THE MODERN THEORY OF SOLIDS

BY

FREDERICK SEITZ, PH.D.
Assistant Professor of Physics
University of Pennsylvania

FIRST EDITION

1940

INTERNATIONAL SERIES IN PHYSIC

LEE A. DuBRIDGE, CONSULTING EDITOR

THE MODERN THEORY OF SOLIDS

Periodic Chart of the Elements

(The integers are the atomic numbers of the elements)

Common valence											
Electropositive	1	2	3	4							
Electronegative				4	3	2	1	0			
	1 H							2 He			
	3 Li	4 Be	5 B	6 C	7 N	8 O	9 F	10 Ne			
	11 Na	12 Mg	13 Al	14 Si	15 P	16 S	17 Cl	18 A			
First long period	19 K	20 Ca	21 Sc	22 Ti	23 V	24 Cr	25 Mn	26 Fe	27 Co	28 [illegible]	
	29 Cu	30 Zn	31 Ga	32 Ge	33 As	34 Se	35 Br	36 Kr			
Second long period	37 Rb	38 Sr	39 Yt.	40 Zr	41 Cb	42 Mo	43 Ma	44 Ru	45 Rh	46 Pd	
	47 Ag	48 Cd	49 In	50 Sn	51 Sb	52 Te	53 I	54 Xe			
Third long period	55 Cs	56 Ba	Rare earths	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	86 Rd			
	87	88 Ra	89 Ac	90 Th	91	92 U					

The Fundamental Constants

(In cgs units)

Electronic charge	$e = 4.80 \cdot 10^{-10}$	$\hbar = h/2\pi$	$1.05 \cdot 10^{-27}$
Planck's constant	$h = 6.60 \cdot 10^{-27}$	e^2	$= 2.30 \cdot 10^{-19}$
Avogadro's number	$N_A = 6.03 \cdot 10^{23}$	$\hbar^2$	$= 1.10 \cdot 10^{-54}$
Electronic mass	$m = 9.03 \cdot 10^{-28}$	$a_1 = \hbar^2/me^2$	$= 0.531 \cdot 10^{-8}$ cm
Proton mass	$M = 1.67 \cdot 10^{-24}$	$\beta = e\hbar/2mc$	$= 9.30 \cdot 10^{-21}$
Velocity of light	$c = 3.00 \cdot 10^{10}$	e^2/a_1	$= 4.33 \cdot 10^{-11}$ erg
Boltzmann's constant	$k = 1.39 \cdot 10^{-16}$	$e/300$	$= 1.60 \cdot 10^{-12}$

Conversion Factors

1 ev $= 1.60 \cdot 10^{-12}$ erg	1 erg $= 6.25 \cdot 10^{11}$ ev
1 ev/molecule $= 23.05$ kg cal/mol	1 kg cal/mol $= 0.04337$ ev/molecule
1 kg cal $= 4.185 \cdot 10^{10}$ ergs	1 erg $= 2.389 \cdot 10^{-9}$ kg cal.
1 ev $= 11,500\ k$	$k = 8.69 \cdot 10^{-5}$ ev/deg

PREFACE

The present volume was written with a desire to satisfy the requirements of three types of possible reader: First, of course, students of physics and chemistry who desire to learn some details of a particular branch of physics that has general use; second, experimental physicists and chemists, and engineers and metallurgists with mathematical leanings who are interested in keeping an eye on a field of physics that is of possible value to them; and third, theoretical physicists at various stages of development who are interested in the present status of that phase of solid bodies that deals with electronic structure. The author fully realizes that the first two groups of reader do not wish to be concerned with all the intricacies of the theory, and for this reason he has attempted to edit the text by marking the more mathematical sections with an asterisk. It is recommended that readers not desiring to go through all this material read over it with an eye for the qualitative arguments and conclusions.

The author believes that an investigation of the table of contents will tell more concerning the scope of the book than a paragraph or two at this place. Although the book is a large one, it must be admitted that a number of very important topics are not treated. For example, the plastic properties of solids are only touched upon. The reason for this omission is, of course, that the theory involved in this field is not one that grows naturally out of modern quantum theory, and hence might better be treated under separate cover with some of the other structure-sensitive properties of solids. In addition, it was felt necessary to curtail the discussion of many interesting topics simply to avoid making the book much too long. In all such cases, which usually involve rather specialised subjects, an attempt has been made to give the reader reference material from which he may draw further information.

