscillator is to have a certain stable state, the ground state, in which it does not make a transition at all. In this ground state, the amplitude X(n, +1) of the oscillation in the classical theory must be zero, while, according to quantum mechanics, the transition to a state of still smaller quantum number is forbidden i.e., X(n; n - 1) = 0. Since the labelling of the states is arbitrary in the sense that we can choose any one of them to specify the ground state, we set n = 0 for this state and hence the integration constants in both cases become zero. Accordingly, we have

$$|X(n, 1)|^{2} = \frac{1}{8\pi^{2} \nu m} nh, \qquad (32.23)$$

or

$$|X(n; n-1)|^2 = \frac{1}{8\pi^2 \nu m} nh$$
. (32.23')

Eq. (32.23) is nothing but the relationship (19.46') obtained previously in § 19.

By this equation the amplitude is determined when n is specified, but as mentioned before, we should in Eq. (32.23) allow n to take any positive values, since Eq. (32.23) was obtained by integrating Eq. (32.22) from zero to n. One has to notice, on the other hand, that n in Eq. (32.23') should be an integer. The relationship (32.23') has no meaning at all for non-integral values of n because this relationship was derived by "integrating" the difference equation (32.22') step by step from n = 0 to n = 1, and from n = 1 to n = 2 and so on. One now sees precisely how the discrete nature of the quantum theoretical quantities appears and how this is related to the discreteness of  $\tau$ . This is a remarkable result of Heisenberg's theory.

The relationship (32.23'), being the correct quantum theoretical expression to replace Bohr's condition J = nh, clearly shows the irrelevancy of the assumption of the old quantum theory, and that it must in some way be re-interpreted. For, according to its literal interpretation, the equation, J = nh, imposes a certain condition to be

satisfied by the amplitude of the electron's orbital motion and Eq. (32.23) is this condition. In the light of the correspondence principle we reinterprete J = nh as imposing a certain condition on the quantum mechanical quantity which is closely related to the transition and Eq. (32.23') is nothing but the condition re-interpreted in this way.

We now examine the energies of the stationary states. To do this, we introduce our solutions (32.21) or (32.21)

To do this, we introduce our solutions (32.21) or (32.21') into the energy function

$$W = \frac{1}{2}m x^{2} + \frac{1}{2}m (2\pi\nu)^{2}x^{2}. \qquad (32.24)$$

We should then obtain the Fourier components and the quantum mechanical components, respectively, of the energy and these would completely tell us the values of the energy. We begin first with the Fourier components of the energy in the classical theory. The Fourier component of  $\dot{x}^2$  is

$$-(2\pi)^{2} \sum_{\tau'} \nu(n,\tau')\nu(n,\tau-\tau') X(n,\tau') X(n,\tau-\tau')$$
$$\cdot \exp\left\{2\pi i\nu(n,\tau)t\right\} \qquad (32.25)$$

which, upon substitution of Eq. (32.21), becomes

 $\begin{aligned} &8\pi\nu^{2} | X(n, 1) |^{2} & \text{for } \tau = 0 , \\ &- (2\pi)^{2}\nu^{2} | X(n, 1) |^{2} \cdot \exp \{ 2\pi i\nu(n, 2)t \} & \text{for } \tau = +2 , \\ &- (2\pi)^{2}\nu^{2} | X(n, -1) |^{2} \cdot \exp \{ 2\pi i\nu(n, 2)t \} \text{for } \tau = -2 , \\ &0 & \text{for all other } \tau \text{'s.} \end{aligned}$ 

In the meantime, the Fourier component of  $x^2$  is, from (32.9),

$$\begin{array}{ll} 2 \mid X(n, 1) \mid^{2} & \text{for } \tau = 0 , \\ \mid X(n, 1) \mid^{2} \cdot \exp \left\{ 2\pi i \nu(n, 2) t \right\} & \text{for } \tau = + 2 , \\ \mid X(n, -1) \mid^{2} \cdot \exp \left\{ -2\pi i \nu(n, 2) t \right\} & \text{for } \tau = - 2 , \\ 0 & \text{for all other } \tau \text{'s.} \end{array}$$

Making use of these results in Eq. (32.24), we readily obtain the Fourier component of the energy W; it is

$$\begin{cases} 8\pi^2 v^2 m \mid X(n, 1) \mid^2 & \text{for } \tau = 0 \\ 0 & \text{for all other } \tau' \text{s.} \end{cases}$$
(32.28)

Сн. 5, § 32]

Therefore the Fourier expansion of the energy W has a non-vanishing term only for  $\tau = 0$ , i.e. for the time independent term. In other words, the energy W is a constant with respect to time and its value is given by  $W = 0.2^{2} t + W(-1) t^{2}$ 

$$W = 8\pi^2 \nu^2 m \mid X(n, 1) \mid^2,$$

or, by the use of Eq. (32.23),

$$W = nh\nu . \tag{32.29}$$

These are the familiar energy levels of the harmonic oscillator, provided that n is restricted to integers.

We now turn to the translation of this result into quantum mechanical language. In order to accomplish this, we start with the quantum mechanical transition component of  $\dot{x}^2$  which is

$$- (2\pi)^2 \sum_{\tau'} \nu(n; n - \tau')\nu(n - \tau'; n - \tau) \cdot X(n; n - \tau')X(n - \tau'; n - \tau) \cdot \exp\{2\pi i\nu(n, n - \tau)t\}.$$
(32.25')

Introducing Eq. (32.21') into this expression we obtain the transition component of  $\dot{x}^2$  as

$$\begin{array}{ll} (2\pi)^2 \nu^2 \left\{ \left| X(n;n-1) \right|^2 + \left| X(n;n+1) \right|^2 \right\} & \text{for } \tau = 0 , \\ - (2\pi)^2 \nu^2 X(n;n-1) X(n-1;n-2) \cdot \exp \left\{ 2\pi i \nu(n;n-2) t \right\} & \\ & \text{for } \tau = +2 , \\ - (2\pi)^2 \nu^2 X(n;n+1) X(n+1;n+2) \cdot \exp \left\{ 2\pi i \nu(n;n+2) t \right\} & \\ & \text{for } \tau = -2 . \\ 0 & \text{for all other } \tau \text{'s.} & (32.26') \end{array}$$

In the meantime, Eq. (32.9') gives also the transition components of  $x^2$  as

$$\{ | X(n; n-1) |^2 + | X(n; n+1) |^2 \}$$
 for  $\tau = 0$ ,  

$$X(n; n-1)X(n-1; n-2) \cdot \exp \{ 2\pi i \nu(n; n-2)t \}$$
 for  $\tau = +2$ ,  

$$X(n; n+1)X(n+1; n+2) \cdot \exp \{ 2\pi i \nu(n; n+2)t \}$$
 for  $\tau = -2$ .  

$$(32.27')$$

Upon substitution of these results into Eq. (32.24), we arrive at the result that the quantum mechanical transition component of the energy W is

$$(2\pi\nu)^{2} \{ | X(n; n-1) |^{2} + | X(n; n+1)^{2} \} \text{ for } \tau = 0 .$$
  
0 for all other  $\tau$ 's. (32.28')

[сн. 5, § 32

The result is that here also only the  $\tau = 0$  component, i.e., the time independent component, of W has a non-vanishing value, all the others being zero. This result can be interpreted as implying that the energy has a constant value in each stationary state. However, the content of this statement is not yet quite clear since the quantum theoretical transition component can not be summed up to build a function of time. Nevertheless, we assume in correspondence with the classical theory that when all the transition components except the ones of  $\tau = 0$ of a physical quantity are zero this physical quantity has a constant value in each of the stationary states. We then have as the energy value in the *n*-th stationary state

 $W_{n} = (2\pi\nu)^{2}m \{ |X(n; n-1)|^{2} + |X(n; n+1)|^{2} \},\$ 

or, using Eq. (32.23'),

$$W_n = (n + \frac{1}{2})h\nu$$
,  $n = 0, 1, 2, \dots$  (32.29')

These are the quantum mechanical energy levels of the harmonic oscillator. In the above we assumed that these are the energies of the oscillator in the stationary states and this is supported by the observation that the use of Bohr's frequency condition gives for the frequency associated with the transition  $n \rightarrow n - 1$ 

$$\nu(n; n-1) = \nu = \frac{W_n - W_{n-1}}{h}$$

which is identical to the frequency previously obtained in Eq. (32.21'). Therefore, we see that the transition components do in fact oscillate with frequencies statifying Bohr's frequency condition. The general proof of this statement, not only for the harmonic oscillator, will be given in the next section.

According to Eq. (32.21'), the transition components of the coordinate,  $X(n; n - \tau)$ , are all zero except for  $\tau = \pm 1$ . This immediately leads to the conclusion that, if the oscillator has an electric charge e, no transitions can occur with changes of the quantum number larger than unity. This conclusion is the same as that arrived at in § 23 (iii). The absolute square of the transition element for the transition  $n \rightarrow n - 1$  is given by Eq. (32.23'). According now to the basic assumption of the correspondence principle, the light thereby emitted is to be identical,

# CH. 5, § 32] HEISENBERG'S DISCOVERY

in the statistical sense, to that which would be emitted by an electric dipole moving with a Fourier component which is numerically the same as X(n; n-1). The statistical mean energy emitted per unit time by this transition is then given by

$$\frac{\mathrm{d}E_{n \to n-1}}{\mathrm{d}t} = \frac{(2\pi)^4 \nu^4}{3c^3} 4c^2 |X(n; n-1)|^2$$

$$= \frac{8\pi^2}{3} \cdot \frac{e^2 \nu^2 h}{mc^3} \cdot n ,$$
(32.31)

and the transition probability, obtained by dividing this by  $h\nu$ , is

$$A_{n \to n-1} = \frac{8\pi^2}{3} \frac{e^2 \nu^2}{mc^3} n . \qquad (32.32)$$

This is identical with the result, Eq. (23.31), obtained by the old correspondence principle.

Eq. (32.29') shows that the values of the energy of the oscillator are different from the corresponding values in the old theory. The difference is  $\frac{1}{2}h\nu$  and it indicates that the oscillator has a certain amount of energy even in the ground state. This is called the zero-point energy and its existence had been anticipated also from experiment.

The reader should have in this way understood the basic patterns of Heisenberg's theory. The theory, however, is still a very formal one and involves many elements whose basic physical meanings are obscure. For instance, it is entirely unclear yet in what way the cartesian coordinate x of the electron (which in this theory means only the aggregate of the transition elements) is related to the classical concept of the location of the electron. Moreover, many phenomena show that Heisenberg's theory applies not only to electrons but also to all other particles. If the location of an electron and of other particles were to become such an abstract aggregate, or matrix, how can we explain in this theory the track of a particle observed commonly in a Wilson chamber? It was Lorentz who said: "Can you imagine me to be nothing but a matrix? It is hardly to be believed that all this is real". Furthermore, this theory is scarcely more complete than the old correspondence principle as far as the description of light emission in a transition is concerned. As a matter of fact, the new achievement of

223

this theory (though a large one) is only that the quantum mechanical transition components are used instead of the Fourier components of the classical orbital motion. One has then to introduce the hypothetical dipole, oscillating with the same Fourier components, and calculate the intensity and the polarization of the light emitted by this dipole according to the Maxwell theory. The results thus derived are then re-interpreted statistically. This is the way one handles a problem in this theory and hence the theory is not complete by itself. As far as the behavior of the atom is concerned one does not indeed have to rely upon the classical theory coupled with a subsequent quantum theoretical re-interpretation, but can instead, by Heisenberg's method of calculation, directly derive conclusions starting from relationships between quantum mechanical quantities; but with light, the true law of interaction between an atom and radiation is yet unknown, and one must still make calculations with the aid of the classical theory, and give, ad hoc, a quantum mechanical meaning to the results. Since this is true, we realize that the theory of Heisenberg is not the ultimate one and is still far from explaining the dual nature of light.

However, it is rather surprising that, when we apply the Heisenberg theory to various problems, neglecting for the time being this unsatisfactory feature of the theory, we arrive at the correct results. Therefore, this theory must contain at least a part of the correct theory to come. What we shall do now is to refine the mathematical form of this theory to make it easy to use. This refinement was due to Born, Jordan and Heisenberg himself.

# § 33. Matrix Mechanics

As was explained in the last section, the coordinate x, or the momentum p, of an electron could not, according to the Heisenberg theory, be considered as an ordinary function of time as it was in the classical theory and should be treated instead as an ensemble of the transition elements  $X(n; n - \tau) \exp \{2\pi i \nu(n; n - \tau)t\}$  for the coordinate, and  $P(n; n - \tau) \exp \{2\pi i \nu(n; n - \tau)t\}$  for the momentum. Each element of these ensembles has been labelled by specifying the two numbers n and  $\tau$  but it is more convenient to use n and  $n' = n - \tau$  to label it. Each element then corresponds in a one to one manner to the transition  $n \to n'$ . сн. 5, § 33]

#### MATRIX MECHANICS

Any physical quantity in general is, like the coordinate x or the momentum p, expressible by an ensemble of its transition components: the coordinate x is the ensemble of

$$X_{nn'} \exp\left\{2\pi \mathrm{i}\nu_{nn'}t\right\}$$

and the momentum p is that of

$$P_{nn'}\exp\left\{2\pi\mathrm{i}\nu_{nn'}t\right\}.$$

We arrange these components which have two suffixes in the form of a two-dimensional pattern, for instance,

$$x = \begin{bmatrix} X_{11} \exp(2\pi i\nu_{11}t) & X_{12} \exp(2\pi i\nu_{12}t) & X_{13} \exp(2\pi i\nu_{13}t) \dots \\ X_{21} \exp(2\pi i\nu_{21}t) & X_{22} \exp(2\pi i\nu_{22}t) & X_{23} \exp(2\pi i\nu_{23}t) \dots \end{bmatrix} .$$
(33.1)

The equality sign in the expression

$$x = \begin{bmatrix} \cdots \\ \cdots \\ \cdots \end{bmatrix}$$

implies that the quantity x is the ensemble of the components on the right hand side. In mathematics such two-dimensional ensembles are called "matrices" provided that they obey certain addition and multiplication rules. Each component of a matrix will hereafter be called a "matrix element" of that matrix.

In our case, as was explained in the last section, the coordinate is to be a quantity like a real number and the frequencies have to satisfy Bohr's condition. Then we have, from Eqs. (32.2') and (32.3'),

$$X_{nn'} = X_{n'n}^*, \quad v_{nn'} = -v_{n'n}, \text{ and } v_{nn} = 0.$$
 (33.2)

Denoting  $X_{nn}$ , exp  $(2\pi i \nu_{nn} t)$  by  $x_{nn}$ , Eq. (33.1) becomes

$$x = \begin{bmatrix} x_{11} & x_{12} & x_{13} \dots \\ x_{21} & x_{22} & x_{23} \dots \\ x_{31} & x_{32} & x_{33} \dots \\ \vdots & \vdots & \vdots \end{bmatrix}$$

The  $x_{nn}$ 's then are generally functions of time and satisfy

$$x_{nn'} = x_{n'n}^* \,. \tag{33.4}$$

The elements with n = n',  $x_{nn}$ , are exceptional in that they are independent of time and

$$x_{nn} = \text{real number}$$
 (33.4')

 $x_{nn'}$  is called the (n, n') element of x and the elements with n = n' the diagonal elements.

Analogously, we write

$$p = \begin{bmatrix} p_{11} & p_{21} & p_{13} \cdots \\ p_{21} & p_{22} & p_{23} \cdots \\ p_{31} & p_{32} & p_{33} \cdots \\ \vdots & \vdots & \vdots \end{bmatrix},$$
(33.5)

where all the  $p_{nn'}$  except  $p_{nn}$ , are functions of time and satisfy

$$p_{nn'} = p_{n'n}^* , \qquad (33.6)$$

while  $p_{nn}$  is independent of time and

$$p_{nn} = \text{real number}$$
. (33.6')

Any dynamical quantity, A, not only a coordinate or momentum but also an angular momentum or an energy for instance, is now considered to be represented by a matrix of the form

$$A = \begin{bmatrix} A_{11} & A_{12} & A_{13} \dots \\ A_{21} & A_{22} & A_{23} \dots \\ A_{31} & A_{32} & A_{33} \dots \end{bmatrix},$$
 (33.7)

where each of the matrix elements  $A_{nn}$ , is a function of time and is supposed to vary as exp  $(2\pi i \nu_{nn}.t)$ . In case the quantity A is a quantity like a real number we have

$$A_{nn'} = A_{n'n}^{\bullet} \tag{33.8}$$

and, hence,

$$A_{nn} = \text{real number} \tag{33.9}$$

226

### MATRIX MECHANICS

which is independent of time. That is, the diagonal elements of the matrix representing a real physical quantity are all real numbers and independent of time.

In general, when the relation (33.8) holds the matrix (33.7) is said to be hermitian.

For a given dynamical system, we consider each of the various physical quantities of the system as a matrix. The aim of quantum mechanics is then to determine the matrix elements, amplitudes and frequencies, of these matrices. We have shown in the last section how to calculate these quantities, taking a simple example. The recipe for this calculation, given by Heisenberg, can be summarized as follows, in the new terminology:

*Recipe* (I): Consider that each quantity is a matrix and that, when the quantity is real-number-like, the matrix is hermitian.

Recipe (II): Assume that the (n, n') element of the matrix of a physical quantity oscillates, as a function of time, as  $\exp(2\pi i \nu_{nn'} t)$ .

Recipe (III): For the frequencies there is the combination law

$$v_{nn'} + v_{n'n''} = v_{nn''}$$
, (III.1)

from which it results further that

$$v_{nn} = 0$$
, and  $v_{nn'} = -v_{n'n}$ . (III.2)

*Recipe* (IV): The time derivative of a physical quantity should be defined by the matrix whose elements are the time derivatives of the corresponding elements of the matrix representing the original quantity.

For instance, when the coordinate x is expressed as

$$x = \begin{bmatrix} x_{11} & x_{12} & x_{13} \dots \\ x_{21} & x_{22} & x_{23} \dots \\ x_{31} & x_{32} & x_{33} \dots \\ \vdots & \vdots & \vdots \end{bmatrix},$$

the velocity is given by

$$v = \begin{bmatrix} \dot{x}_{11} & \dot{x}_{12} & \dot{x}_{13} \dots \\ \dot{x}_{21} & \dot{x}_{22} & \dot{x}_{23} \dots \\ \dot{x}_{31} & \dot{x}_{32} & \dot{x}_{22} \dots \\ \vdots & \vdots & \vdots \end{bmatrix}.$$
 (IV.1)

Or, by means of (I) and (II),

$$v = 2\pi i \begin{bmatrix} 0 & \nu_{12} x_{12} & \nu_{13} x_{13} \dots \\ \nu_{21} x_{21} & 0 & \nu_{23} x_{23} \dots \\ \nu_{31} x_{31} & \nu_{32} x_{32} & 0 \dots \\ \vdots & \vdots & \vdots \end{bmatrix}.$$
(IV.2)

We can write this briefly as

$$v_{nn'} = (\dot{x})_{nn'} = 2\pi i v_{nn'} x_{nn'} . \qquad (IV.3)$$

In general, for any physical quantity of real-number-like nature, a relationship similar to (IV.3) holds, i.e.,

$$(A)_{nn'} = 2\pi i \nu_{nn'} A_{nn'} . \tag{IV.4}$$

Recipe (V): The sum of the two physical quantities A and B, i.e. the quantity (A + B), should be defined by the matrix whose elements are the sum of the corresponding elements of the two matrices representing A and B.

In our notation, when

$$A = \begin{bmatrix} A_{11} & A_{12} & A_{13} \dots \\ A_{21} & A_{22} & A_{23} \dots \\ A_{31} & A_{32} & A_{33} \dots \end{bmatrix} \text{ and } B = \begin{bmatrix} B_{11} & B_{12} & B_{13} \dots \\ B_{21} & B_{22} & B_{23} \dots \\ B_{31} & B_{32} & B_{33} \dots \\ \vdots & \vdots & \vdots & \vdots \end{bmatrix}, \quad (V.1)$$

then

$$A + B = \begin{bmatrix} A_{11} + B_{11} & A_{12} + B_{12} & A_{13} + B_{13} \dots \\ A_{21} + B_{21} & A_{22} + B_{22} & A_{23} + B_{23} \dots \\ A_{31} + B_{31} & A_{32} + B_{32} & A_{33} + B_{33} \dots \end{bmatrix}.$$
 (V.2)

Or, more briefly,

$$(A + B)_{nn'} = A_{nn'} + B_{nn'}.$$
 (V.3)

In particular when A = B, i.e. when the corresponding matrix elements of the two matrices are all equal, we write (A + A) as 2A. Then,

$$(2A)_{nn'} = 2A_{nn'} \,. \tag{V.3'}$$

сн. 5, § 33]

MATRIX MECHANICS

We generalize this and define for any arbitrary number  $\alpha$ 

$$(\alpha A)_{nn'} = \alpha A_{nn'} \,. \tag{V.4}$$

The addition of physical quantities being defined as above, it is clear that the following laws hold;

A + B = B + A (the commutative law), (V.5)

 $\alpha(A + B) = \alpha A + \alpha B$  (the distributive law), (V.6)

$$A + (B + C) = (A + B) + C$$
 (the associative law). (V.7)

Furthermore, for any two arbitrary numbers  $\alpha$  and  $\beta$ , both

$$(\alpha + \beta)A = \alpha A + \beta B \tag{V.8}$$

$$\alpha(\beta A) = (\alpha \beta) A \tag{V.9}$$

are obviously true.

Recipe (VI): The product of two physical quantities A and B, i.e., the quantity AB, should be defined, in conformity with Eqs. (32.12') and (32.13') of the last section, by

$$AB = \begin{bmatrix} \sum_{n} A_{1n}B_{n1} & \sum_{n} A_{1n}B_{n2} & \sum_{n} A_{1n}B_{n3} \dots \\ \sum_{n} A_{2n}B_{n1} & \sum_{n} A_{2n}B_{n2} & \sum_{n} A_{2n}B_{n3} \dots \\ \sum_{n} A_{3n}B_{n1} & \sum_{n} A_{3n}B_{n2} & \sum_{n} A_{3n}B_{n3} \dots \end{bmatrix}.$$
 (VI.1)

Or, more briefly,

$$(AB)_{nn'} = \sum_{n''} A_{nn''} B_{n''n'} . \qquad (VI.2)$$

This multiplication rule, and the addition rule given in (V.2) are familiar ones in mathematics as the rules for matrix multiplication and addition.

Thanks to this definition of (VI), the (n, n') element of the product AB has a time dependence with the same frequency  $\nu_{nn'}$  as the corresponding element of A or B. Thus in the matrix elements of the product AB there never appear any frequencies beyond those which appear in the separate factors. Or, more generally speaking, when a

229

new physical quantity is derived as a function of the various physical quantities belonging to the same dynamical system by adding or multiplying them as defined above, the (n, n') element of the new physical quantity has the same frequency  $v_{nn'}$  as the corresponding elements of the original quantities.

As a result of the above definition of multiplication, we have to face the fact that

$$AB = BA$$
 (the commutative law) (VI.3)

does not generally hold. When, in the exceptional cases, Eq. (VI.3) holds, the two matrices A and B are said to commute. Though the commutative law does not hold, it can easily be shown that the other laws of multiplication are still valid, i.e.,

(A + B)C = AC + BC (the distributive law), (VI.4)

(AB)C = A(BC) (the associative law), (VI.5)

for arbitrary physical quantities, or matrices, A, B, and C, and in addition,

$$(\alpha A)B = A(\alpha B) \tag{VI.6}$$

for an arbitrary number  $\alpha$ .

*Recipe* (VII): Using the definition of time derivative, sum, and product as given above, insert the coordinates and their time derivatives into the equation of motion characteristic of the given dynamical system.

The equation of motion here is given in the form of a differential equation for the coordinate which results from equating to zero a function of the coordinate and its time derivatives, velocity or acceleration. The construction of this function, the functional form of which is determined by the structure of the given dynamical system, is carried out in accordance with the above defined time differentiation and the algebraical operations. The statement that a matrix is equal to zero is to be understood to mean that all its matrix elements are zero. Since our definition of the equality of two matrices was naturally the simultaneous equality of all of their corresponding matrix elements, the meaning of the above statement becomes more apparent, if we understand zero as a matrix the elements of which are all zero. In what follows the notation 0 will be frequently used as a matrix in this sense and we shall call it the "zero matrix".

230

сн. 5, § 33]

### MATRIX MECHANICS

Definition: The zero matrix is the one, all of whose elements are zero.

It seems appropriate to define two special matrices before we proceed to the next hypothesis. These are the unit matrix and a diagonal matrix.

*Definition*: The unit matrix is one whose diagonal elements are all unity and of which all the others, the off-diagonal elements, are zero. We denote this matrix by 1, i.e.,

$$1 = \begin{bmatrix} 1 & 0 & 0 & 0 \dots \\ 0 & 1 & 0 & 0 \dots \\ 0 & 0 & 1 & 0 \dots \\ 0 & 0 & 0 & 1 \dots \\ \vdots & \vdots & \vdots & \vdots & \vdots \end{bmatrix}.$$
 (33.9)

Or, by introducing the notation

$$\delta_{nn'} = \begin{cases} 1 & \text{for } n = n', \\ 0 & \text{for } n \neq n', \end{cases}$$
(33.10)

where, by the way, 1 on the right hand side has the usual meaning of unity, Eq. (33.9) can also be written briefly as

$$(1)_{nn'} = \delta_{nn'} \,. \tag{33.11}$$

The following equation involving the unit matrix,

$$lA = Al = A$$

holds for any matrix A. This can easily be shown by making use of Eq. (33.11) in (VI.2).

Definition: A diagonal matrix is one which has all its off-diagonal elements equal to zero. Thus, a diagonal matrix D has the form

$$D = \begin{bmatrix} D_1 & 0 & 0 & 0 & \dots \\ 0 & D_2 & 0 & 0 & \dots \\ 0 & 0 & D_3 & 0 & \dots \\ 0 & 0 & 0 & D_4 & \dots \\ \vdots & \vdots & \vdots & \vdots & \vdots \end{bmatrix},$$
(33.13)

or, in terms of the  $\delta$ -symbol of Eq. (33.10),

$$D_{nn'} = D_n \delta_{nn'} \,. \tag{33.14}$$

With these preliminary remarks in mind, we proceed to the next recipe, which is

Recipe (VIII): Assume that the momentum p, conjugate to the coordinate x, satisfies the relationship,

$$p\mathbf{x} - \mathbf{x}p = \frac{h}{2\pi i}\mathbf{1} , \qquad (\text{VIII.1})$$

where 1 on the right hand side is the unit matrix defined above.

This relationship is called the canonical commutation relation. In general, when two quantities A and B are given, the relationship specifying (AB-BA) is called the commutation relationship for A and B.

A little more explanation may be necessary concerning recipe (VIII). We derived Eq. (32.15') in the last section from the decoding of the quantum condition  $\oint p dx = nh$ . In our new notation, it is,

$$\frac{h}{2\pi i} = \sum_{n'} p_{nn'} x_{n'n} - \sum_{n'} x_{nn'} p_{n'n} . \qquad (VIII.2)$$

According to our definition of matrix multiplication (VI.1),  $\sum_{n'} p_{nn'} x_{n'n}$  is the (n, n) element of the product px while  $\sum_{n'} x_{nn'} p_{n'n}$  is that of the product xp, so that (VIII.2) can be written as

$$\frac{h}{2\pi i} = (px)_{nn} - (xp)_{nn}$$
. (VIII.3)

We now add to this relationship derived from the quantum condition  $\oint p dx = nh$ , a new hypothesis

$$0 = (px)_{nn'} - (xp)_{nn'} \quad \text{for } n \neq n'. \quad (VIII.4)$$

Combining (VIII.3) and (VIII.4), we obtain

$$(px)_{nn'} - (xp)_{nn'} = \frac{h}{2\pi i} \delta_{n,n'}$$

which is nothing but (VIII.1).

In the case of the harmonic oscillator treated in the last section, only Recipes (I) through (VII) together with (VIII.3) were necessary for obtaining all the required quantities, except for the undetermined phase factor which multiplies the non-vanishing matrix element X(n; n-1). But this is an exceptional case and in more general cases Eq. (VIII.4), in addition to these recipes, is necessary for determining the matrix elements of physical quantities. This is because we obtain the same number of equations as that of the unknowns only after we include the relationships (VIII.4).

The justification for this additional recipe is then to be obtained by observing the correctness of the conclusions derived from it.

When the dynamical system under consideration has more degrees of freedom and, hence, a number of coordinates,  $x_1, x_2, x_3, \ldots, x_f$  are involved, we need the following assumption. Namely,

*Recipe* (IX): The momenta  $p_1, p_2, p_3, \ldots$ , and  $p_f$  conjugate to the coordinates  $x_1, x_2, x_3, \ldots$ , and  $x_f$ , respectively, should satisfy the relations

$$\begin{aligned} x_s x_{s'} &- x_{s'} x_s = 0 \\ p_s p_{s'} &- p_{s'} p_s = 0 \\ x_s p_{s'} &- p_{s'} x_s = 0 \\ p_s x_s &- x_s p_s = \frac{h}{2\pi i} 1. \end{aligned}$$
 (IX.1)

In other words, any two coordinates, any two momenta, and a coordinate with one of the momenta conjugate to other coordinates, all commute, while a coordinate and its conjugate momentum do not commute but satisfy the canonical commutation relation.

The simplest example of an actual calculation on the basis of these recipes is the case of the oscillator explained in the last section. A summary of the procedures used there is as follows.

From the matrix of the coordinate x we calculate that of the acceleration by the aid of (IV), i.e.,

$$\left(\frac{\mathrm{d}^2 x}{\mathrm{d}t^2}\right)_{nn'} = (2\pi \mathrm{i})^2 (v_{nn'})^2 x_{nn'} . \qquad (33.15)$$

We next derive the matrix of the force

$$f_{nn'} = -m \left\{ (2\pi\nu)^2 x + \lambda x^2 \right\}_{nn'} \\ = -m \left\{ (2\pi\nu)^2 x_{nn'} + \lambda \sum_{n''} x_{nn''} x_{n''n''} \right\}, \qquad (33.16)$$

where (V) and (VI) have been used in producing the last expression. These expressions are then introduced into the equation of motion

$$m \frac{\mathrm{d}^2 x}{\mathrm{d}t^2} = f \tag{33.17}$$

to obtain the relations between the matrix elements

$$(2\pi i)^2 v_{nn'}^2 x_{nn'} = - (2\pi v)^2 x_{nn'} - \lambda \sum_{n''} x_{nn''} x_{n''n'} . \qquad (33.17')$$

The number of these relations is equal to the number of all possible combinations of n and n'.

In addition, we have the relations due to the quantum condition (VIII), i.e.,

$$2\pi im \sum_{n''} \{ v_{nn''} x_{nn''} x_{nn''} - v_{n''n'} x_{nn''} x_{nn''n'} \} = \frac{\hbar}{2\pi i} \delta_{nn'}. \qquad (33.18)$$

The number of these relations is also equal to the number of all possible combinations of n and n'.

Therefore, the number of available equations is twice the number of all possible combinations of n and n' and is equal to the number of the unknowns  $v_{nn'}$  and  $x_{nn'}$ , which is also twice that number, so that we can determine these unknowns uniquely except for the undetermined phase factors which, however, have no effect on the physical results as will be explained towards the end of this chapter. In the actual calculation in the last section, by the way, we put  $\lambda = 0$ . Next, in the example there, we introduced the solutions thus obtained into the expression for the energy

$$W = \frac{1}{2}m\dot{x}^{2} + \frac{1}{2}m(2\pi\nu)^{2}x + \frac{1}{3}\lambda x^{3}, \qquad (33.19)$$

also with  $\lambda = 0$ , to obtain the matrix of the energy W. We then found all of its matrix elements equal to zero except for the non-vanishing elements, n = n', i.e., the transition components of  $\tau = 0$  in the terminology used there. That is, in our new notation

$$W_{nn'} = W_n \delta_{nn'} \,. \tag{33.20}$$

In other words, the energy matrix had become a diagonal matrix. The  $W_n$ 's in Eq. (33.20) are all real numbers and independent of the time.

The fact that only time-independent matrix elements of the energy had non-vanishing magnitudes was interpreted as meaning that the energy had a certain constant value in a stationary state. This interpretation corresponds to the fact that, in the classical theory, the energy as a function of time has only Fourier components which are independent of time. Thus the  $W_n$  on the right hand side of Eq. (33.20) has been given the meaning of the energy value of the dynamical system in the *n*-th stationary state.

Furthermore, the  $W_n$ 's calculated in the last section, together with the  $v_{nn'}$  already obtained have been shown to satisfy Bohr's relation

$$\boldsymbol{v}_{nn'} = \frac{W_n - W_{n'}}{h}. \tag{33.21}$$

This then is a summary of the calculations of the last section. In more general cases, however, calculations of this kind become quite involved. This is because solutions of the simultaneous equations, corresponding to Eqs. (33.17') and (33.18), for more general cases become extremely difficult to obtain. Under such circumstances it is difficult to see by direct calculation whether Bohr's relationship (33.21) actually holds or not.

However, without a proof that this relationship is in fact fulfilled generally with the assumptions introduced in Recipes (I) to (IX), we can not claim with certainty that the theory is valid. In order to obtain this proof we have to make an additional proposition that, in matrix mechanics also, the equations of motion can be put in the form of the canonical equations of Hamilton. When the equations of motion are put in this form, a new approach becomes possible for obtaining the solutions by integrating them.

# § 34. The Canonical Equations of Motion and Bohr's Frequency relation

In the last section we enumerated Recipes (I) to (IX) as the basis of calculations in Heisenberg's theory. According to these recipes we obtain relationships between matrix elements from the given equation of motion (33.17). Furthermore, in calculating the energy of the system, the matrix representing the energy, as a function of coordinates and velocities, was assumed to have the same functional form as in the classical theory. Namely, it was assumed that the energy of the system and the force appearing in the equation of motion were matrices related to each other in the same manner as in classical mechanics. Generalizing this assumption, we introduce

Recipe (X): When the energy of a system as a function of the coordinate x and the momentum p is given by

$$H(x, p) , \qquad (X.1)$$

which is a real-number-like matrix, and hence hermitian, we assume that the equations of motion of the system are given by the canonical equations due to Hamilton, i.e. by

$$\frac{\mathrm{d}x}{\mathrm{d}t} = \frac{\partial H}{\partial p} \tag{X.2}$$

$$\frac{\mathrm{d}p}{\mathrm{d}t} = -\frac{\partial H}{\partial x} \,.$$

We call this H the Hamiltonian of the system as we did in classical mechanics. The Hamiltonian H may sometimes involve the time explicitly but in this section we deal with only those dynamical systems which have Hamiltonians which do not involve the time explicitly, i.e., only conservative systems.

When Recipe (X) is introduced, we can give a general proof of the facts explained at the end of the last section, that the energy matrix becomes a time independent diagonal matrix and that Bohr's relationship (33.21) actually holds. The reason why these are true for the example of § 32 is that the equation of motion and the energy used in that example happened to satisfy assumption (X).

In what follows we give a general proof but, for the sake of simplicity, shall restrict ourselves to the case of a one-dimensional dynamical system.

Before we start the actual proof, we must define the partial derivative appearing in (X.2) and give a few other mathematical preliminaries.

## (i) MATHEMATICAL PRELIMINARIES

The definition of partial differentiation is as follows.

Definition: When f(x, y, z, ...), a matrix, is given as a function of the variables, x, y, z, ... which are also matrices but not necessarily

236
# CH. 5, § 34] THE CANONICAL EQUATIONS OF MOTION

cartesian coordinates, the partial derivative with respect to x of this function f, i.e.  $\partial f/\partial x$ , is defined by

$$\frac{\partial f}{\partial x} = \lim_{a \to 0} \frac{f\left(x + a1, y, z, \ldots\right) - f\left(x, y, z, \ldots\right)}{a}, \qquad (34.1)$$

where a is an ordinary number and 1 is the unit matrix. Here by a function f(x, y, z, ...) is meant, for the sake of definiteness, the matrix defined by an equation involving summations and multiplications of variable matrices x, y, z, ... with each other or with ordinary numbers, such as  $f = c_0 + c_1 x + c_2 y + ... + c_{12} x y + c_{13} x z + ..., c_0, c_1, c_2$ , etc. being ordinary numbers. The number of terms may also be infinite provided that certain convergence conditions are satisfied. Sometimes this definition is too narrow; it is so, for example, in the case where we have to deal with the Coulomb potential 1/r, but we shall be content with this simple definition in the discussions to follow. A more general definition will be given in the later volume.

It is obvious from the definition (34.1) that

$$\frac{\mathrm{d}x}{\mathrm{d}x} = 1 , \qquad (34.2)$$

where the notation d/dx is used instead of  $\partial/\partial x$  since there is only one variable here. Furthermore,  $dx^2/dx$  is given by

$$\begin{pmatrix} \frac{\mathrm{d}x^2}{\mathrm{d}x} \end{pmatrix}_{nn'} = \lim_{a \to 0} \frac{1}{a} \{ (x+a1)^2 - x^2 \}_{nn'} \\ = \lim_{a \to 0} \frac{1}{a} \{ \sum_{n''} (x+a1)_{nn''} (x+a1)_{n''n'} - \sum_{n''} x_{nn''} x_{n''n'} \} \\ = \lim_{a \to 0} \frac{1}{a} \{ \sum_{n''} (x_{nn''} x_{n''n'} + a \delta_{nn''} x_{n''n'} + a x_{nn''} \delta_{n''n'} + a \delta_{nn''} \delta_{n''n'} + a \delta_{nn''} x_{n''n'} \} \\ = 2x_{nn'},$$

which proves the relation

$$\frac{\mathrm{d}x^2}{\mathrm{d}x} = 2x \,. \tag{34.3}$$

In the same way, we can also prove that

$$\frac{\mathrm{d}x^n}{\mathrm{d}x} = nx^{n-1} \,. \tag{34.4}$$

When two matrix functions f and g are given, the following theorem holds concerning the derivatives of their sum and their product, just as for ordinary functions.

Theorem: When two matrix functions f and g are given as functions of the matrix variables  $x, y, z, \ldots$ , we have

$$\frac{\partial}{\partial x}(f+g) = \frac{\partial f}{\partial x} + \frac{\partial g}{\partial x},$$
 (34.5)

and

$$\frac{\partial}{\partial x}(fg) = f \cdot \frac{\partial g}{\partial x} + \frac{\partial f}{\partial x} \cdot g . \qquad (34.6)$$

An important precaution is necessary concerning the order of the multiplying factors in Eq. (34.6), i.e., we have to preserve the same order on the right hand side as on the left. For instance, it is not allowed to change the equation into

$$\frac{\partial}{\partial x}(f \cdot g) = f \cdot \frac{\partial g}{\partial x} + g \cdot \frac{\partial f}{\partial x} \, .$$

The proof of this theorem is quite easy and is reserved for the reader.

The partial derivative of a matrix function being defined as above, we have the following important theorem. That is,

Theorem: When a matrix f(x, p) is given as a function of the coordinate and the momentum p which satisfy  $px - xp = (h/2\pi i)$  1, we have

$$\frac{\partial f}{\partial x} = \frac{2\pi i}{h} (pf - fp)$$

$$\frac{\partial f}{\partial p} = -\frac{2\pi i}{h} (xf - fx) .$$
(34.7)

*Proof*: If we first take p itself as f, we can easily verify Eq. (34.7) for this choice of f. Namely, the first equation of (34.7) gives  $\partial p/\partial x = 0$  while the second becomes  $\partial p/\partial p = 1$  due to Recipe (VIII.1) and these are what we should have for the derivatives. The validity of the theorem for the case of x as f can be verified in an analogous manner. The next step is to show that when the theorem holds for two functions  $f_1$  and  $f_2$  separately it holds also for their sum  $(f_1 + f_2)$ . This can

and

# CH. 5, § 34] THE CANONICAL EQUATIONS OF MOTION

readily be done by making use of the fact that the distribution law holds also for the multiplication of matrices, i.e.,

$$p(f_1 + f_2) = pf_1 + pf_2,$$
  

$$(f_1 + f_2)p = f_1p + f_2p,$$

so that

$$\frac{2\pi i}{h} \left\{ p(f_1 + f_2) - (f_1 + f_2)p \right\} = \frac{2\pi i}{h} (pf_1 - f_1p) + \frac{2\pi i}{h} (pf_2 - f_2p)$$
$$= \frac{\partial f_1}{\partial x} + \frac{\partial f_2}{\partial x} = \frac{\partial}{\partial x} (f_1 + f_2).$$

By a similar argument we can also prove

$$-\frac{2\pi i}{h} \{ x(f_1 + f_2) - (f_1 + f_2)x \} = \frac{\partial}{\partial p} (f_1 + f_2) .$$

We can further prove that, when Eq. (34.7) holds for  $f_1$  and  $f_2$  separately, it holds also for their product  $(f_1f_2)$ . Namely from Eq. (34.6) we obtain

$$\frac{\partial (f_1 f_2)}{\partial x} = \frac{\partial f_1}{\partial x} f_2 + f_1 \frac{\partial f_2}{\partial x},$$

which, by the use of the assumption that Eq. (34.7) holds for  $f_1$  and  $f_2$  separately, becomes

$$=\frac{2\pi i}{h}\left\{\left(pf_{1}-f_{1}p\right)f_{2}+f_{1}(pf_{2}-f_{2}p)\right\},\$$

and with the help of the distribution law, this expression reduces further to

$$\frac{2\pi i}{h} \{ pf_1 f_2 - f_1 pf_2 + f_1 pf_2 - f_1 f_2 p \} = \frac{2\pi i}{h} \{ p(f_1 f_2) - (f_1 f_2)p \} .$$

We have thus verified that the first equation of (34.7) holds for the product  $(f_1 f_2)$ . The second equation of (34.7) can also be proved in the present case. Summarizing the results obtained so far, the theorem

239

holds for the choice of x or p as f, and also for the sum  $(f_1 + f_2)$  and the product  $(f_1 f_2)$  provided that it is valid both for  $f_1$  and  $f_2$ separately. This implies that the theorem holds for any polynomials of x and p or more generally that it holds for any power series of xand p provided that uniform convergence is guaranteed. We have hence arrived at the result that the theorem (34.7) holds for any function f of x and p.

Turning now to the time derivatives, we have the following theorem.

Theorem: For two given physical quantities f and g, which are matrix functions, we have

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(f+g\right) = \frac{\mathrm{d}f}{\mathrm{d}t} + \frac{\mathrm{d}g}{\mathrm{d}t}\,.\tag{34.8}$$

and

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(f\cdot g\right) = f \cdot \frac{\mathrm{d}g}{\mathrm{d}t} + \frac{\mathrm{d}f}{\mathrm{d}t} \cdot g \,. \tag{34.9}$$

The first of these equations can readily be proved and we shall not reproduce the proof here. The second equation can be verified in the following way. The (n, n') element of the left hand side of the equation is, due to (IV.4), given by

$$\left(\frac{\mathrm{d}(f \cdot g)}{\mathrm{d}t}\right)_{nn'} = 2\pi \mathrm{i} \nu_{nn'} (f \cdot g)_{nn'} = 2\pi \mathrm{i} \nu_{nn'} \sum_{n''} f_{nn''} g_{n''n''} .$$

On the other hand, the (n, n') element of the right hand side is given by

$$\left( f \cdot \frac{\mathrm{d}g}{\mathrm{d}t} + \frac{\mathrm{d}f}{\mathrm{d}t} \cdot g \right)_{n\,n'} = \sum_{n''} 2\pi \mathrm{i} v_{n''\,n'} f_{n\,n''} g_{n''\,n'} + \sum_{n''} 2\pi \mathrm{i} v_{n\,n''} f_{n\,n''} g_{n''\,n'} = 2\pi \mathrm{i} \sum_{n''} \left( v_{n\,n''} + v_{n''\,n'} \right) f_{n\,n''} g_{n''\,n'} ,$$

which, by the use of the combination law

 $\nu_{nn''} + \nu_{n''n'} = \nu_{nn'},$ 

becomes

$$= 2\pi i \nu_{nn'} \sum_{n''} f_{nn''} g_{n''n'} \, .$$

This is identical with the expression given above for the (n, n')element of the left hand side and the theorem is proved.

We have now completed the necessary mathematical preliminaries and can proceed with the physical discussion.

# (ii) THE LAW OF ENERGY CONSERVATION, AND BOHR'S RELATIONSHIP

We shall now return to our main subject with the mathematical theorems derived in the last sub-section. Our problem is to prove that when we substitute for x and p any solution of the equation of motion, the Hamiltonian function H can be brought into the form of a diagonal matrix and that Bohr's relationship holds between the diagonal elements of this matrix and the frequency.

To begin with, we first note the fact that the canonical equation of motion (X.2) can be written as

$$\frac{\mathrm{d}x}{\mathrm{d}t} = -\frac{2\pi\mathrm{i}}{h} (xH - Hx) ,$$

$$\frac{\mathrm{d}p}{\mathrm{d}t} = -\frac{2\pi\mathrm{i}}{h} (pH - Hp) ,$$
(34.10)

where the theorem (34.7) has been used.

We next prove that for any arbitrary function g of x and p, it is true that

$$\frac{\mathrm{d}g}{\mathrm{d}t} = -\frac{2\pi\mathrm{i}}{h} \left(gH - Hg\right), \qquad (34.11)$$

provided that g does not depend on t explicitly. The proof of the theorem is quite similar to that of Eq. (34.7). That is, when g is x, or p, Eq. (34.11) is nothing but the canonical equation of motion which is already verified. We then prove that, when it holds for both  $g_1$  and  $g_2$ separately, it holds also for the sum  $(g_1 + g_2)$  and the product  $(g_1 g_2)$ . The proof of this fact is analogous to that of Eq. (34.7). That is, the distribution law.

$$(g_1 + g_2)H = g_1H + g_2H$$
,  
 $H(g_1 + g_2) = Hg_1 + Hg_2$ ,

and

$$H(g_1 + g_2) = Hg_1 + Hg_2$$
,

guarantees that Eq. (34.11) holds for the sum  $(g_1 + g_2)$ . To show the validity of the theorem for the product  $(g_1 g_2)$  we start with the use of Eq. (34.9), i.e.,

$$\frac{\mathrm{d}(g_1g_2)}{\mathrm{d}t} = g_1 \cdot \frac{\mathrm{d}g_2}{\mathrm{d}t} + \frac{\mathrm{d}g_1}{\mathrm{d}t} \cdot g_2 ,$$

which, on using the hypothesis that the theorem holds for both  $g_1$  and  $g_2$  separately, can be transformed to

$$= -\frac{2\pi i}{h} \{ g_1(g_2H - Hg_2) + (g_1H - Hg_1)g_2 \}$$
$$= -\frac{2\pi i}{h} \{ (g_1g_2)H - H(g_1g_2) \}.$$

This implies that it holds also for the product  $(g_1 g_2)$ . Since we have proved that the theorem holds for x and p, and that it holds for the sum  $(g_1 + g_2)$ , and for the product  $(g_1 g_2)$ , provided that both  $g_1$  and  $g_2$  satisfy Eq. (34.11), we arrive at the conclusion that the theorem Eq. (34.11) holds for any function g of x and p.

As a special case of the theorem, we take H itself for g in Eq. (34.11). We then immediately arrive at the result that

$$\frac{\mathrm{d}H}{\mathrm{d}t} = -\frac{2\pi \mathrm{i}}{h} \left(HH - HH\right) = 0 , \qquad (34.12)$$

or, in terms of the matrix element, that

$$2\pi i v_{nn'} H_{nn'} = 0. \qquad (34.12')$$

Since generally  $v_{nn'} \neq 0$  for  $n \neq n'$ , Eq. (34.12') requires that

$$H_{nn'} = 0$$
 for  $n \neq n'$ ,

while for n = n',  $\nu_{nn} = 0$  and hence, the  $H_{nn}$ 's do not have to vanish. Therefore, we arrive at the conclusion that the matrix H must be a diagonal matrix; i.e.,

$$H_{nn'} = W_n \delta_{nn'} , \qquad (34.13)$$

where the diagonal elements  $W_n$ 's do not depend on time. In certain cases, however, there can also be some  $v_{nn'}$ 's of  $n \neq n'$  which happen to vanish and the matrix elements  $H_{nn'}$ , corresponding to these do not have to vanish. The matrix H in such cases is not necessarily reduced to a diagonal matrix but it is still independent of time. The interpretation of such cases will be postponed to a later section and we shall in this section restrict ourselves to the simplest case where the

matrix H does become a diagonal matrix. The Hamiltonian function H of the system is now a diagonal matrix, i.e., only its diagonal elements have non-vanishing, and at the same time time-independent, values. Denoting the (n, n) element by  $W_n$ , we interpret this as the energy value of the system in the *n*-th stationary state. This is the law of the conservation of energy which was to be proved.