The author started to write this book in 1936 when he was at the University of Rochester and gratefully acknowledges the cooperation he received from Professor L. A. DuBridge in connection with it. The book was continued in spare time during the author's stay at the Research Laboratories of the General Electric Company in Schenectady from 1937 to 1939. He would like to say that the atmosphere of this organization proved very stimulating for writing, as well as for many other forms of research, and would like to express his gratitude to the directors for their

interest and support. The book was completed at the University of Pennsylvania with the encouragement of Professor G. P. Harnwell. In addition, the author is obligated to the directors of the Westinghouse Research Laboratories for the privilege of spending a stimulating summer in East Pittsburgh in 1939.

The author also wishes to express his gratitude to the many friends and colleagues to whom he has turned for advice and discussions. Among these he particularly desires to mention Dr. R. P. Johnson, Dr. W. Shockley, Professor L. N. Ridenour, Dr. S. Dushman, Dr. E. U. Condon, Professor E. P. Wigner, and Professor J. H. Van Vleck.

Finally, he should like to acknowledge, though in a very inadequate manner, the constant help and encouragement furnished by his wife, Elizabeth Marshall Seitz, without whose aid this book probably never would have been written.

FREDERICK SEITZ.

PHILADELPHIA,
June, 1940.

CONTENTS

(NOTE: *Sections marked with an asterisk are more mathematical in nature, and may be read only for their qualitative content in a first reading.*)

COMMONLY USED SYMBOLS

a_h	The Bohr radius of the hydrogen atom.
A	The magnetic vector potential.
α_V	The coefficient of volume expansion.
β	The volume compressibility; also, the Bohr magneton.
c	The velocity of light; also, the velocity of elastic waves.
c_v	The specific heat at constant volume.
C_V	The molar or atomic heat at constant volume.
C_P	The molar heat at constant pressure.
γ	The damping frequency of an oscillator.
e	The electronic charge; also, the base of the natural system of logarithms.
E	The electrostatic field intensity.
$\epsilon(\mathbf{k})$	The energy of an electron of wave number $\mathbf{k}$.
ϵ	The dielectric constant.
ϵ'	The energy of the uppermost electron in the filled band.
ϵ_0'	The energy of the uppermost electron in the filled band at absolute zero of temperature.
ϵ_0	The energy of the electron at the bottom of the filled band.
$\bar{\epsilon}$	The mean electronic energy.
$\bar{\epsilon}_0$	The mean electronic energy at absolute zero of temperature.
δ	The delta function.
Δ	The Laplacian operator.
f_E	The Einstein specific-heat function.
f_D	The Debye specific-heat function.
f	A partition function.
$g(r_s)$	The free-electron correlation energy.
h	Planck's constant.
$\hbar$	Planck's constant divided by 2π.
H	A Hamiltonian operator.
H	The magnetic field intensity.
J, I	The current per unit area.
$\mathbf{k}$	The electronic wave-number vector; components k_x, k_y, k_z.
k	Boltzmann's constant; also the optical extinction coefficient.
k_0	The wave number of the electron at the top of the filled band.
$\mathbf{K}$, $\mathbf{K}_\alpha$	Principal vectors in the reciprocal lattice of a crystal.

l	The mean free path between collisions (averaged for all scattering angles).
$l(\mathbf{k})$	The mean free path of the electron having wave number $\mathbf{k}$.
$\mathbf{L}$	Orbital angular momentum vector.
L	The orbital angular momentum quantum number.
λ	Wave length of an electron, light quantum, or lattice vibrational wave.
m	The actual electronic mass.
m^*	The effective electronic mass [determined from the $\epsilon(\mathbf{k})$ curve].
M	The mass of an atom or ion.
$\mathbf{M}$	Orbital magnetic moment.
$\mathbf{M}_d$	Atomic electric dipole moment.
$\mathbf{M}_{ij}$	Matrix component of the atomic dipole moment.
M	Magnetic moment per unit volume.
M_s	The saturation magnetic moment.
$\mathfrak{u}$	The electron spin magnetic moment vector.
μ	The magnetic permeability.
n	Index of refraction; exponent in Born's ion-ion repulsion function.
n_e	The number of particles per unit volume.
n_f	The effective number of free electrons (usually the number of electrons in the band of width kT at the top of the filled region).
N	The total number of atoms or electrons in a system; the complex index of refraction.
N_A	Avogadro's number.
ν	Frequency of vibration.
ν_m	Maximum frequency in elastic vibrational spectrum of a lattice.
P	The electric polarization.
φ	The work function of a metal; also, the azimuthal angle.
r_s	The radius of the sphere having the same volume as the atomic cell.
r_{12}	The distance between two particles.
R	The gas constant; also, the Hall constant; also, the reflection coefficient; also, the refractivity.
ρ	Resistivity; also, the radiation density; also, the charge density.
S	The total spin quantum number.
σ	The electrical conductivity; also, the absolute value of the wave-number vector.
σ_s	The saturation magnetization in units of Bohr magnetons per atom.
$\mathfrak{s}$	The wave-number vector for lattice vibrations; also, the spin angular momentum vector; the components are, σ_x, σ_y, σ_z in both cases.
σ_T	The Thomson heat.

T	The temperature, usually in degrees Kelvin.
Θ	A characteristic temperature, such as that for the specific heat or the electrical conductivity; also, the effective cross section for collisions between electrons and lattice.
Θ_D	The Debye characteristic temperature.
Θ_C	The Curie temperature.
τ	The mean time between collisions.
$\mathbf{a}_i$	The primitive translation vectors of the lattice.
$v(\epsilon_0')$	The velocity of the uppermost electron in the filled electron band.
V	The total volume of a system.
χ	The magnetic susceptibility (per unit volume).
χ_m	The molar magnetic susceptibility.
Z	The atomic number.

THE MODERN THEORY OF SOLIDS

CHAPTER I

EMPIRICAL CLASSIFICATION OF SOLID TYPES

The Five Solid Types.—When using the term "solid" in this book, we shall refer to crystalline aggregates of atoms and molecules; that is, we shall have little to do with substances such as glasses that do not have definite lattice structure. In addition, we usually shall deal with crystals having relatively simple structures because they are most amenable to theoretical treatment. This limitation is not very important so long as we are interested only in general properties of solids, for substances with complicated structures can be classified in the same general way as simple ones. On the other hand, the restriction is very serious if the theory is looked upon as a tool for aid in making practical use of solids. There seems to be no way of removing this restriction other than to continue work along the lines that are developed here.

Although there is no unique way of classifying all the solids found in nature, the division that will be used here has enough natural advantages to make a discussion of alternatives unnecessary. It is based upon a survey of chemical, thermal, electrical, and magnetic characteristics. Briefly, the classification is as follows:

a. Metals.
b. Ionic crystals.
c. Valence crystals.
d. Semi-conductors.
e. Molecular crystals.

Metals, which are distinguished primarily by their good electrical and thermal conductivity, are formed by the combination of the atoms of electropositive elements.

Ionic crystals are distinguished by good ionic conductivity at high temperatures, strong infrared absorption spectra, and good cleavage. They are formed by a combination of highly electropositive and highly electronegative elements, the salts, sodium chloride, magnesium oxide, etc., being the best examples.

Valence crystals, of which diamond and carborundum are examples, have poor electronic and ionic conductivity, great hardness, and poor cleavage. They are formed by combination of the lighter elements in the middle columns of the periodic chart.

Semi-conductors, of which zinc oxide and cuprous oxide are good examples, show a feeble electronic conductivity which increases with increasing temperature. It should be added that there is evidence that these substances are electronic conductors only when impure or when their composition is slightly different from that characteristic of ideal stoichiometric proportions, such as when there is an excess of zinc in zinc oxide. For this reason, semi-conductors are characterized by a tendency to favor addition of impurities and to disobey simple valence rules.

Finally, molecular crystals are the solids formed by inactive atoms such as the rare gases, and saturated molecules such as hydrogen and methane. They are characterized by low melting and boiling points, and they generally evaporate in the form of stable molecules.

As will be shown below, a large number of solids have properties that overlap those of two or more of these ideal groups. For this reason, the divisions should not be regarded as being clean-cut in the sense that a given solid belongs to only one of them.