We shall now derive Bohr's relation Eq. (33.21) in a general way. For this purpose, we start with Eq. (34.11), which in terms of the matrix element is written as

$$2\pi i v_{nn'} g_{nn'} = -\frac{2\pi i}{h} \left\{ \sum_{n''} g_{nn''} H_{n''n'} - \sum_{n''} H_{nn''} g_{n''n'} \right\}.$$

This in turn becomes

$$=-\frac{2\pi \mathrm{i}}{h}g_{nn'}\left(W_{n'}-W_{n}\right),$$

where Eq. (34.13) has been used. Since this holds for any arbitrary function g, the relationship

$$h v_{nn'} = W_n - W_{n'}$$

has to hold in general and we have proved Bohr's relation.

For our later convenience, we shall give the previously obtained solution x of the harmonic oscillator expressed in our new notation. From Eq. (32.21') together with Eq. (32.23'), we have

$$x_{nn-1} = \sqrt{\frac{h}{8\pi^2 \nu m}} \sqrt{n} \exp(2\pi i \nu t + i \delta_n) ,$$
  

$$x_{n-1n} = \sqrt{\frac{h}{8\pi^2 \nu m}} \sqrt{n} \exp(-2\pi i \nu t - i \delta_n)$$
(34.14)  
for  $n = 0, 1, 2, ...,$ 

and

$$x_{nn'} = 0 \qquad \qquad \text{for } n' \neq n \pm 1 .$$

In the meantime, Eq. (32.29') becomes, in the new notation,

$$W_{nn'} = h\nu(n + \frac{1}{2})\delta_{nn'}. \qquad (34.15)$$

Furthermore, the use of the relation p = mx gives for the momentum matrix

$$p_{nn-1} = -\frac{1}{i} \sqrt{\frac{1}{2}hm\nu} \sqrt{n} \exp(2\pi i\nu t + i\delta_n) ,$$

$$p_{n-1n} = \frac{1}{i} \sqrt{\frac{1}{2}hm\nu} \sqrt{n} \exp(-2\pi i\nu t - i\delta_n) \qquad (34.16)$$
for  $n = 0, 1, 2, 3, ...,$ 

$$p_{i} = 0$$
 for  $n' \neq n+1$ 

The validity of the relationship

$$px - xp = \frac{h}{2\pi i} l$$

can in fact be confirmed for the above derived expressions (34.14) and (34.16). The undetermined phase angle  $\delta_n$  appearing in Eqs. (34.14) and (34.16) will be shown in the next section to remain undetermined in principle. Since however the physically meaningful quantities, the transition probability for instance, are given in terms only of the absolute squares of the matrix elements, this indeterminacy of the matrix elements does not give rise to any difficulty.

# § 35. The Eigenvalue Problem

When the theory is given in the Hamiltonian formalism rather than that of Eqs. (33.17) and (33.18), we can solve the problem by utilizing more familiar types of mathematical equations. Both Eqs. (33.17) and (33.18) are not linear in the unknown variables  $x_{nn'}$ , i.e., Eq. (33.18) is always of second order while Eq. (33.17) is so since the force acting on the system contained  $x^2$  as the highest order term. If the force happens to involve still higher order terms in x, the equation we have to solve will become of still higher order. The solutions of such higher order equations are obviously quite complicated. In this section, we shall show that the use of the Hamiltonian formalism can reduce the problem to solving simpler equations which are generally linear in the unknown quantities. The discussion may become somewhat mathematical but it will be of advantage for the reader to become familiar with the concepts in this section, since these will be necessary for the later develop-

244

# сн. 5, § 35]

ments of quantum mechanics. The whole argument will be given in several steps.

# (i) THE INVERSE OF THE LAW OF ENERGY CONSERVATION

Our problem is to find the Hermitian matrices x and p which, assuming that their (n, n') elements oscillate in the manner of  $\exp(2\pi i \nu_{nn'} t)$ , satisfy

$$\frac{\mathrm{d}x}{\mathrm{d}t} = -\frac{2\pi \mathrm{i}}{h} \left( xH - Hx \right) ,$$

$$\frac{\mathrm{d}p}{\mathrm{d}t} = -\frac{2\pi \mathrm{i}}{h} \left( pH - Hp \right) ,$$
(35.1)

together with the commutation relationship

$$1 = \frac{2\pi i}{h} \left( px - xp \right) , \qquad (35.2)$$

for a given, real-number-like, Hamiltonian function H(x, p) of coordinate x and momentum p.

According to the theorem proved in the last section, the solutions x and p of Eqs. (35.1) and (35.2), if they can be obtained, will reduce the matrix H(x, p) to a diagonal matrix of the form

$$H_{nn'} = W_n \,\delta_{nn'} \tag{35.3}$$

and, furthermore, the frequencies  $v_{nn'}$  involved in x and p will satisfy Bohr's relation

$$v_{nn'} = \frac{1}{h} \left( W_n - W_{n'} \right) \,.$$
 (35.4)

We shall now show that in a certain sense the inverse of this theorem also holds. Namely:

Theorem: If, by some way or other, two Hermitian matrices X and P whose matrix elements are all time-independent can be found such that they satisfy

$$1 = \frac{2\pi i}{h} \left( PX - XP \right) \tag{35.2'}$$

and that, when substituted for X and P in H(X, P), they make the Hamiltonian diagonal, i.e.,

$$H(X, P)_{nn'} = W_n \,\delta_{nn'}, \qquad (35.3')$$

then the matrices x and p satisfying both the equation of motion (35.1) and the commutation relation (35.2) can be obtained by the prescription

$$x_{nn'} = X_{nn'} \exp(2\pi i \nu_{nn'} t) , \qquad (35.5)$$
  
$$p_{nn'} = P_{nn'} \exp(2\pi i \nu_{nn'} t) ,$$

where  $v_{nn'}$  is defined by

$$v_{nn'} = \frac{1}{h} \left( W_n - W_{n'} \right) \,.$$
 (35.6)

The statement in the theorem that X and P are independent of time is not quite the same as the similar statement in the last section that the energy W is independent of time. For P and X have to satisfy neither the assumption II nor the equation of motion and hence they do not necessarily become diagonal matrices as the energy does. The statement here that P and X are independent of time merely implies that their matrix elements are all constants.

For the proof of this theorem we need the following preliminary theorem. That is, for any function f(x, p) of x and p we have

$$f(x, p)_{nn'} = f(X, P)_{nn'} \exp(2\pi i \nu_{nn'} t) , \qquad (35.7)$$

provided that the quantities in the expression satisfy the relations (35.5) and (35.6). The proof of this preliminary theorem is not difficult. First, for the case of f = x or p it reduces to Eq. (35.5) which obviously holds by definition. We then prove that it holds for the sum (f + g) and the product (fg) provided f and g themselves satisfy Eq. (35.7) separately. What we prove is

$$\{f(x, p) + g(x, p)\}_{nn'} = \{f(X, P) + g(X, P)\}_{nn'} \exp(2\pi i \nu_{nn'} t) \quad (35.8)$$
  
and

$$\{f(x, p)g(x, p)\}_{nn'} = \{f(X, P)g(X, P)\}_{nn'} \exp(2\pi i \nu_{nn'} t) .$$
(35.9)

The first of these equations is almost self-obvious and the left hand side of the second is transformed to

$$\sum_{n''} f(x, p)_{nn''} g(x, p)_{n''n'} = \sum_{n''} f(X, P)_{nn''} \exp(2\pi i \nu_{nn''} t) \\ \cdot g(X, P)_{n''n'} \exp(2\pi i \nu_{n''n'} t) ,$$

сн. 5, § 35]

which upon the use of (35.6) for  $v_{nn''}$  and  $v_{n''n'}$ , and hence  $v_{nn'} + v_{n'n''} = v_{nn''}$ , becomes

$$= \{ \sum_{n''} f(X, P)_{nn''} g(X, P)_{n''n'} \} \exp(2\pi i \nu_{nn'} t)$$
  
=  $\{ f(X, P)g(X, P) \}_{nn'} \exp(2\pi i \nu_{nn'} t) .$ 

The preliminary theorem (35.7) is now proved for the product fg. Thus it holds for x and p, and for the sum (f + g) and the product fg when f and g satisfy it separately. By a frequently used line of argument we can conclude that Eq. (35.7) holds for any function of x and p.

It can easily be shown by the aid of this preliminary theorem that x and p of Eq. (35.5) satisfy Eq. (35.2). For, by the preliminary theorem,

$$(xp - px)_{nn'} = (XP - PX)_{nn'} \exp(2\pi i \nu_{nn'} t) .$$

The right hand side reduces, by the assumption on X and P, to

$$\frac{h}{2\pi \mathrm{i}} \delta_{nn'} \exp\left(2\pi \mathrm{i} v_{nn'} t\right) ,$$

which, since  $\delta_{nn'} = 0$  for  $n \neq n'$ , in turn can also be written as

$$\frac{h}{2\pi \mathrm{i}} \delta_{nn'} \exp\left(2\pi \mathrm{i} \nu_{nn} t\right) \,.$$

The frequency  $v_{nn} = 0$  in view of its definition (35.5) and hence

$$(xp - px)_{nn'} = \frac{h}{2\pi i} \delta_{nn'}.$$

We shall now show that x and p satisfy the equation of motion (35.1). Substituting Eq. (35.5) into the left hand side of Eq. (35.1) we have

$$\begin{pmatrix} \frac{\mathrm{d}x}{\mathrm{d}t} \end{pmatrix}_{nn'} = 2\pi \mathrm{i}\nu_{nn'} X_{nn'} \exp\left(2\pi \mathrm{i}\nu_{nn'} t\right)$$
  
=  $2\pi \mathrm{i}\nu_{nn'} x_{nn'}$ , (35.10)

while the right hand side becomes, after using the preliminary theorem,

$$\{ xH(x, p) - H(x, p)x \}_{nn'}$$

$$= \{ XH(X, P) - H(X, P)X \}_{nn'} \exp(2\pi i \nu_{nn'} t)$$

$$= \sum_{n''} \{ X_{nn''} H(X, P)_{n''n'} - H(X, P)_{nn''} X_{n''n'} \} \exp(2\pi i \nu_{nn'} t) .$$

The matrix H(X, p) however is by assumption a diagonal matrix, i.e.,

$$H(X, P)_{nn'} = W_n \,\delta_{nn'} \,.$$

Introducing this expression for H into the above, we obtain

$$\{ xH(x, p) - H(x, p)x \}_{nn'}$$

$$= \sum_{n''} \{ X_{nn''} W_{n''} \delta_{n''n'} - W_n \delta_{nn''} X_{n''n'} \} \exp(2\pi i \nu_{nn'} t)$$

$$= (W_{n'} - W_n) X_{nn'} \exp(2\pi i \nu_{nn'} t)$$

$$= -h \nu_{nn'} x_{nn'} .$$
(35.11)

Comparing Eq. (35.11) with Eq. (35.10), we arrive at the conclusion that

$$\left(\frac{\mathrm{d}x}{\mathrm{d}t}\right)_{nn'} = -\frac{2\pi \mathrm{i}}{h} \left\{ xH(x,p) - H(x,p)x \right\}_{nn'}.$$

We have thus proved that the x of Eq. (35.5) satisfies the equation of motion (35.1) and the proof for p can be obtained in an analogous way.

As a result of this inverse theorem, our problem of solving the equation of motion is reduced to finding the time independent Hermitian matrices X and P which satisfy the canonical commutation relation (35.2) and which, in addition, make the Hamiltonian a diagonal matrix. Obviously this does not mean very much if the problem of finding these matrices X and P is not easier than the original problem itself. But the problem in this reduced form is considerably more familier in mathematics than it is in the original form.

### (ii) THE UNITARY TRANSFORMATION

We start with the definition of unitarity.

Definition: A matrix U is unitary when its matrix elements satisfy

$$\sum_{n''} U_{nn''} U_{n'n'}^{\bullet} = \delta_{nn'},$$

$$\sum_{n''} U_{n''n}^{\bullet} U_{n''n'} = \delta_{nn'}.$$
(35.12)

and

Definition: The Hermitian conjugate matrix of a matrix A is the

one whose (n, n') element is the complex conjugate of the (n', n) element of A and is denoted by  $A^{\dagger}$ , i.e.,

$$(A^{\dagger})_{nn'} = A^*_{n'n} \,. \tag{35.13}$$

In other words, in order to obtain the Hermitian conjugate  $A^{\dagger}$  of a matrix A we change the roles of row and column in A and then take the complex conjugate of each element.

In terms of this definition, the previously mentioned Hermitian matrix is one which satisfies

$$A^{\dagger} = A , \qquad (35.14)$$

and the unitarity of a matrix U can also be expressed by

$$U^{\dagger}U = UU^{\dagger} = 1.$$
 (35.15)

We introduce another definition.

Definition: When a matrix B exists for a given A which satisfies AB = BA = 1, we call B the inverse matrix of A and denote it by  $A^{-1}$ . A unitary matrix U can now be defined also by

$$U^{\dagger} = U^{-1}.$$

Definition: The procedure of obtaining, from a given matrix X and a unitary matrix U, the matrix X' defined by

$$X' = U^{\dagger} X U \tag{35.16}$$

is called the "unitary transformation of X by U".

The following theorem holds concerning unitary transformations. Theorem: When X is Hermitian, X' is also Hermitian. Proof: By definition,

$$X'_{nn'} = \sum_{n''n'''} U^{\bullet}_{n''n} X_{n''n'''} U_{n'''n'}.$$

Taking the complex conjugate of the equation and changing the dummy indices n'' to n''' and n''' to n'', we have

$$X_{nn'}^{\prime \bullet} = \sum_{n''n'''} U_{n'''n} X_{n'''n''}^{*} U_{n''n'}^{*},$$

249

which further becomes, since X is Hermitian by assumption,

$$= \sum_{n''n'''} U_{n'''n} X_{n''n''} U_{n''n'}^{\bullet}$$
$$= \sum_{n''n'''} U_{n''n'}^{\bullet} X_{n''n'''} U_{n'''n}.$$

This last expression is the one which results from the interchange of n and n' in the expression of  $X'_{nn'}$ . Therefore, we arrive at the conclusion that

$$X_{nn'} = X_{n'n},$$

and hence the matrix X' is proved to be Hermitian.

Theorem: When a matrix f(X, Y, Z, ...) is given as a function of a number of matrices X, Y, Z, ..., then the transformed function  $f' = U^{\dagger}fU$  is equal to f(X', Y', Z', ...) or more specifically,

$$U^{\dagger}f(X, Y, Z, \ldots)U = f(U^{\dagger}XU, U^{\dagger}YU, U^{\dagger}ZU, \ldots)$$
. (35.17)

*Proof*: It is clear that the theorem holds for the sum F = f + g when both f and g satisfy (35.17). Namely,

$$U^{\dagger}FU = U^{\dagger}(f+g)U$$

becomes, due to the distribution law,

$$= U^{\dagger}(fU + gU)$$

which in turn, again by the distribution law, becomes

That is,

$$= U^{\dagger}fU + U^{\dagger}gU$$
.

 $U^{\dagger} \{ f(X, Y, Z, \ldots) + g(X, Y, Z, \ldots) \} U$ =  $U^{\dagger} f(X, Y, Z, \ldots) U + U^{\dagger} g(X, Y, Z, \ldots) U$ .

The last expression is by assumption equal to

 $f(U^{\dagger}XU, U^{\dagger}YU, U^{\dagger}ZU, \ldots) + g(U^{\dagger}XU, U^{\dagger}YU, U^{\dagger}ZU, \ldots)$ and, hence,  $F(U^{\dagger}XU, U^{\dagger}YU, U^{\dagger}ZU, \ldots)$ 

 $U^{\dagger}F(X, Y, Z, \ldots)U = F(U^{\dagger}XU, U^{\dagger}YU, U^{\dagger}ZU, \ldots).$ 

The theorem is proved also for the product F = fg on the same assumption. Namely,

$$U^{\dagger}FU = U^{\dagger}fgU$$

since  $UU^{\dagger} = 1$ , is equal to

$$U^{\dagger} f U U^{\dagger} g U$$

which by assumption is

$$= f(U^{\dagger}XU, U^{\dagger}YU, U^{\dagger}ZU, \ldots) \cdot g(U^{\dagger}XU, U^{\dagger}YU, U^{\dagger}ZU, \ldots)$$
$$= F(U^{\dagger}XU, U^{\dagger}YU, U^{\dagger}ZU, \ldots).$$

We have thus proved that

$$U^{\dagger}FU = F(U^{\dagger}XU, U^{\dagger}YU, U^{\dagger}ZU, \ldots)$$

for F = fg. It is obvious that the theorem holds for f or g which is equal to  $X, Y, Z, \ldots$ , respectively. Therefore, by a line of argument often used above the theorem holds for any function of  $X, Y, Z, \ldots$ 

Corollary: When the two matrices X and P satisfy

$$PX - XP = \frac{h}{2\pi i} l , \qquad (35.2')$$

the matrices X' and P' defined respectively by

$$X' = U^{\dagger}XU$$
 and  $P' = U^{\dagger}PU$ 

also satisfy the relationship

$$P'X' - X'P' = \frac{\hbar}{2\pi i} 1 , \qquad (35.2'')$$

and when P and X are Hermitian, so are P' and X'. The proof is very easy and is left for the reader. The fact that, when Eq. (35.2') is satisfied by the un-primed quantities, Eq. (35.2") is also satisfied by the primed quantities can sometimes be expressed by the statement that the canonical commutation relation is invariant with respect to a unitary transformation.

This completes our mathematical preparation.

#### (iii) THE EIGENVALUE PROBLEM

As explained before, our problem is to find two matrices X and P which satisfy the commutation relation

$$PX - XP = \frac{h}{2\pi i} \, 1 \tag{35.2'}$$

and which, in addition, lead to

$$H(X, P) =$$
diagonal matrix , (35.3')

where, however, P and X are supposed to be time independent in the sense explained in (i). This problem can be solved in the following way. We start with two otherwise arbitrary matrices P and X which satisfy the canonical commutation relation. For instance, we may take for such matrices P and X, the values of p and x, respectively, at t = 0 of the harmonic oscillator case. That is, by Eqs. (34.14) and (34.16),

$$X^{\mathbf{0}} = \frac{1}{2\pi} \sqrt{\frac{h}{2}} \begin{bmatrix} 0 & \sqrt{1} & 0 & 0 & 0 & \dots \\ \sqrt{1} & 0 & \sqrt{2} & 0 & 0 & \dots \\ 0 & \sqrt{2} & 0 & \sqrt{3} & 0 & \dots \\ 0 & 0 & \sqrt{3} & 0 & \sqrt{4} & \dots \\ 0 & 0 & 0 & \sqrt{4} & 0 & \dots \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & \sqrt{1} & 0 & \sqrt{2} & 0 & 0 & \dots \\ \sqrt{1} & 0 & \sqrt{2} & 0 & 0 & \dots \\ 0 & \sqrt{2} & 0 & \sqrt{3} & 0 & \dots \\ 0 & \sqrt{2} & 0 & \sqrt{3} & 0 & \dots \\ 0 & 0 & \sqrt{3} & 0 & \sqrt{4} & \dots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \sqrt{4} & 0 & \dots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \end{bmatrix}$$
(35.18)

and

The actual values of  $\nu$  and m in Eqs. (34.14) and (34.16) are irrelevant for our present discussion and have both been set equal to unity.

By  $H^0$  we denote the matrix H with p and x replaced respectively by  $P^0$  and  $X^0$ , i.e.,

$$H^{0} = H(X^{0}, P^{0}) . (35.20)$$

Naturally, the matrix  $H^0$  need not yet be a diagonal matrix. We then look for the solution of the problem in the form

$$X = U^{\dagger} X^{0} U \quad \text{and} \quad P = U^{\dagger} P^{0} U , \qquad (35.21)$$

that is, we are to determine the unitary matrix U which makes

 $H(U^{\dagger}X^{0}U, U^{\dagger}P^{0}U)$ 

a diagonal matrix. The determination of this U implies the solution of our problem since, by the corollary of (ii), X and P of Eq. (35.21) also satisfy the canonical commutation relation and since they bring the matrix H into diagonal form. The problem is further reduced, by the theorem (35.17) of (ii), to finding the unitary matrix U which leads to

$$U^{\dagger}H^{0}U =$$
diagonal matrix.

Denoting this diagonal matrix by W, our problem is to find the unitary matrix U which satisfies

$$U^{\dagger}H^{0}U = W,$$
  
$$W_{nn'} = W_{n} \,\delta_{nn'}. \qquad (35.22)$$

Or, by multiplying the first equation of (35.22) from the left by U and using  $UU^{\dagger} = 1$ ,

 $H^{0}U = UW. \qquad (35.23)$ 

The reader should recall that the order of multiplication can not be arbitrarily changed in matrix algebra.

In terms of the matrix elements, Eq. (35.23) becomes

$$\sum_{n''} H^0_{n\,n''} \, U_{n''\,n'} = \sum_{n''} U_{n\,n''} \, W_{n''} \, \delta_{n''\,n'} = U_{n\,n'} \, W_{n'} \,. \qquad (35.23')$$

The quantities  $H_{n,n'}^0$  in this equation are the matrix elements of H with x and p replaced by  $X^0$  and  $P^0$  respectively, and, hence, are all known. Therefore, the set of equations (35.23') are simultaneous equations for the unknown quantities  $U_{nn'}$  and  $W_n$ . The equations are linear and homogeneous with respect to the unknowns  $U_{nn'}$ . Among the equations, which are equal in number to all the possible combinations of n and n', those belonging to different values of n' are independent of each other. Or more specifically, only those  $U_{n'n'}$  and  $W_{n'}$  having the same value of n' are related among themselves. We are then to solve the equations

$$\sum_{n''} H^0_{nn''} \xi_{n''} - W \xi_n = 0, \qquad n = 1, 2, 3, \dots$$
 (35.24)

Equations of this type appear frequently in the theory of linear equations. Written explicitly, they are

$$(H_{1_{1}}^{0} - W)\xi_{1} + H_{1_{2}}^{0}\xi_{2} + H_{1_{3}}^{0}\xi_{3} + \dots = 0$$
  

$$H_{2_{1}}^{0}\xi_{1} + (H_{2_{2}}^{0} - W)\xi_{2} + H_{2_{3}}^{0}\xi_{3} + \dots = 0$$
  

$$H_{3_{1}}^{0}\xi_{1} + H_{3_{2}}^{0}\xi_{2} + (H_{3_{3}}^{0} - W)\xi_{3} + \dots = 0.$$
  
(35.24')

Such equations can be handled much more easily than Eqs. (33.17) and (33.18) given previously, since they are linear in the unknowns  $\xi$ 's.

As a simple illustration we consider the case where the number of possible values of n, or of n', is finite and is equal to N. The theory of linear equations then tells us that, in order for the equations to have a solution other than the trivial one  $\xi_1 = \xi_2 = \xi_3 = \ldots = 0$ , the equation

$$\begin{vmatrix} H_{11}^{0} - W & H_{12}^{0} & H_{13}^{0} & \dots \\ H_{21}^{0} & H_{22}^{0} - W & H_{23}^{0} & \dots \\ H_{31}^{0} & H_{32}^{0} & H_{33}^{0} - W \dots \end{vmatrix} = 0 \quad (35.25)$$

has to hold. This equation is of N-th order in the unknown quantity W and by solving this N-th order equation we obtain N roots for W. If the matrix  $H^0$  is real-number-like, i.e., if  $H^{0}_{nn'} = H^{0^{\bullet}}_{n'n}$ , all these roots are shown to be real. To prove this, we consider the equations which are complex conjugate to Eq. (35.24)

$$\sum_{n''} H^{0}_{n'' n} \xi^{\bullet}_{n''} - W^{\bullet} \xi^{\bullet}_{n} = 0. \qquad (35.24)^{\bullet}$$

Multiplying Eq. (35.24) by  $\xi_n^*$  and taking the sum over *n*, we obtain

$$\sum_{n\,n''} \,\xi_n^* H_{n\,n''}^0 \xi_{n''} - W \sum_n \xi_n^* \xi_n = 0 \,. \tag{35.26}$$

By a similar procedure, (35.24)\* gives

$$\sum_{n\,n''}\,\xi_{n''}^{*}H_{n''\,n}^{0}\xi_{n}-W^{*}\sum_{n}\,\xi_{n}^{*}\xi_{n}=0\,,$$

or, by interchanging the roles of n and n'' in the first summation,

$$\sum_{n\,n''} \xi_n^* H_{n\,n''}^0 \xi_{n''} - W^* \sum_n \xi_n^* \xi_n = 0 \,. \tag{35.26}^*$$

сн. 5, § 35]

Subtracting Eq. (35.26)\* from Eq. (35.26), we obtain

$$(W-W^{\bullet})\sum_{n}|\xi_{n}|^{2}=0$$
,

which immediately leads to the required result

$$W - W^* = 0$$
  
$$\sum_{n} |\xi_n|^2 \neq 0.$$

since

We shall now denote the N roots of the equation (35.25) by  $W_1, W_2, W_3, \ldots$ , and  $W_n$ . To avoid excessive complexity we assume for the time being that they are all different from each other. When these N values of W are obtained, we use one of them,  $W_n$ , for instance, in Eq. (35.24') to solve it for the unknown quantities  $(\xi_1, \xi_2, \ldots, \xi_N)$ . To specify that the solution corresponds to  $W_{n'}$ , we attach the suffix n' to it as  $(\xi_{1n'}, \xi_{2n'}, \ldots, \xi_{Nn'})$ . These  $\xi$ 's are determined to within a common multiplying factor. By chosings this factor suitably, we can satisfy

$$\sum_{n''} |\xi_{n''n'}|^2 = 1.$$
 (35.27)

When the solution is adjusted to satisfy Eq. (35.27), it is said to be normalized. We shall derive a few properties of this solution  $\xi$ . First of all,

$$\sum_{n} \xi_{nn''}^* \xi_{nn'} = 0 \qquad \text{for } n' \neq n'' . \tag{35.28}$$

Generally, when this relation holds, we say that the solution  $(\xi_{1n'}, \xi_{2n'}, \ldots, \xi_{Nn'})$  corresponding to  $W_{n'}$  is orthogonal to the solution  $(\xi_{1n''}, \xi_{2n''}, \ldots, \xi_{Nn''})$  corresponding to  $W_{n''}$  different from n'. The proof of this fact is not difficult. Since  $(\xi_{1n'}, \xi_{2n'}, \ldots, \xi_{Nn'})$  is the solution of (35.24) corresponding to  $W_{n'}$ , we have

$$\sum_{n'''} H^{0}_{n\,n'''} \xi_{n'''\,n'} - W_{n'} \xi_{n\,n'} = 0 \; .$$

Multiplying this equation by  $\xi_{nn}^*$  and summing over *n*, we obtain

$$\sum_{n n'''} \xi_{n n''}^* H_{n n'''}^0 \xi_{n''' n'} - W_{n'} \sum_{n} \xi_{n n''}^* \xi_{n n'} = 0.$$
 (35.29)

On the other hand,  $(\xi_{1n''}^{\bullet}, \xi_{2n''}^{\bullet}, \ldots, \xi_{Nn''}^{\bullet})$  satisfies

$$\sum_{n'''} H_{n\,n'''}^{0^{*}} \, \xi_{n'''\,n''}^{*} - W_{n''}^{*} \, \xi_{n\,n''}^{*} = 0 \; ,$$

which, upon using the relations  $H_{nn''}^{0*} = H_{n'''n}^{0}$  and  $W_{n''} = W_{n''}^{*}$ , becomes

$$\sum_{n'''} H^{0}_{n'''n} \, \xi^{\bullet}_{n'''n''} - W_{n''} \, \xi^{\bullet}_{nn''} = 0 \; .$$

Multiplying this by  $\xi_{nn'}$  and summing over *n*, we have

$$\sum_{n\,n'''} \xi_{n'''\,n''}^* H_{n'''\,n}^0 \xi_{n\,n'} - W_{n''} \sum_{n} \xi_{n\,n''}^* \xi_{n\,n'} = 0 . \qquad (35.30)$$

After interchanging the dummy suffixes n and  $n^{m}$  in the first summation of Eq. (35.30) we subtract it from Eq. (35.29) to obtain

$$(W_{n''} - W_{n'}) \sum_{n} \xi^{*}_{nn''} \xi_{nn'} = 0.$$

Since we have assumed at the beginning that  $W_{n'} \neq W_{n''}$  for  $n' \neq n''$ , this relationship immediately leads to the required result (25.28).

Combining the normalization condition (35.27) with the orthogonality relation (35.28), we can write

$$\sum_{n} \xi_{nn'}^* \xi_{nn'} = \delta_{n''n'} \,. \tag{35.31}$$

Next we must prove that for the determinant  $\Delta$  constructed from these  $\xi$ 's by

$$\Delta = \begin{vmatrix} \xi_{11} & \xi_{12} & \xi_{13} \dots \xi_{1N} \\ \xi_{21} & \xi_{22} & \xi_{23} \dots \xi_{2N} \\ & \ddots & & \\ \xi_{N1} & \xi_{N2} & \xi_{N3} & \xi_{NN} \end{vmatrix} = \begin{vmatrix} \xi_{11} & \xi_{21} & \xi_{31} \dots \xi_{N1} \\ \xi_{12} & \xi_{22} & \xi_{32} \dots \xi_{N2} \\ & \ddots & & \\ \xi_{1N} & \xi_{2N} & \xi_{3N} & \xi_{NN} \end{vmatrix}$$

we have

$$\Delta \neq 0, \qquad (35.32)$$

- or, more specifically,
- $\Delta \Delta^{\bullet} = 1 . \qquad (35.32')$

сн. 5, § 35]

The proof is quite easy. From the multiplication rules for determinants, we have

$$\Delta \Delta^{*} = \begin{vmatrix} \sum_{n} \xi_{n1} & \xi_{n1} & \sum_{n} \xi_{n1} & \xi_{n2} & \cdots & \sum_{n} \xi_{n1} & \xi_{nN} \\ \sum_{n} \xi_{n2} & \xi_{n1} & \sum_{n} \xi_{n2} & \xi_{n2}^{*} & \cdots & \sum_{n} \xi_{n2} & \xi_{nN}^{*} \\ \vdots & \vdots & \vdots & \vdots \\ \sum_{n} \xi_{nN} & \xi_{n1}^{*} & \sum_{n} \xi_{nN} & \xi_{n2}^{*} & \cdots & \sum_{n} \xi_{nN} & \xi_{nN}^{*} \end{vmatrix},$$

which, upon using relationship (35.31), reduces to

$$\Delta \Delta^* = \begin{vmatrix} 1 & 0 & 0 \dots & 0 \\ 0 & 1 & 0 \dots & 0 \\ 0 & 0 & 1 \dots & 0 \\ & & \ddots & \\ 0 & 0 & 0 \dots & 1 \end{vmatrix} = 1.$$

We have now derived all the necessary properties of the solutions of Eq. (35.24').

If the normalized solutions  $\xi_{1n'}, \xi_{2n'}, \ldots, \xi_{Nn'}$  of equation (35.24') are obtained for all the *n'*-values, the matrix we have been after is given by

$$U_{nn'} = \xi_{nn'} , \qquad (35.33)$$

and we thus have found all the  $U_{nn'}$ 's and  $W_{n'}$ 's which satisfy Eq. (35.23').

In order to show that the matrix  $U_{nn}$ , is in fact the desired solution, we have to prove that it satisfies both equation (35.23') and the unitarity condition (35.12). That  $U_{nn'}$  satisfies Eq. (35.23') is evident from the fact that the  $\xi$ 's satisfy Eq. (35.24). We can prove the unitarity of  $U_{nn'}$  in the following way.

First of all, the second condition of Eq. (35.12) follows directly from Eq. (35.31) and what we must show is that the first condition is also satisfied. Starting with the second condition

$$\sum_{\mathbf{n}''} U_{\mathbf{n}''\mathbf{n}}^* U_{\mathbf{n}''\mathbf{n}'} = \delta_{\mathbf{n}\mathbf{n}'},$$

we multiply it by  $U_{n''n}^*$ , and take the sum over n' to obtain

$$\sum_{n''} \sum_{n'} U_{n'''n'}^{*} U_{n''n}^{*} U_{n''n'} = U_{n'''n}^{*}$$

which can be rewritten to become

$$\sum_{n''} \left\{ \sum_{n'} U_{n'''n'}^* U_{n''n'} - \delta_{n'''n''} \right\} U_{n''n}^* = 0 .$$

Taking the complex conjugate of this equation and making use of Eq. (35.33), we obtain

$$\sum_{n''} \left\{ \sum_{n'} U_{n'''n'} \, U_{n'''n'}^* - \delta_{n'''n''} \right\} \xi_{n''n} = 0 \,. \tag{35.34}$$

Since the determinant  $|\xi_{n'n}|$  does not vanish (by Eq. (35.32')), linear homogeneous equations

$$\eta_{1} \xi_{11} + \eta_{2} \xi_{21} + \ldots + \eta_{N} \xi_{N1} = 0$$
  

$$\eta_{1} \xi_{12} + \eta_{2} \xi_{22} + \ldots + \eta_{N} \xi_{N2} = 0$$
  

$$\vdots$$
  

$$\eta_{1} \xi_{1N} + \eta_{2} \xi_{2N} + \ldots + \eta_{N} \xi_{NN} = 0$$

have no other solutions than  $\eta_1 = \eta_2 = \ldots = \eta_N = 0$ . From Eq. (35.34) therefore follows

$$\sum_{n'} U_{n'''n'} U_{n''n'}^* - \delta_{n'''n''} = 0 .$$

This is in fact the first condition of Eq. (35.12) and we have proved the unitarity of  $U_{nn'}$ .

We have thus obtained  $U_{nn'}$  from  $\xi_{nn'}$  by (35.33). However,  $\xi_{nn'}$  cannot be determined uniquely even with the normalization condition. For the multiplication of  $\xi_{nn'}$  by an arbitrary factor  $e^{i\varphi_{n'}}$ , with  $\varphi_{n'}$  a real number, does not violate the condition (35.27). The matrix  $U_{nn'}$  accordingly has the same kind of indeterminary. In our previous case of the oscillator, we saw that the coordinate x and the momentum p involved certain undetermined phase angles. This really was due to the above indeterminacy of U. We shall, however, see later on that this type of indeterminacy does not affect the physical conclusions of the theory.

We have shown in the above that the problems of matrix mechanics can be reduced to the solving of linear homogeneous equations of the type (35.24'). For the sake of simplicity, we assumed in our considerations that the number of possible *n*-values was finite, but this

assumption does not hold for general dynamical systems of quantum mechanics. Therefore, the problem is actually the solution of Eq. (35.24') as an infinite number of simultaneous equations for an infinite number of unknown variables  $\xi_1, \xi_2, \xi_3, \ldots$ . The ordinary theory of linear algebra is not sufficient for this problem and we need more advanced mathematics. However many of the conclusions arrived at in our simple case remain valid also in the case of an infinite number of variables. For instance, in order for Eq. (35.24') to have non-trivial solutions, W in the equation can not take arbitrary values but instead only a certain, generally infinite, number of allowed values. These values of W allowing the existence of solutions of Eq. (35.24') are characteristic of the matrix  $H^0$  and are called the eigenvalues of the matrix H<sup>0</sup>. The solution of Eq. (35.24'),  $\xi_{1n'}, \xi_{2n'}, \xi_{3n'}, \ldots$ , corresponding to one of these eigenvalues,  $W_{n'}$ , is called the eigensolution corresponding to this eigenvalue. Furthermore, it is shown that the eigenvalues of  $H^0$  are determined by the functional form itself of the function H(x, p) and are not affected by the choice of  $X^0$  and  $P^0$ .

In conclusion, the problem is now reduced to finding the eigenvalues and the eigensolutions of an infinite number of linear homogeneous equations for an infinite number of variables. This is the so-called eigenvalue problem. Though problems of this type in general are still not very easy to solve, it is a substantial step forward that we can now always deal, irrespective of the properties of the forces acting in the dynamical system under consideration, with linear homogeneous equations rather than the original higher order equations such as Eqs. (33.17) and (33.18).

• Nevertheless, it is very cumbersome actually to solve problems by matrix mechanics. Pauli and Dirac could each independently solve the problem of the hydrogen atom by matrix mechanics and by similar methods, and showed that the energy levels of Eq. (18.10) were actually derived, but their calculations were extremely complicated. If it were not for the help of wave mechanics, quantum mechanics would have been too cumbersome to be actually applied. Thanks to the discovery of wave mechanics, however, it is now possible to solve problems by means of more familiar mathematics.



# APPENDIX and EXERCISES

### APPENDIX I

### **Boltzmann's Principle**

We shall derive Boltzmann's principle in this appendix. There are various methods of deriving this principle, but we shall give the one which seems, though not quite free from criticism, to have the clearest physical meaning.

### i) THE TRAJECTORY OF A DYNAMICAL SYSTEM

Let us consider an otherwise arbitrary dynamical system governed only by conservative forces. Let the canonical coordinates describing the system and their conjugate momenta be denoted by  $q_1, q_2, \ldots q_f$ , and  $p_1, p_2, \ldots p_f$ , respectively, where f is the number of degrees of freedom of the system.

Furthermore, let the energy of the system under consideration as a function of q's and p's be denoted by

$$H(q_1, q_2, \ldots, q_f, p_1, p_2, \ldots, p_f).$$
 (I.1)

This function, H(q, p) for short, is the so-called Hamiltonian function of the system and, according to Hamilton's theory, the motion of the system can be determined by solving the equations of motion

$$\frac{\mathrm{d}q_1}{\mathrm{d}t} = \frac{\partial H}{\partial p_1}, \quad \frac{\mathrm{d}q_2}{\mathrm{d}t} = \frac{\partial H}{\partial p_2}, \dots, \quad \frac{\mathrm{d}q_f}{\mathrm{d}t} = \frac{\partial H}{\partial p_f},$$

$$\frac{\mathrm{d}p_1}{\mathrm{d}t} = -\frac{\partial H}{\partial q_1}, \quad \frac{\mathrm{d}p_2}{\mathrm{d}t} = -\frac{\partial H}{\partial q_2}, \dots, \quad \frac{\mathrm{d}p_f}{\mathrm{d}t} = -\frac{\partial H}{\partial q_f},$$
(I.2)

to obtain the q's and p's as functions of the time t.

The positions and velocities of all the constituents of the dynamical system are determined by specifying the values of these q's and p's at each instant. We express this fact by saying that the phase of the system is determined by giving the values of these 2f variables, or, more simply, we sometime call the set of values of these 2f variables the phase of the system.

In describing the time variation of the phase of the system, it is often convenient to introduce the following geometrical picture. Namely, we consider a 2f-dimensional space spanned by the orthogonal axes  $q_1, q_2, \ldots, q_f, p_1, p_2, \ldots, p_f$ . Each phase of the system is, then, represented by a point in this space. We call this space and the point representing the phase the phase space and the representative point, respectively. As the system changes in the course of time, the representative point moves in phase space and the path described is called the trajectory of the system.

Concerning the trajectory, the following properties result directly from the equations of motion (I.2).

(I) The trajectories never cross each other; that is, more than one trajectory never pass through the same point. This is due to the fact that, if the values of  $q_1, q_2, \ldots, q_f$  and  $p_1, p_2, \ldots, p_f$  are given at a certain instant, their values at an adjacent instant are uniquely determined by Eq. (I.2).

(II) Two representative points infinitely close to each other at a certain instant will always be so at any other instant in the course of their motion. Or, more generally, if one of the representative points starts to move at time  $t_1$  from a certain point, while the other starts to move at time  $t_2$  from another point infinitely close to the former point, then the location of the first representative point at time  $t_1 + t$  is infinitely close to the location of the second at time  $t_2 + t$ .

(III) Let us consider a (2f-1)-dimensional surface, defined by the equation  $H(q_1q_2 \dots q_f, p_1p_2 \dots p_f) = \text{const.} = E. \quad (I.3)$ 

The representative point starting from a certain point on this surface remains for all time on the same surface during the course of its motion in the phase space. In other words, the trajectory is contained in this surface. This is the law of energy conservation expressed in geometrical language.

A surface, like the one defined by (I.3) on which the energy has a certain constant value, is called an equi-energy surface.

These properties, (I), (II) and (III), hold quite generally. In cases where the dynamical system under consideration has a certain type of symmetry, conservation laws other than that of energy, e.g., conservation of momentum or of angular momentum, prevail. Denoting such a constant of motion, as a function of q and p, by  $F(q_1, q_2 \ldots q_f,$  $p_1, p_2, \ldots p_f)$ , the trajectory of the system remains not only on an equi-energy surface but also on the surface defined by

$$F(q_1q_2\ldots q_fp_1p_2\ldots p_f) = \text{const.} = K.$$

APPENDIX I

Thus, the trajectory is restricted to lie on the intersection of this surface with the equi-energy surface. Such a constant of motion, however, exists only for dynamical systems of very simple structure. For instance, the conservation of angular momentum is realized only when the dynamical system has perfect spherical symmetry, a condition which can hardly by expected to hold for real systems. The dynamical systems which are actually handled in statistical mechanics are of such a complex structure that they have no constant of motion other than the energy.

What shape of the trajectory then does the representative point describe on the equi-energy surface? The answer to this question can only be obtained when one solves Eq. (I.2) for each specific case. However, this is utterly impossible for the complex dynamical systems which are usually treated in statistical mechanics. We may expect that the more complex the dynamical system becomes, the more so would be its trajectory, so that it would cover almost every part of the equi-energy surface.

With this in mind we assume the following, namely that the trajectory will cover the entire surface of the equi-energy surface. We call this the ergodic hypothesis. From the mathematical point of view, there are a number of arguments to be made concerning this hypothesis but we shall not go into any details here. We now make a further hypothesis about our dynamical system. That is, we assume that the equi-energy surface does not extend to infinity, but forms a closed surface in the phase space. This hypothesis is fulfilled for the dynamical systems one usually deals with.

Under these hypotheses the dynamical system assumes, for a given energy, all the phases on the equi-energy surface as time goes on. In general, the aggregate of all the phases which the dynamical system assumes during its temporal development is called the temporal ensemble. According to our ergodic hypothesis, this temporal ensemble of the system is made up of all the phases on the equi-energy surface.

## ii) THE PROBABILITY DENSITY

Let us take a small domain O on the equi-energy surface on which the representative phase point under consideration describes the trajectory.

Observing the representative point for a very long time T, one will

find that it goes through this domain again and again during its journey on the equi-energy surface. Let  $\tau$  be the total time during which the point is within the domain O. Then the ratio

$$W_T = \frac{\tau}{T}$$

indicates what fraction of the total time T the representative point stays within the domain O. As T goes to infinity, this ratio  $W_T$  will approach a certain finite limit. Or rather we assume here that it does, and put

$$W = \lim_{T \to \infty} \frac{\tau}{T} ; \qquad (I.5)$$

then W represents the probability for the phase point to be found within the domain O. Furthermore, when the domain O becomes infinitesimally small, the ratio W/O will approach a certain limit, which will hereafter be designated by w,

$$w = \lim_{o \to 0} \frac{W}{O} \,. \tag{I.6}$$

This quantity w we shall call the probability density for the representative point to be in the infinitesimal domain O.

Under the hypotheses introduced in the preceding subsection, one can calculate this probability density without actually solving the equations of motion. We proceed as follows. Consider an infinitesimal domain  $dO_1$  on the equi-energy surface. To each phase point in  $dO_1$ there is a corresponding phase point in which the system will be found after a certain time t. We consider the set of all such phase points.

This set will, according to the property (II) in the last subsection, be contained in another infinitesimal domain  $dO_2$ . Conversely, to each phase point in  $dO_2$ , there is in  $dO_1$  a corresponding phase point in which the system was found a time t before.

Therefore, there is a one to one correspondence between the two sets of phase points, those in  $dO_1$  and those in  $dO_2$ . Moreover, since they are connected by a definite time interval t, the time which the representative point spends in  $dO_1$  is equal to the time it does in  $dO_2$ . Accordingly, we have

$$w_1 \, \mathrm{d}O_1 = w_2 \, \mathrm{d}O_2$$
, (1.7)

APPENDIX I

where  $w_1$  and  $w_2$  denote the probability densities at  $dO_1$  and  $dO_2$ , respectively. This implies that we can calculate the ratio of the two probability densities,  $w_1$  and  $w_2$ , when we have a method of calculating the ratio of the two areas,  $dO_1$  and  $dO_2$ . The latter ratio can in fact be calculated by making use of the following important theorem.

## iii) LIOUVILLE'S THEOREM

Let us consider a 2f-dimensional domain  $G_1$  in the phase space which is also of dimension 2f. The phase points in this domain will move in phase space as the time goes on. The set of phase points contained in this domain at one time will constitute another domain  $G_2$  at a later time. Then we have the theorem: The volume of  $G_2$ is equal to that of  $G_1$  independently of the time. In other words, we can simply say that the volume of the phase space is conserved.

This theorem is called Liouville's theorem and the proof goes as follows.

**Proof**: We must find what domain will be formed after a certain time by the set of phase points which once constituted the domain  $G_1$ . For this purpose, it is enough to observe the temporal variation of the phase points constituting the boundary surface of the domain. The set of these phase points will move, starting from this surface and forming all the time a certain surface. If the original surface is continuous, it will always remain so because of the property (II) of the last section. The volume contained in this moving surface is our domain  $G_2$ since the phase points originally inside the surface will never leak out of the surface due to the properties (I) and (II).

We set the starting time at t = 0 and investigate the surface after the infinitesimal time dt.

Let the equation determining the surface be

$$F(q_1q_2...q_f, p_1p_2...p_f) = 0$$
 (I.8)

and let the interior and the exterior of the surface be given by F < 0, and F > 0, respectively. The volume of the domain  $G_1$  is then given by

$$V_1 = \iiint \mathrm{d}q_1 \, \mathrm{d}q_2 \dots \mathrm{d}p_f \,, \tag{I.9}$$

where the integration is to be carried out in the domain

$$F(q_1q_2\dots p_f) \leq 0. \tag{I.10}$$

BOLTZMANN'S PRINCIPLE

Here we have used for simplicity the triple integral sign for the 2f-fold integration. Further, we will use the double integral sign for a (2f-1)-fold integration on an equi-energy surface and call the domain of integration the area.

We now look for the equation determining the surface formed at t = dt by the phase points which were on the surface (I.8) at t = 0. In other words, we are to find the equation to be satisfied at t = dt by the representative points. This can easily be obtained in the following way.

We look at one of the representative points at t = dt and designate its coordinates by  $q'_1, q'_2, \ldots, q'_f, p'_1, p'_2, \ldots, p'_f$ . According to the equation of motion (I.2), the coordinates of this representative point at t = 0 are given by

$$q_{1} = q'_{1} - \left(\frac{\partial H}{\partial p'_{1}}\right) dt ,$$

$$q_{2} = q'_{2} - \left(\frac{\partial H}{\partial p'_{2}}\right) dt$$

$$\dots \qquad (I.11)$$

$$p_{1} = p'_{1} + \left(\frac{\partial H}{\partial q'_{1}}\right) dt ,$$

$$p_{2} = p'_{2} + \left(\frac{\partial H}{\partial q'_{2}}\right) dt ,$$

$$\dots \qquad \dots$$

where  $(\partial H/\partial p')$  and  $(\partial H/\partial q')$  stand for  $\partial H(p', q')/\partial p'$  and  $\partial H(p', q')/\partial q'$ , respectively.

The coordinates of the representative point thus found must, by assumption, satisfy Eq. (I.8). Thus

$$F\left(q'_{1}-\left(\frac{\partial H}{\partial p'_{1}}\right) \mathrm{d}t, q'_{2}-\left(\frac{\partial H}{\partial p'_{2}}\right) \mathrm{d}t, \dots, \\p'_{1}+\left(\frac{\partial H}{\partial q'_{1}}\right) \mathrm{d}t, p'_{2}+\left(\frac{\partial H}{\partial q'_{2}}\right) \mathrm{d}t, \dots\right)=0.$$

This is the required equation to be satisfied at t = dt by the coordinates of the representative point which originally was on the surface (I.8)

267

APPENDIX I

at t = 0. The volume contained in this surface, is then given by

$$V_2 = \iiint \mathrm{d}q'_1 \dots \mathrm{d}q'_f \mathrm{d}p'_1 \dots \mathrm{d}p'_f$$
(I.12)

where the integration is to be carried out over the domain

$$F\left(q'_{1}-\left(\frac{\partial H}{\partial p'_{1}}\right) \mathrm{d}t, q'_{2}-\left(\frac{\partial H}{\partial p'_{2}}\right) \mathrm{d}t, \dots,$$

$$p'_{1}+\left(\frac{\partial H}{\partial q'_{1}}\right) \mathrm{d}t, p'_{2}+\left(\frac{\partial H}{\partial q'_{2}}\right) \mathrm{d}t, \dots\right) \leq 0. \qquad (I.13)$$

The integration can easily be done by changing the variables from  $q'_1, \ldots, q'_f, p'_1, \ldots, p'_f$  to  $q_1, \ldots, q_f, p_1, \ldots, p_f$  defined by (I.11).

The integral then becomes

$$V_2 = \int \int \int \frac{\partial (q'_1 \dots q'_f p'_1 \dots p'_f)}{\partial (q_1 \dots q_f p_1 \dots p_f)} \, \mathrm{d}q_1 \dots \, \mathrm{d}q_f \, \mathrm{d}p_1 \dots \, \mathrm{d}p_f \tag{I.14}$$

where the domain of integration is again defined by (I.10). We now calculate the Jacobian

$$\frac{\partial(q'_1\ldots q'_f p'_1\ldots p'_f)}{\partial(q_1\ldots q_f p_1\ldots p_f)}.$$

This can easily be done by using (I.11), since

$$\left\{ \frac{\partial (q'_1 \dots q'_f p'_1 \dots p'_f)}{\partial (q_1 \dots q_f p_1 \dots p_f)} \right\}^{-1}$$

$$= \begin{vmatrix} 1 - \frac{\partial}{\partial q'_1} \left( \frac{\partial H}{\partial p'_1} \right) dt, & -\frac{\partial}{\partial q'_1} \left( \frac{\partial H}{\partial p'_2} \right) dt, \dots \end{vmatrix}$$

$$= \begin{vmatrix} -\frac{\partial}{\partial q'_2} \left( \frac{\partial H}{\partial p'_1} \right) dt, & 1 - \frac{\partial}{\partial q'_2} \left( \frac{\partial H}{\partial p'_2} \right) dt, \dots \end{vmatrix}$$

which, retaining only first order terms in dt, becomes

$$= 1 - \left\{ \sum_{s} \frac{\partial}{\partial q'_{s}} \left( \frac{\partial H}{\partial p'_{s}} \right) - \sum_{s} \frac{\partial}{\partial p'_{s}} \left( \frac{\partial H}{\partial q'_{s}} \right) \right\} \mathrm{d}t \,.$$

A. I]

Making use of the identity

$$\frac{\partial}{\partial q'_s} \left( \frac{\partial H}{\partial p'_s} \right) = \frac{\partial}{\partial p'_s} \left( \frac{\partial H}{\partial q'_s} \right),$$

we arrive at the result

$$\frac{\partial (q'_1 \dots q'_f p'_1 \dots p'_f)}{\partial (q_1 \dots q_f p_f \dots p_f)} = 1$$

Thus (I.14) is reduced to

$$V_2 = \iiint \mathrm{d}q_1 \dots \mathrm{d}q_f \,\mathrm{d}p_1 \dots \mathrm{d}p_f \,, \tag{1.15}$$

the domain of integration now being defined by (I.10). We therefore have the result

$$V_2 = V_1 \tag{I.16}$$

up to first order with respect to dt, which means that the difference  $V_2 - V_1$  is of the order of  $(dt)^2$ , so that

$$\frac{\mathrm{d}V}{\mathrm{d}t} = 0 \;. \tag{I.16'}$$

We have thus proved that the volume of the phase space is conserved.

# iv) CALCULATION OF THE PROBABILITY DENSITY

The relation between  $dO_1$  and  $dO_2$  of Eq. (I.7) can readily be obtained by the use of Liouville's theorem. In addition to the equienergy surface H = E on which  $dO_1$  and  $dO_2$  lie, we consider an auxiliary surface  $H = E + \delta E$ , with an infinitesimally small energy difference  $\delta E$ . Let the distance between the two surfaces be  $\delta n$ . This distance  $\delta n$  varies from place to place on the surface and its values at  $dO_1$  and  $dO_2$  on the surface H = E are designated by  $dn_1$  and  $dn_2$ , respectively. Very close to the surface element  $dO_1$  we consider a surface element  $dO'_1$  on the auxiliary surface  $H = E + \delta E$  in such a way that its area is equal to that of  $dO_2$ . The domain  $dO'_1$  moves in the course of time on the auxiliary surface and arrives at  $dO'_2$  after the time t. This  $dO'_2$  will then be very close to  $dO_2$  since in the limit of  $\delta E \to 0$   $dO'_2$  tends to  $dO_2$ . The situation is illustrated schematically in Fig. (A.1).



Fig. A.1. The relationship between the separation of equi-energy surfaces and the surface element.

Now let us consider two cylindrical domains between the equi-energy surfaces H = E and  $H = E + \delta E$ , one with  $dO_1$  and  $dO'_1$  as the bases and the other with  $dO_2$  and  $dO'_2$ . The representative points starting at t = 0 from the former domain will all be found at t = tin the latter domain. Therefore, Liou-

ville's theorem requires that the two volumes be equal; i.e.,

$$\mathrm{d}O_1 \delta n_1 = \mathrm{d}O_2 \delta n_2 \;. \tag{I.17}$$

This relationship (I.17) reduces the calculation of the ratio of  $dO_1$  to  $dO_2$  to that of the ratio of  $\delta n_1$  to  $\delta n_2$ . The distance between the two equi-energy surfaces can, on the other hand, be obtained from the equation

$$\delta E = \sqrt{\left(\frac{\partial H}{\partial q_1}\right)^2 + \left(\frac{\partial H}{\partial q_2}\right)^2 + \ldots + \left(\frac{\partial H}{\partial q_f}\right)^2} \,\delta n \,. \tag{I.18}$$

If, for simplicity, we employ the symbols of ordinary 3-dimensional vector analysis for our 2*f*-dimensional phase space, Eq. (I.18) can be written as  $\delta E = 4 \operatorname{grad} H + \delta r$ 

$$\delta E = |\operatorname{grad} H| \,\delta n \,. \tag{I.18'}$$

Using this (I,17) becomes

$$\left(\frac{\mathrm{d}O}{|\operatorname{grad} H|}\right)_{1} = \left(\frac{\mathrm{d}O}{|\operatorname{grad} H|}\right)_{2}.$$
 (I.17')

Introducing this result into Eq. (I.7), we obtain

 $(w | \text{grad } H |)_1 = (w | \text{grad } H |)_2.$  (I.19)

Since Eq. (I.19) holds for any choice of  $dO_1$  and  $dO_2$ , on the equienergy surface H = E, it can be expressed more generally as

$$w = \frac{c}{|\operatorname{grad} H|}, \qquad (I.19')$$

or by using Eq. (I.18') it can also be written in the alternative from

$$w = c \frac{\delta n}{\delta E} \,. \tag{I.19"}$$

The quantity c in Eq. (I.19') has a constant value on an equi-energy surface but assumes different values on different equi-energy surfaces. It is then in general a function of the energy E. As will be shown later, the functional form of this quantity, c(E), can also be calculated without much difficulty.

Once the Hamiltonian function  $H(q_1, q_2, \ldots, p_f)$  is given, we are now able to calculate the probability density w from Eq. (I.19') together with the formula for c(E).

The meaning of Eqs. (I.19') and (I.19") can be visualized in the following manner. Let us imagine a set of equi-energy surfaces with the same energy difference  $\delta E$  from one to the other in phase space. The probability density on one of these surface will then be smaller at places where these surfaces lie close to each other, i.e.,  $\delta n/\delta E$  is smaller, while it will be larger at places where the surfaces are more widely separated, i.e.,  $\delta n/\delta E$  is larger.

The situation can further be understood in terms of the velocity of representative point in the phase space, which is given by

$$v = \sqrt{\left(\frac{\mathrm{d}q_1}{\mathrm{d}t}\right)^2 + \left(\frac{\mathrm{d}q_2}{\mathrm{d}t}\right)^2 + \ldots + \left(\frac{\mathrm{d}q_f}{\mathrm{d}t}\right)^2} = \sqrt{\left(\frac{\partial H}{\partial p_1}\right)^2 + \left(\frac{\partial H}{\partial p_2}\right)^2 + \ldots + \left(\frac{\partial H}{\partial p_f}\right)^2} = |\operatorname{grad} H|.$$
(I.20)

Thus, the representative point moves faster in places where the equienergy surfaces are closer to each other and slower in places where these surfaces are more widely separated. Then we can infer that by definition, the probability will be large when the velocity is small and small when the velocity is large.

Having found the probability density, we can now calculate the probability for any finite domain O on an equi-energy surface by

$$W_o = c \iint_o \frac{\mathrm{d}O}{|\operatorname{grad} H|}, \qquad (I.21)$$

where the double integral sign is for the (2f-1)-fold integral over the domain O on the equi-energy surface.

If the domain O is the entire surface of the equi-energy surface, the probability is by definition unity, i.e.

271

A. I]

APPENDIX I

$$1 = c \int \int \frac{\mathrm{d}O}{|\operatorname{grad} H|}, \qquad (I.22)$$

where the integration is to be carried out over the entire surface H = E. The quantity c is thus determined.

To perform the integration in Eq. (I.22) it is convenient to use a geometrical consideration along with Eq. (I.19''). Then the integral is

$$\int \int \frac{\mathrm{d}O}{|\operatorname{grad} H|} = \frac{1}{\delta E} \int \int \delta n \cdot \mathrm{d}O \; .$$

The integral on the right had side has the simple geometrical meaning that it gives the volume of the shell shaped region between the two equi-energy surfaces H = E and  $H = E + \delta E$ . Denoting the volume of the phase space enclosed by the equi-energy surface H = E which by assumption has a finite extension, by V(E), the shell shaped volume between H = E and  $H = E + \delta E$  is

$$\frac{\mathrm{d}V}{\mathrm{d}E}\delta E$$

and hence,

$$\int \int \delta n \, \mathrm{d}O = \frac{\mathrm{d}V}{\mathrm{d}E} \, \delta E \; .$$

V(E), is, on the other hand, defined by

$$V(E) = \iiint \mathrm{d}q_1 \dots \mathrm{d}q_f \,\mathrm{d}p_1 \dots \mathrm{d}p_f, \qquad (I.23)$$

4 1 / E

where the integral is carried out over the entire domain defined by

II/a a

$$r(q_1q_2\dots p_f) \ge E ,$$

$$c = \frac{1}{\frac{\mathrm{d}V}{\mathrm{d}E}}, \qquad (I.24)$$

We then have for c

and, accordingly, for the probability density

$$w = \frac{1}{\frac{\mathrm{d}V}{\mathrm{d}E}} \frac{1}{|\operatorname{grad}H|}$$
(I.19"')

and for the probability for a finite domain O

$$W_o = \frac{1}{\frac{\mathrm{d}V}{\mathrm{d}E}} \iint_o \frac{\mathrm{d}O}{|\operatorname{grad} H|} \,. \tag{I.21'}$$

U

272
#### BOLTZMANN'S PRINCIPLE

The integration over the domain O on the equi-energy surface can be most easily performed as follows. We specify a point on the surface by the 2f-1 variables  $q_1, q_2, \ldots, q_f, p_1, p_2, \ldots, p_{f-1}$ , with the last coordinate  $p_f$  given by solving  $H(q_1, q_2, \ldots, q_f, p_1, p_2, \ldots, p_f) = E$ . Then the surface element dO is expressed as

$$\mathrm{d}O = \frac{|\operatorname{grad} H|}{\frac{\partial H}{\partial p_f}} \,\mathrm{d}q_1 \dots \,\mathrm{d}q_f \,\mathrm{d}p_1 \dots \,\mathrm{d}p_{f-1} \,. \tag{I.25}$$

where  $p_f$  in H and  $\partial H/\partial p_f$  is understood to belong to the solution of H = E and hence to be a function of the 2f-1 variables  $q_1q_2...q_fp_1p_2...p_{f-1}$ . The relationship (I.25) can be inferred in the following way. On the surfaces H = E and the adjacent surface H = E + dE, let us consider the surface element dO and dO' both specified by  $dq_1, \ldots, dq_f, \ldots, dp_{f-1}$ ; then the volume of the infinitesimal cylinder between two surfaces having dO and dO' as bases is given by

$$\mathrm{d}v = \mathrm{d}q_1 \,\mathrm{d}q_2 \ldots \mathrm{d}q_f \,\mathrm{d}p_1 \ldots \mathrm{d}p_f$$

with

$$\mathrm{d}p_f = \frac{1}{\frac{\partial H}{\partial p_f}} \,\mathrm{d}E \;.$$

On the other hand, the volume of the same cylinder is given by

$$\mathrm{d}v = \mathrm{d}O\,\mathrm{d}n$$
,

dn being the hight of the cylinder, i.e., the distance between dO and dO', which is given by

$$\mathrm{d}n = \frac{1}{|\operatorname{grad} H|} \,\mathrm{d}E \,.$$

These four equations immediately give rise to Eq. (I.25). Making use of Eq. (I.25) in Eq. (I.21') we obtain

$$W_o = \frac{1}{\frac{\mathrm{d}V}{\mathrm{d}E}} \iint \frac{1}{\frac{\partial E}{\partial p_f}} \mathrm{d}q_1 \dots \mathrm{d}q_f \,\mathrm{d}p_1 \dots \mathrm{d}p_{f-1} \,, \qquad (\mathrm{I.21''})$$

where the integration is to be carried out over the projection of the domain O on the coordinate plane  $p_f = 0$ .

A. I]

APPENDIX I

Furthermore, the integration over the entire projection should give  $W_o = 1$  from which we obtain

$$\frac{\mathrm{d}V}{\mathrm{d}E} = \iint_{\substack{\text{entire}\\\text{projection}}} \frac{\frac{1}{\partial H}}{\frac{\partial p_f}{\partial p_f}} \mathrm{d}q_1 \dots \mathrm{d}q_f \,\mathrm{d}p_1 \dots \mathrm{d}p_{f-1} \,. \tag{I.25'}$$

#### v) A SYSTEM IMMERSED IN A THERMAL RESERVOIR

The preceding subsections have begun the consideration of the statistical behavior of a complicated dynamical system. We have seen that when the dynamical system under consideration satisfies the assumptions of section (i), we can calculate the probability for the system to be in a certain particular phase without actually solving the equations of motion for the system. The dynamical systems so far considered were isolated ones, i.e., those having no interactions with the surroundings and hence having a given constant energy all the time. The dynamical systems of practical interest in statistical mechanics, however, are the ones which are immersed in certain types of thermal reservoirs and, accordingly, which constantly exchange energy with the reservoir. We must then proceed to this kind of problem starting with the arguments given in the last sections for isolated systems.

Since a thermal reservoir is by itself a dynamical system, we can combine it with our dynamical system to make a composite dynamical system, which then can be considered as an isolated system. For this isolated system we can directly apply the arguments of the last sections.

Let the coordinates and the momenta of the dynamical system be denoted by  $q_1, q_2, \ldots, q_f, p_1, \ldots, p_f$ , and those of the thermal reservoir by  $q_1^0, q_2^0, \ldots, q_{f'}^0, p_f^0, \ldots, p_{f'}^0$ , respectively. Since the thermal reservoir is in general much larger than the dynamical system under consideration, the number  $f^0$  of its degrees of freedom is usually much larger than that of the dynamical system f. Furthermore, the Hamiltonian function of the dynamical system will be denoted by  $H(q_1, q_2, \ldots, p_f)$  and that of the thermal reservoir by  $H^0(q_1^0, \ldots, p_{f'}^0)$ . The Hamiltonian of the composite system is then given by

$$H^{\bullet}(q_{1} \dots p_{f} q_{1}^{0} \dots p_{f}^{0}) = H(q_{1} \dots p_{f}) + H^{0}(q_{1}^{0} \dots p_{f}^{0}).$$
(I.26)

BOLTZMANN'S PRINCIPLE

Strictly speaking, there should be an additional Hamiltonian besides those of the two systems, one which describes the mutual interaction between them. We simply neglect it as negligibly small, since only the existence of such an interaction is necessary for statistical considerations and its actual functional form is of little importance.

Hereafter all the quantities concerning the composite system will be denoted by the index<sup>\*</sup>, as in  $H^*$ , and those concerning the reservoir by the index 0, as in  $H^0$ .

When the composite system has an energy  $E^*$ , its representative point will move on the equi-energy surface  $H^* = E^*$  in  $2f^* = 2(f + f^0)$ -dimensional phase space. The probability for finding the phase point in the domain  $dO^*$  on this surface is, in accordance with Eq. (I.21"), then given by

$$W_{o*} = \frac{1}{\frac{\mathrm{d}V^*}{\mathrm{d}E^*}} \iint \frac{1}{\frac{\partial H^*}{\partial p_f^0}} \mathrm{d}q_1 \dots \mathrm{d}p_f \, \mathrm{d}q_1^0 \dots \mathrm{d}p_{f^{0-1}}^0, \qquad (I.27)$$

where the last variable  $p_{f^0}^0$  is understood to be the solution of  $H^*(q_1, \ldots, p_f, q_1^0, \ldots, p_{f^0}^0) = E^*$  and, hence, a function of the other  $\{2(f+f^0)-1\}$  variables

$$q_1, \ldots, p_f, q_1^0, \ldots, p_{f^0-1}^0$$

Since Eq. (I.26) holds for  $H^*$ , the last variable  $p_{f^0}^0$  in Eq. (I.27) is the solution of

$$H^{0}(q_{1}^{0}\dots p_{f^{0}}^{0}) = E^{\bullet} - H(q_{1}\dots p_{f}),$$
 (I.28)

and we note that

$$\frac{\partial H^{\bullet}}{\partial p_{f^{\circ}}^{0}} = \frac{\partial H^{\circ}}{\partial p_{f^{\circ}}^{0}}.$$
 (I.29)

We are interested in Boltzmann's principle and therefore in the probability for the phase of the dynamical system to be found in the domain  $dq_1 dq_2 \ldots dp_f$  around the point specified by  $q_1, q_2, \ldots, p_f$  with no restriction on the phase of the thermal reservoir. Accordingly, the domain of concern  $O^*$  is the one with

 $q_1$  in the specified range  $(q_1, q_1 + dq_1)$  $q_2$  in the specified range  $(q_2, q_2 + dq_2)$ 

A. I]

#### APPENDIX I

 $p_f$  in the specified range  $(p_f, p_f + dp_f)$   $q_1^0$  any value ,  $q_2^0$  any value ,  $p_{f^{0-1}}^0$  any value .

The probability we are looking for is then given by

$$W_{o*} = \frac{1}{\frac{\mathrm{d}V^{*}}{\mathrm{d}E^{*}}} \mathrm{d}q_{1} \,\mathrm{d}q_{2} \dots \mathrm{d}p_{f} \iint \frac{1}{\frac{\partial H^{0}}{\partial p_{f^{0}}^{0}}} \mathrm{d}q_{1}^{0} \,\mathrm{d}q_{2}^{0} \dots \mathrm{d}p_{f^{0-1}}^{0}, \qquad (I.30)$$

where Eq. (I.29) has been used.

In this expression,  $q_1, q_2, \ldots, p_f$  have specified values and the integration is to be carried out over the entire range of  $q_1^0, q_2^0, \ldots, p_{f^{0-1}}^0$ . In Eq. (I.30) the last variable  $p_{f^0}^0$  is a function of other variables obtained by solving Eq. (I.28).

The variables,  $q_1, q_2, \ldots p_f$ , of the dynamical system having fixed values,  $H(q_1, q_2, \ldots p_f)$  in Eq. (I.28) also has a definite value, which can be interpreted as the energy of this dynamical system. We denote this by *E*. Making use of Eq. (I.25') for the integral in Eq. (I.30), we obtain

$$\iint \frac{1}{\frac{\partial H^0}{\partial p_{f^0}^0}} \mathrm{d}q_1^0 \, \mathrm{d}q_2^0 \dots \mathrm{d}p_{f^{0-1}}^0 = \left(\frac{\mathrm{d}V^0}{\mathrm{d}E^0}\right)_{E^0 = E^* - E},\qquad(\mathrm{I.31})$$

where  $V^{0}(E^{0})$  is the V-function for the reservoir only, defined by

$$V^{0}(E^{0}) = \iiint \mathrm{d}q_{1}^{0} \,\mathrm{d}q_{2}^{0} \dots \mathrm{d}p_{f^{0}}^{0} \tag{I.32}$$

with the domain of integration

$$H^{0}(q_{1}^{0}, q_{2}^{0}, \ldots, p_{f^{0}}^{0}) \leq E^{0}$$
 (I.33)

We thus have derived the desired probability  $W_0^*$ , which is designated in the text, § 2, by  $P(q_1, q_2, \ldots, q_f, p_1, p_2, \ldots, p_f) dq_1 dq_2 \ldots dq_f dp_1 dp_2 \ldots dp_f$ ,

$$W_{o*} = P(q_1q_2 \dots q_f p_1 p_2 \dots p_f) dq_1 dq_2 \dots dq_f dp_1 dp_2 \dots dp_f$$
$$= \frac{1}{\frac{\mathrm{d}V^*}{\mathrm{d}E^*}} \left(\frac{\mathrm{d}V^0}{\mathrm{d}E^0}\right)_{E^0 = E^* - E} dq_1 dq_2 \dots dq_f dp_1 dp_2 \dots dp_f. \quad (I.34)$$

# vi) INTRODUCTION OF THE TEMPERATURE

The probability which is used in Boltzmann's principle is given by Eq. (I.34). However, it still contains the quantity,  $dV^0/dE^0$ , which depends on the specific structure of the thermal reservoir. In this section we explain how we can approximate this quantity by a certain function which does not depend on the structure. This approximation will be valid under the circumstance that the number of degrees of freedom of the reservoir is extremely large. A quantity called "temperature" will be introduced in this process.

We shall take an ideal gas in a large box as the reservoir. The volume function  $V^0(E^0)$  of this reservoir can be calculated as follows. Let the number of gas molecules be  $N^0$  and the coordinates of the representative point be  $q_1^0, q_2^0 \ldots q_{3N^0}^0, p_1^0, p_2^0, \ldots, p_{3N^0}^0$ . The Hamiltonian function is then given by

$$H^{0} = \frac{1}{2m} \left\{ p_{1}^{0^{2}} + p_{2}^{0^{2}} + \ldots + p_{3N^{0}}^{0^{2}} \right\}.$$
 (1.35)

Since this Hamiltonian does not contain the  $q^{0}$ 's, the integrations with respect to the  $q^{0}$ 's in Eq. (I.32) simply give the  $N^{0}$ -th power of the volume of the box. Denoting this volume by  $\Omega$ , we obtain

$$V^{0}(E^{0}) = \Omega^{N^{0}} \iiint \mathrm{d}p_{1}^{0} \mathrm{d}p_{2}^{0} \dots \mathrm{d}p_{3N^{0}}^{0}$$
,

where the integration with respect to the  $p^{0}$ 's is to be carried out over the domain:

$$\frac{1}{2m}(p_1^{0^2}+p_2^{0^2}+\ldots+p_{3N^0}^{0^2})^2 \leq E^0$$

in the  $3N^{0}$ -dimensional  $p^{0}$ -space. In other words, the integral is the volume of a sphere of radius  $\sqrt{2mE^{0}}$  in this  $3N^{0}$ -dimensional  $p^{0}$ -space. Although the integration can be performed in a straightforword way, it suffices for our purpose to make use of the obvious fact that the volume of this sphere is proportional to the  $3N^{0}$ -th power of the radius. Therefore, the dependence of  $V^{0}(E^{0})$  on  $E^{0}$  is expressed as

$$V^{0}(E^{0}) = \text{const.} (E^{0})^{\frac{1}{2}N^{0}} = \text{const.} (E^{*} - E)^{\frac{1}{2}N^{0}}.$$
 (I.36)

Using this result in Eq. (I.34) we have

$$P(q_1,\ldots,p_f) = \text{const.} (E^*) E^{\frac{3}{2}N^0 - 1} (1 - \frac{E}{E^*})^{\frac{3}{2}N^0 - 1}, \quad (I.34')$$

A. I]

APPENDIX I

where the proportionally constant may depend on  $E^*$  but not on E.

So far we have made no approximations. We shall now make use of the fact that  $N^0$  is extremely large and  $E/E^*$  is very small.

First of all, the exponent  $(\frac{3}{2}N^0 - 1)$  of the last of Eq. (I.34') can be approximated by  $\frac{3}{2}N^0$  and

$$P(q_1, \dots, p_f) = (\text{const.}) \left(E^{\bullet}\right)^{\frac{1}{2}N^0 - 1} \left(1 - \frac{E}{E^{\bullet}}\right)^{\frac{1}{2}N^0}$$
$$= (\text{const.})' \left(1 - \frac{2}{3N^0} \frac{3N^0}{2E^{\bullet}} E\right)^{\frac{3}{2}N^0}. \quad (I.34'')$$

We note here that the thermal reservoir must be much larger than the dynamical system immersed in it and, hence, that the energy of the reservoir is, in almost all the phases of the composite system, much larger than the energy of the dynamical system. In order that the dynamical system have the definite average energy E the total energy  $E^*$  of the composite system must be larger for larger  $N^0$  in such a way that  $3N^0/E^*$  remains constant. Keeping this in mind, we approximate Eq. (I.34") for large  $N^0$ , using the formula

$$\lim_{n \to \infty} \left( 1 - \frac{x}{n} \right)^n = e^{-x} , \qquad (I.37)$$

to obtain

$$P(q_1, q_2, \dots, p_f) = \text{const. e}^{-\frac{3N^0}{2E^*}E},$$
 (I.34"")

or, by putting

$$\frac{3N^0}{2E^*} = \frac{1}{\theta}, \qquad (I.38)$$

we can write this as

$$P(q_1, q_2, \ldots, p_f) = \text{const. } e^{-\frac{\mu}{\theta}} . \tag{I.39}$$

When the thermal reservoir is not an ideal gas, the volume  $V^0(E^0)$  can not be calculated so easily. In any case, however, this quantity is the volume of a domain in some phase space of a great many dimensions and its over-all radius will certainly increase with the energy  $E^0$ . The volume will then increase very rapidly with  $E^0$ , i.e., with a very high power of  $E^0$ . We can not claim that  $V^0(E^0)$  will in general be represented accurately by a simple form like const.  $(E^0)^{\alpha N^0}$ , with  $N^0$  the number of molecules and  $\alpha$  a constant, but in many cases  $V^0(E^0)$  can still be

A. I

approximated over a considerable range of  $E^{0}$  by  $C \cdot (E^{0})^{\alpha N^{0}}$  with a proper choice of the constants C and  $\alpha$ . In other words, if we determine C and  $\alpha$  by requiring that  $V^{0}(E^{0})$  and  $dV^{0}/dE^{0}$  at  $E^{0} = E^{*}$  coincide with  $C \cdot (E^{0})^{\alpha N^{0}}$  and  $dC \cdot (E^{0})^{\alpha N^{0}}/dE^{0}$ , respectively,  $V^{0}(E^{0})$  will be well approximated by  $C \cdot (E^{0})^{\alpha N^{0}}$  over some range of  $E^{0}$  in the vicinity of  $E^{0} = E^{*}$ . Choosing, therefore, C and  $\alpha$  so as to satisfy

$$C \cdot (E^{\bullet})^{\alpha N^{0}} = V^{0}(E^{\bullet})$$

$$C \alpha N^{0} \cdot (E^{\bullet})^{\alpha N^{0}-1} = \left(\frac{\mathrm{d}V^{0}}{\mathrm{d}E^{0}}\right)_{E^{0}=E^{\bullet}},$$
(I.40)

we shall employ the approximation

$$V^{0}(E^{0}) = C \cdot (E^{0})^{\alpha N^{0}} . \tag{I.36'}$$

With this approximation, the arguments given for the ideal gas hold equally for our general case and we shall arrive at the same result, Eq. (I.39), except for a difference in the meaning of  $\theta$ . Namely, in this case,  $\alpha$  has to satisfy Eq. (I.40) and hence

$$\frac{1}{\theta} = \frac{\alpha N^0}{E^*} = \frac{1}{V^0(E^*)} \left(\frac{\mathrm{d}V^0(E^0)}{\mathrm{d}E^0}\right)_{E^0 = E^*}$$
$$= \left(\frac{\mathrm{d}\log V^0}{\mathrm{d}E^0}\right)_{E^0 = E^*}.$$
(I.41)

Thus the formula (I.39) becomes of general applicability.

We shall now investigate the physical meaning of  $\theta$ . For this purpose we consider a thermometer inserted in a thermal reservoir. Specifically we take as the thermometer a unit mole of an ideal gas contained in a closed box of unit volume and measure the temperature by its pressure. Considering this ideal gas as the dynamical system in the above arguments, the probability function for these gas molecules is given by Eq. (I.39) with  $\theta$  defined by Eq. (I.41). Then we have the velocity distribution of molecules in this ideal gas, from which we can calculate, through the elementary kinetic theory of gases, the pressure on the wall of the containing box. We shall not repeat this calculation but merely quote the result that the pressure is given by

$$P = N heta$$
 ,  $(I.42)$ 

where N is the number per mole of the gas molecules.

APPENDIX II

Now, the fact that the gas pressure is given by Eq. (I.42) indicates that the temperature of the gas has the value

$$T = \frac{P}{R} = \frac{\theta}{k}, \qquad (I.43)$$

where R is the gas constant and k = R/N; we can then interpret  $\theta$  as

$$\theta = kT \,. \tag{I.43'}$$

and Eq. (I.39) becomes

$$P(q_1, q_2, \dots, p_f) = \text{const. } e^{-E/kT}$$
. (I.39')

The proportionality constant can be determined by the requirement that the total probability for  $q_1, \ldots p_f$  to assume any of their possible values is unity. The constants appearing in Eqs. (2.1) and (2.2) in the text have been calculated in this way.

Through the considerations of this section, one can clearly see that it is simply the approximation formula (I.37) which gives rise to the magical factor  $e^{-E/kT}$  which appears almost everywhere in statistical mechanics.

### APPENDIX II

## A Simple Model of a Crystal

Since the vibrations of a general crystal lattice are too complicated



Fig. A.2. A simple model of a onedimensional crystal lattice.

for us to deal with, we consider a simple one-dimensional model of a crystal in order to give a general idea of the problem.

The model consists of N identical balls of mass m arranged with equal inter-separations on an elastic string of length L which is fixed at both its ends, as in Fig. (A.2). Let the tension of the string be  $\kappa$ . The mass of the string will be neglected. For simpli-

city, we consider only transverse vibrations of this dynamical system. The balls are labelled  $1, 2, 3, \ldots, l, \ldots, N$  from the left and their coordinates, measured from their equilibrium positions, are designated by  $q_1, q_2, q_3, \ldots, q_l, \ldots, q_N$ , respectively. In general, the distance between the (l-1)-th and the *l*-th ball is given by  $\sqrt{a^2 + (q_l - q_{l-1})^2}$ , where *a* is L/(N+1). Since the distance at equilibrium is *a*, the increase in potential energy due to this displacement is  $\kappa \{\sqrt{a^2 + (q_l - q_{l-1})^2} - a\}$ . The potential energy of the whole system is then given by

$$E_{pot} = \kappa \sum_{l} \{ \sqrt{a^2 + (q_l - q_{l-1})^2} - a \},$$

which, in the case of small displacements, reduces to

$$E_{pol} = \frac{\kappa}{2a} \sum_{l} (q_{l} - q_{l-1})^{2} .$$
 (II.1)

The kinetic energy is obviously

$$E_{\rm kin} = \frac{1}{2}m\sum_{l}\dot{q}_{l}^{2}$$
, (II.2)

and the total energy becomes

$$E = \frac{1}{2}m\sum_{l}\dot{q}_{l}^{2} + \frac{\kappa}{2a}\sum_{l}\left(q_{l} - q_{l-1}\right)^{2}.$$
 (II.3)

We shall now use, for describing the system, the N quantities  $Q_1, \ldots, Q_s, \ldots, Q_N$  defined by

$$q_l = \sum_{s=1}^{N} Q_s \sin \frac{ls\pi}{N+1}, \quad l = 1, 2, \dots, N.$$
 (II.4)

These N relations determine the Q's uniquely from the given q's and vice versa, and hence these variables  $Q_1, Q_2, \ldots, Q_N$  can as well be used to describe the system.

Substituting Eq. (II.4) into Eqs. (II.1) and (II.2), we obtain after some straightforward calculation

$$E_{\text{pot}} = \frac{N}{a_{\kappa}} \sum_{s} \sin^{2} \frac{s\pi}{2(N+1)} \cdot Q_{s}^{2}$$

$$E_{\text{kin}} = \frac{1}{4} Nm \sum_{s} Q_{s}^{2}.$$
(II.5)

We denote the momentum conjugate to the coordinate  $Q_s$  by  $P_s$  and by definition this is given by

$$P_s = \frac{\partial E_{\rm kin}}{\partial Q_s} = \frac{1}{2} Nm \, Q_s \, .$$

Expressing  $E_{kin}$  in terms of the  $P_s$ 's we have

$$E_{\rm kin} = \frac{1}{Nm} \sum_{s} P_s^2$$

and the total energy of the system becomes

$$E = \sum_{s} \left\{ \frac{1}{Nm} P_s^2 + \frac{N\kappa}{a} \sin^2 \frac{s\pi}{2(N+1)} \cdot Q_s^2 \right\}.$$
 (II.6)

This expression has the standard form of Eq. (2.13) of the text and, thus, these Q's constitute the normal coordinates of our system.

Using this expression as the Hamiltonian, we can write Hamilton's canonical equations of motion for the P's and Q's, i.e.,

$$\dot{Q}_{s} = \frac{\partial E}{\partial P_{s}} = \frac{2}{Nm} P_{s}$$

$$\dot{P}_{s} = -\frac{\partial E}{\partial Q_{s}} = -\frac{2N\kappa}{a} \sin^{2} \frac{s\pi}{2(N+1)} \cdot Q_{s}.$$
(II.7)

The solution is readily obtained as

$$Q_s = A_s \cos(2\pi r_s t + \delta_s), \quad s = 1, 2, \dots, N$$
, (II.8)

where  $A_s$  and  $\delta_s$  are the integration constants and the frequency  $v_s$  is given by

$$v_s = \frac{1}{\pi} \sqrt{\frac{\kappa}{ma}} \sin \frac{s\pi}{2(N+1)}, \quad s = 1, 2, \dots, N.$$
 (II.9)

The original coordinates  $q_1, \ldots, q_N$  are obtained from these Q's as

$$q_{t} = \sum_{s=1}^{N} A_{s} \sin \frac{ls\pi}{(N+1)} \cos \left(2\pi \nu_{s} t + \delta_{s}\right).$$
(II.10)

The coordinate of each ball is thus expressed as a superposition of Nharmonic oscillations with frequencies given by Eq. (II.9). These N harmonic oscillations are the proper oscillations of our dynamical system. When the system is in its s-th proper state of motion,  $q_l$  is given by

$$q_{l} = A_{s} \sin \frac{l s \pi}{N+1} \cos \left(2 \pi v_{s} t + \delta_{s}\right).$$

The N proper frequencies given by Eq. (II.9) are not equally spaced, in contrast to those of a continuous string which will be treated in the

next section. The frequencies for the case N = 8 are shown in Fig. (A.3). The maximum frequency for this case is given by

$$v_{\max} = \frac{1}{\pi} \sqrt{\frac{\kappa}{ma}} \sin \frac{N\pi}{2(N+1)}.$$
 (II.11)

It is seen from Eq. (II.9) that the separation between the neighbouring frequencies,  $\Lambda$ , is for very large N,

$$\Delta = \frac{\pi}{2N} \sqrt{r_{\text{max}}^2 - \nu^2} . \qquad (\text{II.12})$$

Accordingly, the number,  $Z(\nu) d\nu$ , of proper oscillations with frequencies between  $\nu$  and  $\nu + d\nu$  is given by

$$Z(\nu) \, \mathrm{d}\nu = \frac{2N}{\pi \sqrt{\nu_{\max}^2 - \nu^2}} \, \mathrm{d}\nu \,. \quad (\text{II.13})$$

This is the expression corresponding to

$$Z(\boldsymbol{\nu}) \, \mathrm{d}\boldsymbol{\nu} = \frac{2L}{c} \, \mathrm{d}\boldsymbol{\nu}$$

for the case of continuous string.

## APPENDIX III

# Transverse Oscillations of a Continuous String

Consider a string of length L, stretched with tension  $\kappa$ . The string is fixed at both its ends. Its linear density is denoted by  $\sigma$ . Taking the left



Fig. A.4. The vibration of a string.

end of the string as the origin of the coordinates, the position x is measured to the right along the string as in Fig. (A.4). The transverse displacement of the line element of the string at x is denoted by q(x). Due to this

displacement the element of string dx at this position is elongated and becomes

$$\sqrt{1+\left(\frac{\partial q}{\partial x}\right)^2}\,\mathrm{d}x\,.$$



Fig. A.3. The distribution of frequencies of the normal vibrations.

APPENDIX III

Accordingly this part has a potential energy

$$\kappa \left\{ \sqrt{1 + \left(\frac{\partial q}{\partial x}\right)^2} - 1 \right\} \mathrm{d}x = \frac{\kappa}{2} \left(\frac{\partial q}{\partial x}\right)^2 \mathrm{d}x \,.$$

The potential energy of the whole string is then

$$E_{\text{pot}} = \frac{\kappa}{2} \int_0^L \left(\frac{\partial q}{\partial x}\right)^2 \, \mathrm{d}x \,, \qquad (\text{III.1})$$

while the kinetic energy is obviously

$$E_{kin} = \frac{\sigma}{2} \int_{0}^{L} \left(\frac{\partial q}{\partial t}\right)^{2} \mathrm{d}x \;. \tag{III.2}$$

The total energy of the system is the sum of the two, i.e.,

$$E = \frac{\kappa}{2} \int_{0}^{L} \left(\frac{\partial q}{\partial x}\right)^{2} dx + \frac{\sigma}{2} \int_{0}^{L} \left(\frac{\partial q}{\partial t}\right)^{2} dx .$$
(III.3)

We shall now introduce the normal coordinates of the system. For this purpose we expand q(x) in a Fourier series

$$q(x) = \sum_{s=0}^{\infty} Q_s \sin \frac{s\pi x}{L}.$$
 (III.4)

By the way, the non-occurrence of the cosine terms is due to the fact that the string is fixed at both ends. Eq. (III.4) enables us to determine q(x) when its Fourier coefficients,  $Q_1, Q_2, \ldots, Q_s, \ldots$  are given and conversely to determine the Q's when q(x) is given. We can then consider these Q's as the coordinates of the system. The fact that there are an infinite number of these Q's reflects the infinite number of degrees of freedom of our dynamical system. Introducing Eq. (III.4) into Eq. (III.1) as well as into Eq. (III.2), we obtain

$$E_{pol} = \frac{\pi^2 \kappa}{4L} \sum_{s=1}^{\infty} s^2 Q_s^2 ,$$

$$E_{kin} = \frac{L\sigma}{4} \sum_{s=1}^{\infty} \dot{Q}_s^2 .$$
(III.5)

and

$$P_s = \frac{1}{2} L \sigma Q_s \, .$$

In terms of these P's, the total energy is given by

$$E = E_{\text{pot}} + E_{\text{kin}}$$
  
=  $\sum_{s} \left\{ \frac{1}{L\sigma} P_{s}^{2} + \frac{\pi^{2}\kappa}{4L} s^{2} Q_{s}^{2} \right\},$  (III.6)

which has the standard form (2.13) of the text. Therefore, these Fourier coefficients constitute the normal coordinates of our dynamical system.

From the expression (III.6) follow Hamilton's equations of motion for the Q's and P's

$$\dot{Q}_{s} = \frac{2}{L\sigma} P_{s},$$

$$s = 1, 2, \dots$$

$$\dot{P}_{s} = -\frac{\pi^{2}\kappa}{2L} s^{2}Q_{s}.$$
(III.7)

The solution is readily obtained as

$$Q_{s} = A_{s} \cos\left(2\pi \nu_{s} t + \delta_{s}\right)$$
, (III.8)

where the frequency  $v_s$  has the value

$$\boldsymbol{\nu}_s = \frac{s}{2L} \sqrt{\frac{\kappa}{\sigma}} \,. \tag{III.9}$$

The frequencies are thus seen to be equally spaced with the separation

$$\Delta = \frac{1}{2L} \sqrt{\frac{\kappa}{\sigma}}.$$
 (III.10)

Substituting Eq. (III.8) in Eq. (III.4), an oscillation is found in general to be expressed as a superposition of sinusoidal oscillations,

$$q(x) = \sum_{s=1}^{\infty} A_s \sin \frac{s\pi x}{L} \cos \left(2\pi \nu_s t + \delta_s\right). \qquad \text{(III.11)}$$

In particular, when only the s-th mode is excited

$$q(x) = A_s \sin \frac{s\pi x}{L} \cos \left(2\pi \nu_s t + \delta_s\right), \qquad (\text{III.11'})$$

APPENDIX IV

The wave length of the oscillation (III.11') is given by

$$\hat{\lambda} = \frac{2L}{s} , \qquad (\text{III.12})$$

and then the wave velocity c on this string is, from  $c = v\lambda$ , given by

$$c = \sqrt{\frac{\kappa}{\sigma}}.$$
 (III.13)

In terms of this quantity c, Eq. (III.9) becomes

$$\nu_s = \frac{c}{2L} s , \qquad (\text{III.14})$$

and Eq. (III.10) becomes

$$\Delta = \frac{c}{2L}, \qquad (\text{III.15})$$

which immediately leads to

$$Z(\nu) \,\mathrm{d}\nu = \frac{2L}{c} \,\mathrm{d}\nu \,. \tag{III.16}$$

The introduction of Eq. (III.11) into Eq. (III.5) gives the energy of the oscillation as a function of  $A_1, A_2, \ldots$ , i.e.,

$$E = \frac{1}{4}L\sigma \sum_{s} (2\pi\nu_{s})^{2}A_{s}^{2} .$$
 (III.17)

The energy per unit length of the string, the "energy density", is thus found to be

$$U = \frac{1}{4}\sigma \sum_{s} (2\pi\nu_{s})^{2} A_{s}^{2} . \qquad (\text{III.18})$$

### APPENDIX IV

# **Dynamical treatment of Electromagnetic Oscillations**

In this section we shall show that electromagnetic oscillations can also be treated, like those of an ordinary dynamical system, in terms of normal coordinates and that they can be formulated in Hamiltonian form.

Electromagnetic oscillations in vacuum satisfy Maxwell's equations, i.e.,

$$\operatorname{curl} \mathbf{E} + \frac{1}{c} \dot{\mathbf{H}} = 0, \quad \operatorname{div} \mathbf{H} = 0,$$

$$\operatorname{curl} \mathbf{H} - \frac{1}{c} \dot{\mathbf{E}} = 0, \quad \operatorname{div} \mathbf{E} = 0,$$
(IV.1)

where E and H signify the electric and magnetic field, respectively. We introduce, as usual, the vector potential A defined by

$$\mathbf{E} = -\frac{1}{c}\dot{\mathbf{A}}$$
 and  $\mathbf{H} = \operatorname{curl}\mathbf{A}$ . (IV.2)

The first half of the equations (IV.1) is, then, automatically satisfied and we are left with the wave equation

$$\Delta \mathbf{A} - \frac{1}{c^2} \ddot{\mathbf{A}} = 0 \tag{IV.3}$$

together with

$$\operatorname{div} \mathbf{A} = \mathbf{0} \,. \tag{IV.4}$$

The energy of the electromagnetic field is then given by

$$H = \frac{1}{8\pi} \int_{v} (E^{2} + H^{2}) dv$$

$$= \frac{1}{8\pi} \int_{v} \left\{ \frac{1}{c^{2}} \dot{A}^{2} + (\operatorname{curl} A)^{2} \right\} dv,$$
(IV.5)

where the integration is to be carried out over the entire volume under consideration.

Let us suppose that our oscillation is confined in a cubic cavity of side L with mirror walls which perfectly reflect light.

Both the tangential component of E and the normal component of H have to vanish at the wall and, hence, the tangential component of A

APPENDIX IV

also has to vanish there. Under this boundary condition, A(x, y, z) is expressed in a Fourier series of the form

$$A_{x}(x, y, z) = \sum_{s} Q_{s, x} \cos \frac{\pi s_{x} x}{L} \cdot \sin \frac{\pi s_{y} y}{L} \cdot \sin \frac{\pi s_{z} z}{L}$$
$$A_{y}(x, y, z) = \sum_{s} Q_{s, y} \sin \frac{\pi s_{x} x}{L} \cdot \cos \frac{\pi s_{y} y}{L} \cdot \sin \frac{\pi s_{z} z}{L} \qquad (IV.6)$$
$$A_{z}(x, y, z) = \sum_{s} Q_{s, z} \sin \frac{\pi s_{x} x}{L} \cdot \sin \frac{\pi s_{y} y}{L} \cdot \cos \frac{\pi s_{z} z}{L}$$

where  $s_x$ ,  $s_y$  and  $s_z$  denote positive integers and  $\sum_s$  stands for the summation over all possible values of these integers, i.e., for  $\sum_{s_x} \sum_{s_y} \sum_{s_z}$ . The suffix s on the Q's is likewise to be understood to mean  $Q_{s_x,s_y,s_z}$ . In what follows, we consider  $s_x$ ,  $s_y$  and  $s_z$  to be the three components of a vector s and  $Q_{s,x}$ ,  $Q_{s,y}$  and  $Q_{s,z}$  those of a vector  $Q_s$ .

When A satisfies the condition (IV.4), we have

 $(\mathbf{Q_s} \cdot \mathbf{s}) = 0$  ,

that is, the vector  $Q_s$  is normal to the vector s. This corresponds to the well known fact that light is a transverse wave. We now introduce the two unit vectors  $e_s^{(1)}$  and  $e_s^{(2)}$  which are both normal to the vector s and at the same time normal to each other. In terms of these vectors,

$$Q_s = q_s^{(1)} e_s^{(1)} + q_s^{(2)} e_s^{(2)}$$
, (IV.7)

and

$$A_{x} = \sum_{s} (q_{s}^{(1)} e_{s}^{(1)} + q_{s}^{(2)} e_{s}^{(2)})_{x} \cos \dots,$$
  

$$A_{y} = \sum_{s} (q_{s}^{(1)} e_{s}^{(1)} + q_{s}^{(2)} e_{s}^{(2)})_{y} \sin \dots,$$
  

$$A_{z} = \sum_{s} (q_{s}^{(1)} e_{s}^{(1)} + q_{s}^{(2)} e_{s}^{(2)})_{z} \sin \dots.$$
(IV.8)

The electromagnetic oscillation in the cavity is thus expressed as a superposition of linearly polarized monochromatic waves of various directions, s, of various frequencies, |s|c/2L, and of various polarizations,  $e_s^{(1)}$  and  $e_s^{(2)}$ . The realtionship (IV.8) allows us to determine the  $q^{(1)}$ 's and  $q^{(2)}$ 's from a given A and conversely to determine A from a

[A. IV

#### APPENDIX IV

given set of  $q^{(1)}$ 's and  $q^{(2)}$ 's. Therefore, we can regard the q's as the coordinates of the electromagnetic wave.

Introducing Eq. (IV.6) into Eq. (IV.3), Q, is found to satisfy

$$\frac{\pi^2 s^2 c^2}{L^2} Q_s + \ddot{Q}_s = 0 , \qquad (IV.9)$$

where

$$|s|^2 = s^2 = s_x^2 + s_y^2 + s_z^2$$
. (IV.10)

The solution is readily obtained as

$$q_s^{(1)} = a_s^{(1)} \cos \left(2\pi \nu_s t + \delta_s^{(1)}\right),$$

$$q_s^{(2)} = a_s^{(2)} \cos \left(2\pi \nu_s t + \delta_s^{(2)}\right),$$
(IV.11)

where  $a_s^{(1)}$ ,  $a_s^{(2)}$ ,  $\delta_s^{(1)}$  and  $\delta_s^{(2)}$  are the integration constants and  $\nu_s$  is given by

$$\boldsymbol{\nu}_{s} = \frac{|s|c}{2L} \,. \tag{IV.12}$$

Eq. (IV.8) together with Eq. (IV.11) gives the electromagnetic oscillation as a superposition of a number of oscillations which vary sinusoidally with time. It is then expected that these  $q^{(1)}$ 's and  $q^{(2)}$ 's constitute the normal coordinates of the electromagnetic oscillations and that each oscillation described by these coordinates is a proper oscillation of the cavity radiation. These proper oscillations are labelled by a set of three positive integers  $s_x$ ,  $s_y$ , and  $s_z$ , and, furthermore, there exist two kinds of oscillations, corresponding to  $e_s^{(1)}$  and  $e_s^{(2)}$ , for each set of  $s_x$ ,  $s_y$ , and  $s_z$ . These facts are utilized without explanation in the text.

We shall now show that in terms of these q's the energy of the electromagnetic wave actually takes the standard form (2.13). Introducing Eq. (IV.8) into Eq. (IV.5), after some straightforward calculation, we obtain the energy H as

$$H = \frac{V}{64\pi c^2} \sum_{s} \left\{ \dot{q}_s^{(1)^2} + \dot{q}_s^{(2)^2} + 4\pi v_s^2 \left( q_s^{(1)^2} + q_s^{(2)^2} \right) \right\}.$$
 (IV.13)

This energy consists of two kinds of terms, one containing  $\dot{q}$ 's and the other q's. They can be interpreted as the kinetic and potential energy,

APPENDIX V

[A. IV, V

respectively. We can define the momentum p conjugate to q according to the general procedure

$$p = rac{\partial E_{kin}}{\partial q}$$
 ,

and obtain

$$p_s^{(1)} = \mu \dot{q}_s^{(1)}$$
 and  $p_s^{(2)} = \mu \dot{q}_s^{(2)}$ , (1V.14)

where

$$\mu=\frac{V}{32\pi c^2}\,.$$

By means of Eq. (IV.14), the energy (IV.13) is reduced to

$$H = \frac{1}{2} \sum_{s} \left\{ \frac{p_s^{(1)^2}}{\mu} + 4\pi v_s^2 q^{(1)^2} \right\} + \frac{1}{2} \sum_{s} \left\{ \frac{p^{(2)^2}}{\mu} + 4\pi v_s^2 q^{(2)^2} \right\}, \quad (IV.15)$$

which in fact is seen to have the standard form (2.13) of the text. It is thus ascertained that our q's constitute the normal coordinates for the electromagnetic oscillation.

We further see that Eqs. (IV.14) and (IV.11) can actually be obtained from Hamilton's equations of motion with Eq. (IV.15) as the Hamiltonian, namely from

$$p_s^{(1)} = -\frac{\partial H}{\partial q_s^{(1)}}, \qquad \dot{q}_s^{(1)} = \frac{\partial H}{\partial \dot{p}_s^{(1)}},$$
$$\dot{p}_s^{(2)} = -\frac{\partial H}{\partial q_s^{(2)}}, \qquad \dot{q}_s^{(2)} = \frac{\partial H}{\partial \dot{p}_s^{(2)}}.$$
(IV.16)

Therefore, we may conclude that the electromagnetic oscillations are not at all different from the oscillations of ordinary dynamical systems.

#### APPENDIX V

# Pressure due to Transverse Vibrations of a String

It may sound rather queer that electromagnetic oscillations in a cavity, or vibrations of a string, can give rise to some pressure on the "walls", since at first sight one may get the impression that an oscillation being a repetition of to and fro motions, gives positive and negative effects alternately, to result finally in a vanishing pressure.

However, it will be shown that the statement "an oscillation gives sometimes positive and sometimes negative effects" is erroneous. This has already been shown in the text for the case of a pendulum.

As a further illustration, we calculate the pressure due to the transverse oscillation of a string. The problem is schematically illustrated in Fig. (A.5). A movable partition plate is inserted between the two end plates which fix the string at both its ends Q and R. The string goes through a hole P in the partition plate. We let oscillations occur

only in the region PR. In order not to have a longitudinal oscillation in QP caused by the transverse oscillation in PR, we consider that QP of the string is non-expansible. When there are no oscillations in PR, there



Fig. A.5. A string with a movable partition plate.

is obviously no force acting on the moving plate. However, when PR is oscillating, it can be shown that there is a force pushing the partition plate toward the end plate Q. This is the pressure of the oscillation.

To calculate this pressure, we take a close look at what is happening at the hole P. Fig. (A.6) shows the situation when the string is instantaneously making an angle  $\theta$  with respect to the hole P. Since the length of the string in the vicinity of the hole P is in this state increased by a factor sec  $\theta$ , the tension T of the string is  $\kappa \sec \theta$ . Consider a string



Fig. A.6. The force acting on a movable partition plate.

element O at the very end of the hole P, as in Fig. (A.6). The forces acting on it are the following: first, there is a tension OB of the magnitude  $T = \kappa \sec \theta$  along the right hand part of the string and, second, another tension OD of the same magnitude along the left hand part of the string. The aforementioned non-expansibility of the string in QP, by the way, prevents the pulling-in of this part into the right hand

side of the hole, i.e., it prevents the occurrence of longitudinal oscillations on the left hand side of the hole. The resultant of the above two forces is OC in Fig. (A.6) which has a component OE pushing the partition plate to the left. That is, the string element O is pushing the plate to the left. This is the net force the plate is receiving at this moment and its magnitude, as seen from Fig. (A.6), is given by

A. V]

 $F = \kappa (\sec \theta - 1)$ .

Since also

$$\sec \theta = \left[ \sqrt{1 + \left(\frac{\partial q}{\partial x}\right)^2} \right]_{x=0}$$
$$\approx 1 + \frac{1}{2} \left(\frac{\partial q}{\partial x}\right)_{x=0}^2,$$
$$F = \frac{\kappa}{2} \left(\frac{\partial q}{\partial x}\right)_{x=0}^2.$$
(V.I)

we have

Introducing Eq. (III.11) into this expression, we obtain

$$F = \frac{\pi^2 \kappa}{2L^2} \sum_{s} \sum_{s'} A_s A_{s'} ss' \cos(2\pi \nu_s t + \delta_s) \cos(2\pi \nu_s t + \delta_{s'}) .$$

This instantaneous force is a complicated function of time but the quantity of concern to us, the pressure, is its time average, which is

$$P = \langle F \rangle = \frac{\pi^2 \kappa}{4L^2} \sum_s s^2 A_s^2 , \qquad (V.2)$$

or, using Eq. (III.9),

$$P = \frac{1}{4}\sigma \sum_{s} (2\pi\nu_{s})^{2}A_{s}^{2}.$$
 (V.3)

Recalling the result for the energy density U, Eq. (III.8), we arrive at the conclusion

$$P = U , \qquad (V.4)$$

which is to be compared with

$$P = \frac{1}{3}U$$

in the case of cavity radiation. The reason that the factor  $\frac{1}{3}$  does not appear in our result is simply that our present problem is a one-dimensional one.

It has been shown above that the effect of the transverse oscillation of a string on the partition plate is not "sometimes positive and some-

times negative". What, then, gives rise to the same situation in the case of electromagnetic oscillations?. We shall not go into detailed calculations here but instead give a brief account of it. When an electromagnetic oscillation hits the wall, the electric field in its vicinity takes positive and negative values alternately. When the field is at some instant positive, or negative, there will be induced on the surface of the wall some negative, or positive, charge respectively. Since the electric force is the product of the electric field and the electric charge, this quantity is always negative irrespective of the sign of the varying field, exerting thus a negative pressure on the wall. But if we apply similar arguments to the magnetic field we shall see that the magnetic force is always positive and the resultant force acting on the wall due to the electromagnetic oscillation becomes positive, i.e., it is a pressure.

Now we shall make some remarks concerning adiabatic invariants. Suppose the length of the oscillating string is increased by an amount dL by displacing the movable partition by dL. The work performed through this displacement, which is equal to the increase in the oscillational energy, is given by

or, using Eq. (V.4),

 $\mathrm{d}E = -P \,\mathrm{d}L \,.$  $\mathrm{d}E = - U \,\mathrm{d}L \;.$ 

The equation is further transformed, by E = UL, to

$$\frac{\mathrm{d}E}{E} = -\frac{\mathrm{d}L}{L}\,.$$

On the other hand, from Eq. (III.9),

$$\frac{\mathrm{d}\boldsymbol{v}_{s}}{\boldsymbol{v}_{s}} = -\frac{\mathrm{d}L}{L}$$

and, hence,

 $\frac{\mathrm{d}E}{E} = \frac{\mathrm{d}\mathbf{v}_{\mathrm{s}}}{\mathbf{v}_{\mathrm{s}}},$ 

which implies

 $\frac{E}{v} = \text{constant}$ . (V.6)

A. V]

APPENDIX VI

$$\frac{E_s}{v_s} = \text{constant}$$
 (V.7)

cannot be obtained by mere energy balance considerations. This proof will be given in the next appendix.

### APPENDIX VI

## Adiabatic Invariants of an Oscillating System

Consider an arbitrary oscillating system which may be a simple one like a pendulum or a very complex one like the electromagnetic wave. We shall prove that  $E_s/\nu_s$  for each proper oscillation remains invariant under infinitely slow changes of the system caused by some external effects.

Let the normal coordinates of the system be  $q_1, q_2, \ldots, q_s, \ldots$ . Then the energy of the system, H, will have the form

$$H = \sum_{s} E_{s} ,$$

$$E_{s} = \frac{1}{2} (\alpha_{s} \dot{q}_{s}^{2} + \beta_{s} q_{s}^{2}) .$$
(VI.1)

We shall examine the case where no external influences act on the system. Although we have already shown that in this case  $q_s$  performs a sinusoidal motion, we shall here for later convenience, use a different approach. We define new coordinates by

$$\sqrt{\alpha_{\rm s}} \, q_{\rm s} = Q_{\rm s} \,. \tag{VI.2}$$

In terms of this  $Q_s$ ,

$$E_s = \frac{1}{2} \left( \dot{Q}_s^2 + \frac{\dot{\beta}_s}{\alpha_s} Q_s^2 \right), \qquad (VI.3)$$

and the equation of motion becomes

$$\ddot{Q}_s = - (2\pi\nu_s)^2 Q_s , \qquad (\text{VI.4})$$

where  $v_s$  is given by

$$v_s = \frac{1}{2\pi} \sqrt{\frac{\beta_s}{\alpha_s}}.$$
 (VI.5)

The solution is readily obtained as

$$Q_s = A_s \cos\left(2\pi\nu_s t + \delta_s\right), \qquad (\text{VI.6})$$

where the amplitude  $A_s$  and the phase  $\delta_s$  are the integration constants. The energy of the oscillation as described by Eq. (VI.6) is found to be

$$E_s = \frac{1}{2} (2\pi \nu_s)^2 A_s^2 . \qquad (VI.7)$$

We shall now allow an external influence to act on the system and to change it very slowly. Mathematically this process is described by assuming  $\alpha_s$  and  $\beta_s$  of Eq. (VI.1) to be functions of time, because these quantities are functions of length, tension, and mass density of the string which are to be varied by an external influence. By very slow variation we mean that  $\alpha_s$  and  $\beta_s$  can be treated as constants over a considerable number of periods of motion, or more specifically, that

$$\frac{\alpha_s}{\alpha_s} \ll v_s$$
, and  $\frac{\beta_s}{\beta_s} \ll v_s$ . (VI.8)

With these properties of  $\alpha_s$  and  $\beta_s$  in mind, we again define  $Q_s$  according to Eq. (VI.2). Then we have

$$\dot{q}_s = \frac{1}{\sqrt{\alpha_s}} \left( \dot{Q}_s - \frac{1}{2} \frac{\dot{\alpha}_s}{\alpha_s} Q_s \right)$$

and inserting this into Eq. (VI.1) we obtain

$$E_s = \frac{1}{2} \left\{ \left( \dot{Q}_s - \frac{1}{2} \frac{\dot{\alpha}_s}{\alpha_s} Q_s \right)^2 + \frac{\beta_s}{\alpha_s} Q_s^2 \right\}.$$
(VI.3')

Our result (VI.3') differs from Eq. (VI.3), but taking Eq. (VI.8) into account we arrive at

$$\frac{1}{2}\frac{\alpha_s}{\alpha_s}Q_s\ll\dot{Q}_s,$$

since  $\dot{Q}_s$  is of the order of  $\nu_s Q_s$ , and thus Eq. (VI.3') can be approximated by

$$E_s = \frac{1}{2} \left( \dot{Q}_s^2 + \frac{\beta_s}{\alpha_s} Q_s^2 \right)$$

A. VI]

APPENDIX VI

which again has the form (VI.3). This expression for the energy leads to equations of motion

$$\ddot{Q}_s = -(2\pi\nu_s)^2 Q_s$$
,  
 $\nu_s = \frac{1}{2\pi} \sqrt{\frac{\beta_s}{\alpha_s}}$ .
(VI.9)

Here, however,  $\alpha_s$  and  $\beta_s$  are both functions of time and, hence, so is  $v_s$ . It is in general quite difficult to solve Eq. (VI.9) with time dependent  $v_s$  but an approximate solution can be obtained with the aid of condition (VI.8).

We proceed as follows. Due to the time dependence of  $v_s$  the motion is no longer sinusoidal as given by Eq. (VI.6), but it will still be represented within a considerable number of periods of motion by a sinusoidal function and its amplitude, phase, and frequency will vary appreciably only after a time much longer than a period. Thus we shall represent the solution of Eq. (VI.9) in the form

$$Q_{s}(t) = A_{s}(t) \cos \{ 2\pi v_{s}(t)t + \delta_{s}(t) \}, \qquad (\text{VI.10})$$

and proceed to determine  $A_s(t)$  and  $\delta_s(t)$  for a given  $v_s(t)$ . Since in Eq. (VI.10) two unknown functions  $A_s(t)$  and  $\delta_s(t)$  are used to express only one unknown function  $Q_s(t)$ , we can impose one arbitrary condition to be satisfied by  $A_s(t)$  and  $\delta_s(t)$ . For convenience in later calculations, we choose the condition

$$\dot{A}_s \cos \left(2\pi \nu_s t + \delta_s\right) = \left(2\pi \nu_s t + \dot{\delta}_s\right) A_s \sin \left(2\pi \nu_s t + \delta_s\right) \qquad (\text{VI.11})$$

and consider  $A_s(t)$  as unknown,  $\delta_s$  being determined by this condition. The convenience of this choice will be seen shortly.

In order to find  $A_s(t)$ , we introduce Eq. (VI.10) into the equation of motion (VI.9). From Eq. (VI.10) we have

$$\begin{split} \dot{Q}_s &= -2\pi \nu_s A_s \sin\left(2\pi \nu_s t + \delta_s\right) \\ &+ \dot{A}_s \cos\left(2\pi \nu_s t + \delta_s\right) - A_s (2\pi \dot{\nu}_s t + \dot{\delta}_s) \sin\left(2\pi \nu_s t + \delta_s\right) , \end{split}$$

which is simplified by the condition (VI.11) to

$$Q_s = -2\pi \nu_s A_s \sin\left(2\pi \nu_s t + \delta_s\right). \qquad (VI.12)$$

Furthermore,  $\ddot{Q}_s$  is given, again after the use of Eq. (VI.11), by

$$\ddot{Q}_s = -2\pi (\dot{\nu}_s A_s + \nu_s \dot{A}_s) \sin (2\pi\nu_s t + \delta_s) - (2\pi\nu_s)^2 Q_s$$
$$-2\pi\nu_s \dot{A}_s \frac{\cos^2 (2\pi\nu_s t + \delta_s)}{\sin (2\pi\nu_s t + \delta_s)}.$$

Substituting this  $Q_s$  into Eq. (VI.9), we obtain

$$\dot{\nu}_s A_s + 2\nu_s A_s = \dot{\nu}_s A_s \cos(4\pi\nu_s t + 2\delta_s)$$
 (VI.13)

for the equation to be satisfied by  $A_s(t)$ .

We now calculate the change of  $E_s/v_s$ . Introducing Eqs. (VI.10) and (VI.12) into Eq. (VI.3") we obtain

$$E_{s} = \frac{1}{2} (2\pi \nu_{s})^{2} A_{s}^{2}$$

which is of the same from as Eq. (VI.3). Then we get

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{E_s}{v_s}\right) = 2\pi^2 A_s (\dot{v}_s A_s + 2v_s \dot{A}_s)$$

or by making use of Eq. (VI.13) together with Eq. (VI.3)

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{E_s}{v_s}\right) = \left(\frac{E_s}{v_s}\right)\frac{\dot{v}_s}{v_s}\cos\left(4\pi v_s t + 2\delta_s\right). \quad (\mathrm{VI.14})$$

In order to express mathematically the fact that the change of  $v_s$  is infinitely slow, we replace  $v_s(t)$  by  $v_s(t/T)$  and let T tend to infinity. Thus v(t/T) changes from  $v_s(0)$  to  $v_s(1)$  when t changes from 0 to T. By letting T be infinitely large, this change become infinitely slow.

Now the change of  $E_s/\nu_s$  caused by changing  $\nu_s$  from  $\nu_s(0)$  to  $\nu_s(1)$  is obtained by integrating Eq. (VI.14),

$$\left(\frac{E_s}{v_s}\right)_T = \left(\frac{E_s}{v_s}\right)_0 e^{F(T)}$$
 (VI.15)

where

$$F(T) = \int_{0}^{T} \frac{\mathrm{d}\nu_{s}(t/T)}{\mathrm{d}t} \frac{1}{\nu_{s}(t/T)} \cos\{4\pi\nu_{s}(t/T)\cdot t + 2\delta_{s}(t/T)\}\,\mathrm{d}t \,. (VI.16)$$

We now prove that for  $T \to \infty$  this F(T) vanishes.

To this purpose we transform the integration variable from t to  $\tau$  defined by

$$\tau = t/T$$
.

Then we get

$$F(T) = \int_0^1 \frac{\mathrm{d} \nu_s(\tau)}{\mathrm{d} \tau} \frac{\mathrm{l}}{\nu_s(\tau)} \cos \left\{ 4\pi \nu_s(\tau) \cdot T\tau + 2\delta_s(\tau) \right\} \mathrm{d} \tau \,.$$

If T is very large the cosine term in the integrand is a very rapidly oscillating function of  $\tau$ , while  $v_s(\tau)$  and  $\delta_s(\tau)$  are slowly varying ones. This fact gives rise to a vanishingly small value of F(T). To see this one has just to observe the fact that in each period of the cosine one may replace  $v_s(\tau)$  and  $\delta_s(\tau)$  by their mean values during this period, since the period is very small, of the order of  $1/Tv_s$ , and therefore neither  $v_s$  nor  $\delta_s$  changes appreciably. The necessary correction to be added to this approximate integrand is of the order of  $1/Tv_s$ . It is obvious that in this approximation F(T) vanishes, since in each period  $v_s$  and  $\delta_s$  are replaced by constants so that the integral vanishes in each period. On the other hand the correction term to F(T) is only a small small quantity of the order of  $1/Tv_s$ , that is the order of the integration interval. We see thus that F(T) is a small quantity of the order of  $1/Tv_s$  which vanishes for infinite T.

It is instructive to apply the general method given in § 19(iii) to our dynamical system, but this problem will be left to the reader.

Having thus proved the existence of the adiabatic invariants for the general oscillating system, the pressure of an oscillation can be calculated by considering, as in the text, the energy balance without making a close investigation of what is happening at the wall.

The approach in this section for proving the theorem of adiabatic invariance is quite elementary but a more general and more elaborate proof is found in § 19 of the text.

#### APPENDIX VII

# Fluctuation of the Energy in the Oscillation of a String

We shall show for the oscillation of a string that a fluctuation of the energy occurs in the wave theory due to a kind of beat and that its mean square  $\langle \Delta E^2 \rangle$  satisfies

$$\langle \Delta E^2 \rangle = \frac{\langle E^2 \rangle}{Z_{\nu}(\nu) \, \mathrm{d}\nu},$$
 (VII.1)

### OSCILLATION OF A STRING

where  $\langle E^2 \rangle$  is the mean square of the energy and  $Z_v(v) dv$  is the number of proper oscillations of frequency between v and v + dv in the chamber of volume v. In the case of cavity radiation,

$$Z_v(\mathbf{v}) \,\mathrm{d}\mathbf{v} = rac{8\pi v}{c^3} \,\mathbf{v}^2 \,\mathrm{d}\mathbf{v}$$
 ,

and in the case of a string oscillation

$$Z_{l}(\nu) \,\mathrm{d}\nu = \frac{2l}{c} \,\mathrm{d}\nu \;.$$

In the text we considered a small chamber v in a large one V. The small chamber is surrounded by walls and there is a small opening in one of the walls. From considerations more general than that in the text the same result (VII.1) for the energy fluctuation can be obtained without specifically setting up the small chamber, i.e., by merely considering the energy in an imaginary small part v of the large chamber V. The actual calculation of the fluctuation is simpler in the latter scheme, and we shall, in the following, derive the equation (VII.1) for the oscillation of string by means of this scheme.

For a stretched string of length L, we consider a small part l which can in principle be anywhere on the string. For simplicity however we take the part l at the left end of the string.

The energy contained in this part of the string is given by changing the integration limit of Eq. (III.3) from (0, L) to (0, l) i.e., by

$$E_{l} = \frac{\kappa}{2} \int_{0}^{l} \left(\frac{\partial q}{\partial x}\right)^{2} dx + \frac{\sigma}{2} \int_{0}^{l} \left(\frac{\partial q}{\partial t}\right)^{2} dx .$$
 (VII.2)

We shall now extract from this energy the part which is due to the oscillations of frequency between v and v + dv. This part of the energy will be denoted simply by E, without any suffix, to avoid excessive complication. The time variation of this E is what we shall be concerned with, and it will be seen that E in fact shows a "beat" phenomenon and that the mean square of this "beat" around the average value is given by Eq. (VII.1).

We shall impose the following conditions for proving these facts:

(1) The length l under consideration is very much smaller than the total length L of the string; i.e.,  $l \ll L$ .

A. VII]

APPENDIX VII

(3) The range of frequency  $d\nu$  is sufficiently large to contain a sizable number of proper oscillations of the string of length l; i.e.,  $d\nu \gg c/2l$ .

(4)  $d\nu$  is, however, small enough so that the intensity of oscillation does not vary appreciably over this frequency range. Denoting the amplitude of the oscillation as a function of  $\nu$  by  $A(\nu)$ ,  $A(\nu)$  can thus be considered as constant over the frequency range  $d\nu$  and it will be denoted by  $A_{\nu}$ .

In what follows, the limits of the frequency range v and v + dv are sometimes replaced by  $v_1$ , and  $v_2$ , respectively. This is to avoid a possible confusion with the symbol of integration variable dv.

We shall, to begin with, calculate the time variation of E which is to be obtained by introducing Eq. (III.11) into Eq. (VII.2). After some algebraic calculations we have

$$E = \frac{1}{4}\sigma \sum_{s=1}^{\infty} \sum_{s'=1}^{\infty} (2\pi v_s) (2\pi v_{s'}) A_s A_{s'}$$

$$\cdot \left[ \frac{\sin\left\{\frac{2\pi}{c} (v_s - v_{s'})l\right\}}{\frac{2\pi}{c} (v_s - v_{s'})} \cos\left\{2\pi (v_s - v_{s'})t + \delta_s - \delta_{s'}\right\} + \frac{\sin\left\{\frac{2\pi}{c} (v_s + v_{s'})l\right\}}{\frac{2\pi}{c} (v_s + v_{s'})} \cos\left\{2\pi (v_s + v_{s'})t + \delta_s + \delta_{s'}\right\} \right],$$

where it is understood that

$$\frac{\sin\left\{\frac{2\pi}{c}\left(\mathbf{v}_{s}-\mathbf{v}_{s'}\right)l\right\}}{\frac{2\pi}{c}\left(\mathbf{v}_{s}-\mathbf{v}_{s'}\right)}=l$$

for s = s', since this factor originates with the integral

$$\int_{0}^{1} \cos \{ 2\pi/c \ (v_{s} - v_{s'}) \} x \, dx .$$

#### A. VII]

This expression gives the total energy contained in l but what we are seeking is that part resulting from the oscillations of frequencies between  $v_1$  and  $v_2$ :

$$E = \frac{1}{4}\sigma \sum_{v_{1} < v_{s} < v_{s}} \sum_{v_{1} < v_{s}' < v_{2}} (2\pi v_{s}) (2\pi v_{s}') A_{s} A_{s'}$$

$$\cdot \left[ \frac{\sin\left(\frac{2\pi}{c} (v_{s} - v_{s'})l\right)}{\frac{2\pi}{c} (v_{s} - v_{s'})} \cos\left\{2\pi (v_{s} - v_{s'})t + \delta_{s} - \delta_{s'}\right\} + \frac{\sin\left(\frac{2\pi}{c} (v_{s} + v_{s'})l\right)}{\frac{2\pi}{c} (v_{s} + v_{s'})} \cos\left\{2\pi (v_{s} + v_{s'})t + \delta_{s} + \delta_{s'}\right\} \right].$$
(VII.3)

We see from this expression that E varies as a superposition of two sets of oscillations, one with frequencies of type  $2\pi(v_s - v_{s'})$  and the other with those of type  $2\pi(v_s + v_{s'})$ . From the condition (2) both  $v_s$  and  $v_{s'}$  are very large but their difference  $v_s - v_{s'}$  can become as small as c/2l. Therefore the first set contains oscillations of much smaller frequencies than the frequency v under consideration. These slow oscillations, their frequencies being the differences between the preexisting ones, are of the same nature as the well known "beating". We can thus call the variation due to the first set the "beat". The variation due to the second set, on the other hand, is extremely rapid.

The expression (VII.3) further tells us that the amplitudes of the first kind of oscillation are proportional to

$$\frac{\sin\left\{\frac{2\pi}{c}\left(\nu_{s}-\nu_{s'}\right)l\right\}}{\frac{2\pi}{c}\left(\nu_{s}-\nu_{s'}\right)}$$

while those of the second kind to

$$\frac{\sin\left\{\frac{2\pi}{c}\left(\nu_{s}+\nu_{s'}\right)l\right\}}{\frac{2\pi}{c}\left(\nu_{s}+\nu_{s'}\right)}$$

APPENDIX VII

and, hence, that the slow oscillations have much larger amplitudes than the rapid ones; the main part of the time variation of the energy E is therefore due to the interference effect between an oscillation of frequency  $v_s$  and those of the neighbouring frequencies  $v_{s'}$ . Taking this main part only, Eq. (VII.3) can be approximated, without appreciable error (thanks to condition (2)), by

$$E = \frac{1}{4}\sigma \sum_{\nu_{1} < \nu_{s} < \nu_{2}} \sum_{\nu_{1} < \nu_{s'} < \nu_{2}} (2\pi\nu_{s}) (2\pi\nu_{s'})A_{s}A_{s'}$$

$$- \frac{\sin\left\{\frac{2\pi}{c} (\nu_{s} - \nu_{s'})l\right\}}{\frac{2\pi}{c} (\nu_{s} - \nu_{s'})} \cos\left\{2\pi(\nu_{s} - \nu_{s'})l + \delta_{s} - \delta_{s'}\right\}. \quad (VII.4)$$

Using condition (4), it can further be simplified to

$$E = \frac{1}{4}\sigma (2\pi\nu)^2 A_{\nu}^2 \sum_{\nu_1 < \nu_s < \nu_2} \sum_{\nu_1 < \nu_s < \nu_2} \frac{\sin\left\{\frac{2\pi}{c}(\nu_s - \nu_{s'})l\right\}}{\frac{2\pi}{c}(\nu_s - \nu_{s'})} \cdot \cos\left\{2\pi(\nu_s - \nu_{s'})l + \delta_s - \delta_{s'}\right\}.$$
 (VII.5)

The time variation of E being thus obtained, the time average of this expression gives the average energy  $\langle E \rangle$  in the frequency range  $\nu$  and  $\nu + d\nu$  contained in l. After this averaging process, only the terms s = s' remain, while the others vanish:

$$\langle E \rangle = \frac{1}{4} \sigma (2\pi\nu)^2 A_{\nu}^2 l \sum_{\nu_1 < \nu_s < \nu_2} l$$
, (VII.6)

where  $\sum_{v_1 < v_2 < v_2} l$  is the number of proper oscillations of frequencies between  $v_1$  and  $v_2$ , that is between v and v + dv, and can be denoted by  $Z_L(v)dv$ . The subscript L on Z(v) is to distinguish it from the same kind of quantity for the separated part l which will be denoted by  $Z_l(v)$ . According to (III.16),

$$Z_{L}(\mathbf{v}) = \frac{2l}{c},$$

$$Z_{L}(\mathbf{v}) = \frac{2L}{c}.$$
(VII.7)

OSCILLATION OF A STRING

We are thus led to

$$\langle E \rangle = \frac{1}{4}\sigma \left(2\pi\nu\right)^2 A_{\nu}^2 l Z_L(\nu) \,\mathrm{d}\nu \,. \tag{VII.8}$$

Now, the fluctuation around the average,  $E - \langle E \rangle$ , is given by omitting in Eq. (VII.5) those terms with s = s' and its square becomes  $(AE)^2 = (E - \langle E \rangle)^2$ 

$$= \frac{1}{16} \sigma^{2} (2\pi\nu)^{4} A_{\nu}^{4} \sum_{\substack{s \neq s' \\ v_{1} < (v_{s}, v_{s'}) < v_{2} \\ v_{1} < (v_{r}, v_{r'}) < v_{2}}} \left[ \frac{\sin \left\{ \frac{2\pi}{c} (v_{s} - v_{s'}) l \right\}}{\frac{2\pi}{c} (v_{s} - v_{s'})} \cdot \frac{\sin \left\{ \frac{2\pi}{c} (v_{r} - v_{r'}) l \right\}}{\frac{2\pi}{c} (v_{r} - v_{r'})} \right] \\ \cdot \cos \left\{ 2\pi (v_{s} - v_{s'}) l + \delta_{s} - \delta_{s'} \right\} \cos \left\{ 2\pi (v_{r} - v_{r'}) l + \delta_{r} - \delta_{r'} \right\}.$$

In taking the average of this quantity, we do so not only with respect to time t but also with respect to the phase angles  $\delta_s$ ,  $\delta_{s'}$ ,  $\delta_v$  and  $\delta_v$ . since, in actual oscillation processes, particulary when the system is interacting with a thermal reservoir, these phase angles will change randomly through the influence of the boundary walls. We notice here that

$$\langle \cos \{ 2\pi (\mathbf{v}_{s} - \mathbf{v}_{s'})t + \delta_{s} - \delta_{s'} \} \cos \{ 2\pi (\mathbf{v}_{r} - \mathbf{v}_{r'})t + \delta_{r} - \delta_{r'} \} \rangle$$

$$= \frac{1}{2} \langle \cos \{ 2\pi (\mathbf{v}_{s} - \mathbf{v}_{s'} + \mathbf{v}_{r} - \mathbf{v}_{r'})t + \delta_{s} - \delta_{s'} + \delta_{r} - \delta_{r'} \} \rangle$$

$$+ \frac{1}{2} \langle \cos \{ 2\pi (\mathbf{v}_{s} - \mathbf{v}_{s'} - \mathbf{v}_{r} + \mathbf{v}_{r'})t + \delta_{s} - \delta_{s'} - \delta_{r} + \delta_{r'} \} \rangle,$$

and that, in order to have a non-vanishing average value, we should have either s = r' and s' = r simultaneously or s = r and s' = r' simultaneously. Thus we obtain

$$\langle \Delta E^{2} \rangle = \frac{1}{16} \sigma^{2} (2\pi v)^{c} A_{v}^{4} \sum_{s \neq s'} \left[ \frac{\sin \left\{ \frac{2\pi}{c} (v_{s} - v_{s'}) l \right\}}{\frac{2\pi}{c} (v_{s} - v_{s'})} \right]^{2}. \quad (VII.9)$$

The summation on the right hand side can, under the conditions (1) and (3), be replaced by an integration, giving rise to

$$\langle \Delta E^2 \rangle = \frac{1}{16} \sigma^2 (2\pi\nu)^4 A_{\nu}^4 Z_L^2(\nu) \int_{\nu_1}^{\nu_2} \int_{\nu_1}^{\nu_2} \left[ \frac{\sin\left\{\frac{2\pi}{c} (\nu - \nu')l\right\}}{\frac{2\pi}{c} (\nu - \nu')} \right]^2 d\nu d\nu'.$$
(VII.10)

A. VII]

APPENDIX VII

We now introduce a new variable  $\xi$  defined by

$$\frac{2\pi}{c}(\nu-\nu')l=\xi.$$

Then the integral

$$\int_{\nu_1}^{\nu_2} \left[ \frac{\sin\left\{\frac{2\pi}{c} \left(\nu - \nu'\right)l\right\}}{\frac{2\pi}{c} \left(\nu - \nu'\right)} \right]^2 d\nu'$$

becomes

$$\frac{cl}{2\pi}\int_{\xi_1}^{\xi_2}\left(\frac{\sin\xi}{\xi}\right)^2\mathrm{d}\xi\,,$$

where

$$\xi_2 = \frac{2\pi(\nu_2 - \nu)l}{c}$$
, and  $\xi_1 = \frac{2\pi(\nu_1 - \nu)l}{c}$ .

Since  $\nu$  varies from  $\nu_1$  to  $\nu_2$ ,  $\xi_2$  is positive while  $\xi_1$  is negative and they are both very large as compared to unity on account of condition (3), which says that  $d\nu = \nu_2 - \nu_1 \gg c/2l$ . The integration can then be carried out from  $-\infty$  to  $+\infty$  without introducing a noticeable error in the result. Thus with sufficient accuracy we get

$$\int_{\nu_1}^{\nu_2} \left[ \frac{\sin\left\{\frac{2\pi}{c}\left(\nu-\nu'\right)l\right\}}{\frac{2\pi}{c}\left(\nu-\nu'\right)} \right]^2 d\nu' = \frac{cl}{2\pi} \int_{-\infty}^{+\infty} \left(\frac{\sin\xi}{\xi}\right)^2 d\xi \qquad (\text{VII.11})$$
$$= \frac{4}{cl},$$

and accordingly,

$$\langle \Delta E^2 \rangle = \frac{1}{16} \sigma^2 (2\pi\nu)^4 A_{\nu}^4 Z_L(\nu)^2 \frac{1}{2} cl \int_{\nu_1}^{\nu_2} d\nu ,$$

which further is transformed to

$$\langle \Delta E^2 \rangle = \frac{1}{16} \sigma^2 (2\pi\nu)^4 A_{\nu}^4 l^2 Z_L(\nu)^2 Z_l(\nu)^{-1} d\nu$$
 (VII.12)

where Eq. (VII.7) and the fact that  $\nu_2 - \nu_1 = d\nu$  have been made use of. Comparing this expression with Eq. (VII.8), we arrive at our result

$$\langle \Delta E^2 \rangle = \frac{\langle E \rangle^2}{Z_l(\nu) \, \mathrm{d}\nu}.$$

A. VII, VIII]

In the case of actual cavity radiation, the calculation becomes much more involved since the problem is three dimensional but it can be done essentially in the same way to arrive at the same result.

### APPENDIX VIII

### Another Proof of T = dJ/dW

The method used in the text for the proof of T = dJ/dW was of

such an analytic nature that the train of reasoning involved might have been obscured. We shall in this section give another proof which should help clarify matters.

The trajectory of the system is assumed to be a closed curve as in Fig. (A.7) in which, besides the trajectory corresponding to the energy E, we add for later purposes, the trajectory corresponding to the energy  $E + \delta E$ .

The equations of motion,



Fig. A.7. The relation between the separation of equi-energy surfaces and the velocity of the trajectory.

$$\frac{\mathrm{d}q}{\mathrm{d}t} = \frac{\partial H}{\partial p}$$
 and  $\frac{\mathrm{d}p}{\mathrm{d}t} = -\frac{\partial H}{\partial q}$ , (VIII.1)

give for the velocity of the representative point (q, p) on the trajectory

$$v = \sqrt{\left(\frac{\mathrm{d}q}{\mathrm{d}t}\right)^2 + \left(\frac{\mathrm{d}p}{\mathrm{d}t}\right)^2}$$

$$= \sqrt{\left(\frac{\partial H}{\partial p}\right)^2 + \left(\frac{\partial H}{\partial q}\right)^2},$$
(VIII.2)

and hence the amount of time required for the representative point to complete a full orbit on the trajectory is given by

$$T = \oint \frac{\mathrm{d}s}{\sqrt{\left(\frac{\partial H}{\partial p}\right)^2 + \left(\frac{\partial H}{\partial q}\right)^2}}, \qquad (\text{VIII.3})$$

where ds is a line element of the trajectory and the integration runs over one full turn of the closed curve.

APPENDIX VIII

Let the distance between the trajectories of energies E and  $E + \delta E$ ,  $\delta E$  being very small, be denoted by  $\delta n$  as in Fig. (A.7). Then

$$\delta E = \sqrt{\left(\frac{\partial H}{\partial q}\right)^2 + \left(\frac{\partial H}{\partial p}\right)^2} \,\delta n \,, \qquad (\text{VIII.4})$$

$$\frac{1}{\sqrt{\left(\frac{\delta H}{\partial q}\right)^2 + \left(\frac{\delta H}{\partial p}\right)^2}} = \frac{\delta n}{\delta E}, \qquad (\text{VIII.5})$$

and the expression (VIII.3) becomes

$$T = \frac{1}{\delta E} \ \oint \delta n \ \mathrm{ds} \ . \tag{VIII.6}$$

But  $\delta n$ ds is the area of the shaded domain in Fig. (A.7) and hence the integral on the right-hand side of Eq. (VIII.6) gives the area of the ring shaped domain between the two trajectories. Recalling the definition of J(E), this area is equal to  $J(E + \delta E) - J(E)$ , so that we have

$$T = \frac{1}{\delta E} \left\{ J(E + \delta E) - J(E) \right\}.$$
(VIII.7)

In the limit,  $\delta E \rightarrow 0$ , this leads to the result

$$T = \frac{\mathrm{d}J}{\mathrm{d}E} \tag{VIII.7'}$$

which was to be proved.

## Exercises

### CHAPTER I

(1) Consider a diatomic molecule in which the two atoms are bound together by an elastic force. Prove that the classical specific heat of a gas consisting of such molecules is  $\frac{2}{2}R$ .

(2) Find Z(v) for a box of sides  $L_x$ ,  $L_y$  and  $L_z$ , containing radiation.

(3) In a given box with mirror walls there exists radiation propagating not isotropically but only in the direction normal to one of the surrounding walls. Prove that the pressure on this wall is given by P = U.

(4) Calculate the specific heat of a simple one-dimensional model crystal as in Appendix II. Observe the difference between the results with the Debye approximation on the one hand and with the use of spectrum (II.13) on the other.

### CHAPTER II

(1) When light hits a free electron, which is bound in neither an atom nor an external force field, the Compton effect can occur while the photoelectric effect cannot. Find the reason.

(2) Use the corpuscular nature of light in proving that, when a light ray is reflected back by a mirror which is normal to the direction of its propagation the pressure exerted on this mirror is given by P = U.

(3) A mirror is placed facing a light source of 1 candle at 3 meter distance. Calculate the pressure on the mirror.

(4) Discuss the Compton effect produced by a photon incident upon an electron of initial velocity v.

(5) Discuss the possibility of explaining the Doppler effect on the basis of the corpuscular nature of light.

### CHAPTER III

(1) One used to consider a simple model of an atom in which an electron made simple harmonic oscillations under the influence of an elastic-type force. Discuss the motion in detail for the cases where

#### EXERCISES

such a dynamical system is placed in a uniform electric or a uniform magnetic field.

(2) A stream of particles is incident upon an elastic sphere of radius a. Calculate the angular distribution of the scattered particles as well as the scattering cross section. Show that the total cross section is equal to  $\pi a^2$  when the size of the particle is negligible and is equal to  $\pi (a + r)^2$  when the incident particle is a sphere of radius r.

(3) Show that the coordinates of the electron in a hydrogen atom as a function of time can be expressed as a Fourier series with a single fundamental frequency. Show further that, when a uniform magnetic field is switched on, this becomes a double Fourier series with two fundamental frequencies one of which is the Larmor frequency.

(4) Draw a figure corresponding to Fig. 28 of § 18 for the case n = 10 and observe that for large *n*'s the classical frequencies agree well with the quantum-mechanical ones.

(5) Show that the value of  $\oint p dq$  is always the same for any choice of canonical coordinate q.

(6) Treat the simple model of the atom of (1) quantum-theoretically, first with ordinary rectangular coordinates and then with the polar coordinates. Observe that the quantum-theoretically allowed motions differ in the different coordinate systems considered. This is a paradox similar to that of azimuthal quantization.

(7) Find the energy levels for an electron in the force potential  $-e/r + \alpha^2/r^2$ , instead of the ordinary potential -e/r, and show that they depend also on the subsidiary quantum number k.

(8) Show that the electron motion of the last problem is, in contrast to the case of a Coulomb force, described by a double Fourier series with two fundamental frequencies.

(9) Show that the selection rules of  $\S 28$  (iv) actually hold for the hydrogen atom, the atomic model of problem (1), and that of problem (7).

(10) Use Boltzmann's principle for calculating the specific heat of a diatomic molecule considered as a rotator and show that the specific heat-temperature relationship of Fig. 2 actually results.

(11) For a quantum-theoretical gas contained in a cubic box of volume 1 cm<sup>3</sup>, find the temperature at which the specific heat starts deviating from the classical value due to the quantization of translational motion.
#### EXERCISES

(12) Find the temperature at which the motion of the atomic electrons starts contributing to the specific heat.

## CHAPTER IV

(1) Investigate the Rydberg correction for the atomic model of problem (7) of the last chapter. The constant  $\alpha$  is assumed to be very small.

(2) The X-rays of the continuous spectrum from an X-ray tube are due to the stopping of the cathode-ray electrons by the anticathode. As the applied voltage is increased, the spectrum shifts to shorter wave lengths. Discuss the physical reason behind this fact and find the relationship between the wave length and the applied voltage. Calculate the shortest possible wave length in the continuous X-ray spectrum from a 1000 volt X-ray tube.

(3) The size of an atom is known experimentally to vary with the same period as seen in the periodic table. Find the explanation and predict where in the periodic table it shows maxima and minima.

(4) Use Pauli's principle in explaining why the ground states of atoms with only closed shells, i.e., inert gases, have zero angular momenta.

(5) Due to Pauli's principle the so-called electron gas exhibits many differences from ordinary gases and from a photon gas. As an example, investigate the specific heat-temperature and the pressure-temperature relationships for this gas. Find how they differ from those of an ordinary gas and a photon gas.

#### CHAPTER V

(1) For two given matrices A and B, prove that

$$(AB)^{\dagger} = B^{\dagger} A^{\dagger}$$

$$(AB)^{-1} = B^{-1}A^{-1}$$

(2) Is the statement correct that the product of two Hermitian matrices is also Hermitian? If the statement is not generally correct find the condition under which it is so.

(3) Is the product of two unitary matrices also a unitary matrix?

and

#### EXERCISES

(4) When two matrices X and P satisfy  $PX - XP = (h/2\pi i) 1$ , prove that the eigenvalues of H(X, P), a function of X and P, are identical to those of  $H(U^{\dagger}XU, U^{\dagger}PU)$  where U is an arbitrary unitary matrix. Finite dimensionality may be assumed since the proof for the infinite dimensional case is very difficult. This problem proves the result in the text that the eigenvalues of  $H^{\circ}$  are independent of the choice of  $X^{\circ}$  and  $P^{\circ}$ .

(5) Find the unitary matrix which transforms  $\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$  into a diagonal matrix. What is the transformed diagonal matrix?

#### INDEX

## A

Action variable, 118 Adiabatic invariant, 21 — of oscillating system, 294 Generalized —, 110 Atom Bohr's theory of the —, 97 Nagaoka model of the —, 82 Periodicity of the —, 180 Rutherford model of the —, 90 Shell structure of the —, 167, 177 Structure of the —, 78 Thomson model of the —, 82 Azimuthal quantization, 130

## B

Balmer formula, 94 Balmer series, 96 Bergman series, 168 Black body radiation, 2, 19 Boltzmann constant, 5 Boltzmann's principle, 5, 262 Bohr magneton, 107 Bohr radius, 106 Bohr's correspondence principle, 103, 139 Bohr's frequency condition, 99, 235 Brackett series, 96

# С

Canonical equation of motion, 235 Cavity, 2, 43 Closed shell, 177, 180 Cloud chamber, 70 Characteristic X-ray, 174 Combination law, 96 Commutation relation, 232 Compton effect, 64 Compton wave length, 69 Compton-Simon experiment, 71 Constant of motion, 121, 263 Continuous X-ray, 173 Core electron, 169 Corpuscular nature of light, 43, 75 Correspondence principle, 142 Cross section, 88

#### D

Debye function, 41 Debye temperature, 41 Degenerate level, 130 Dipole moment, 140 D-line, 171 D-orbit, 129 Doublet, 190 D-term, 129, 170 Dulong-Petit's law, 10

#### Е

Ehrenfest's adiabatic hypothesis, 107 Eigensolution, 259 Normalization of -, 255 Orthogonality of ---, 255 Eigenvalue, 259 — problem, 244 Einstein's relationship, 74 Energy level, 100 - of hydrogen, 100 Doublet ---, 190 Singlet —, 190 Triplet -, 190 Energy quantum, l Equi-energy surface, 263 Equipartition law, 5 Ergodic system, 264

## F

Fluctuation — of the energy, 45, 298 — with Planck's formula, 55 — with the Rayleigh-Jeans formula, 49 — with Wien's formula, 50 Franck-Hertz experiment, 136 F-term, 170 Fundamental frequency, 93

## H

Hamiltonian, 112, 236, 290 — function, 111, 262 Hamiltonian's equation of motion, 113, 290 Harmonic oscillator, 150 Heisenberg's matrix mechanics, 204 Hermitian matrix, 227 Hermitian conjugate, 249

# I

Inner quantum number, 189 Intensity of light, 141

## K

K-shell, 180

## L

Larmor angular velocity, 81 Larmor precession, 81 Law of equipartition of energy, 5 Lenard's experiment, 57 Level number, 190 Light quantum, 43 Liouville's theorem, 266 L-shell, 180 Lyman series, 96

## М

Magnetic moment Intrinsic —, 192 Orbital —, 197 Magnetic quantum number, 132 Magneton, 107 Matrix, 225 — element, 225 Associative law of --, 229, 230 Commutative law of -, 229, 230 Diagonal —, 231 Distribution law of -, 229, 230 Hermitian -, 227 Hermitian conjugate of —, 249 Inverse —, 249 Multiplication of -, 229 Unit ---, 231 Unitary -, 248 Zero ---, 231 Michelson's interferometer, 76 Millikan's experiment, 60 Momentum of light quantum, 64

Moscley's relation, 174 M-shell, 180 Multi-periodic motion, 93

## N

Normal coordinate, 9 Normal mode — of crystal, 280 — of string, 16, 283 — of electromagnetic field, 13, 283

## 0

Old quantum mechanics, 78 Optical electron, 169 Optical spectra, 167

## Р

Pauli's principle, 200 Pendulum, 21 Periodic table, 181 Phase, 262 — space, 263 Photo-chemical reaction, 62 Photoelectric effect, 57 Photo-electron, 71 Photon gas, 72 Planck's constant, 33 Planck's formula, 32 Planck's oscillator, 216 Polarization of light Circular —, 142 Linear ---, 141 P-orbit, 129 Pressure - of radiation, 24 — of string, 290 Principal quantum number, 128 Principal series, 168 Proper oscillations, 9 - of crystal, 280 — of electromagnetic field, 13, 286 — of string, 16, 283 P-term, 129, 170

# Q

Quantization, 107 — of angular momentum, 124 — of rotational motion, 163 — of translational motion, 164

## 312

#### INDEX

Quantum — of action, 33 — of energy, 34 Quantum condition, 107 — for multi-periodic system, 119 — for one-dimensional system, 118 Quantum jump, 99 Quantum number, 107 Inner —, 189 Magnetic —, 132 Principal —, 128 Subsidiary —, 128 Quantum theory — of specific heat, 39 — of X-ray scattering, 66

## R

Rayleigh-Jeans formula, 14 Recoil electron, 69 Ritz's combination law, 96 Rutherford's formula, 84 Rydberg's constant, 95 Rydberg's correction, 168 Rydberg's formula, 95

## S

Selection rule, 151 Shell structure, 177 Singlet, 190 Sommerfeld-Landé theory, 193 S-orbit, 129 Specific heat, 3 — of crystalline substance, 9 - of diatomic molecules, 8 — of monatomic molecules, 6 - of vacuum, 12 Spectral series Balmer —, 96 Bergman —, 168 Brackett —, 96 Lyman —, 96 Paschen —, 96 Principal —, 168 Subordinate —, 168 Spectral terms, 96, 175 D —, 129 Multiplet structure of —, 189 P ---, 129 S —, 129

Spectrum of atoms, 94 Spin of electron, 192 Stationary state, 100 — of hydrogen atom, 124 Stefan-Boltzmann law, 26 Stefan's constant, 33 Stefan's law, 14 S-term, 129, 170 Stern-Gerlach experiment, 138 Subordinate series, 168 Subsidiary quantum number, 128

## T

Term, 96, 175 Thermal reservoir, 274, 278 Trajectory of motion, 111, 263 Transition, 99 -- component, 206 -- probability, 147

## U

Unitary matrix, 248 Unitary transformation, 248

## W

Wave nature, 75 Wien's displacement law, 20, 29 Wien's formula, 31

## Х

X-ray — spectrum, 173 Characteristic ---, 174 Continuous --, 173 K, L, M, N-groups of ---, 174 X-ray scattering Quantum theory of ---, 66 Thomson's theory of ---, 64

## Y

Young's experiment, 75

## Z

Zeeman effect, 78 Anomalous —, 198 Normal —, 198