We shall now give a more detailed discussion[1] of each of the five solid types.

2. Monatomic Metals.—Metals may be divided conveniently into two major classes, namely, monatomic metals and alloys. The literature relating to alloys naturally is much larger than that for monatomic metals. Since we shall not be interested in developing the theory of alloys beyond an elementary stage, we shall not give them a proportional amount of space.

We may recognize a further subdivision of metals into two groups depending upon whether the *d* shells[2] of the constituent atoms are filled or not. If the *d* shells are completely filled or completely empty, the properties of the metal usually are simpler than if they are not. The two cases will be discussed separately, the designation "simple metals" being used if the *d* shells are completely filled or completely empty, and "transition metals" in the alternative case.

[1] Since the methods used in obtaining most of this experimental material can be found in other places, they usually will not be discussed here; however, a few lesser known experiments concerning semi-conductors and ionic crystals are discussed in Secs. 4 and 6.

[2] Throughout this book, we shall use the conventional notation for the electronic orbital momentum quantum numbers (*cf.*, for example, H. E. White, *Introduction to Atomic Spectra* (McGraw-Hill Book Company, Inc., New York, 1934). In this notation the letters *s*, *p*, *d*, *f*, *g*, etc., designate the states having orbital angular momentum quantum numbers 0, 1, 2, 3, 4, etc., respectively.

a. Cohesion of Monatomic Metals.—The heat of sublimation, which is the energy required to dissociate a mol of substance into free atoms, is a convenient measure of the cohesion of a metal. Numerical values of the heats of sublimation that have been taken from the compilation by Bichowsky and Rossini[1] are given in Table I. The few values that appear in parenthesis do not occur in these tables and have been estimated by use of Trouton's rule, namely,

$$L = 0.0235T,$$

TABLE I.—THE HEATS OF SUBLIMATION OF MONATOMIC METALS
(In kg cal/mol at room temperature)

Monovalent Metals					
Li	39.0			Cu	81.2
Na	25.9			Ag	68.0
K	19.8			Au	92.0
Rb	18.9				
Cs	18.8				
Divalent Metals					
Be	75			Zn	27.4
Mg	26.3			Cd	26.8
Ca	47.8			Hg	14.6
Sr	47				
Ba	49				
Ra	(72.7)				
Trivalent Metals					
Al	55			Ga	52
Sc	70			In	32
Yt	90			Tl	40
La	90				
Tetravalent Metals					
Ti	100			Ge	85
Zr	110			Sn	78
Hf	(>72)			Pb	47.5
Th	177				
Pentavalent Metals					
As	30.3				
Sb	40.				
Bi	47.8				
Transition Element Metals					
V	85	Nb	(>68)	Ta	(>97)
Cr	88	Mo	160	W	210
Mn	74	Ma		Re	
Fe	94.0	Ru	120	Os	125
Co	85.0	Rh	115	Ir	120
Ni	85.0	Pd	110	Pt	127
		U	220		

[1] F. R. BICHOWSKY and F. D. ROSSINI, *The Thermochemistry of the Chemical Substances*, Reinhold Publishing Corporation, New York, 1936.

where L is the heat of sublimation in kilogram calories per mol at the boiling point and T_b is the boiling temperature in degrees Kelvin. There are many interesting relationships among these values. One of the most

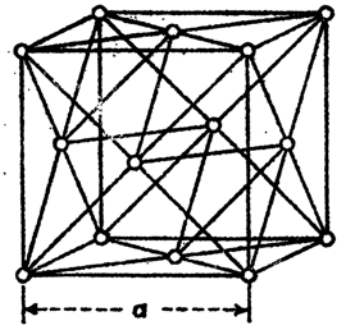

FIG. 1.—The face-centered cubic lattice.

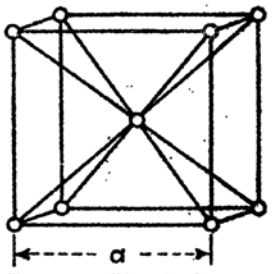

FIG. 2.—The body-centered cubic lattice.

striking ones is the fact that the atoms of transition metals on the whole are more tightly bound together than those of simple metals.

b. Crystal Structures.—Most of the monatomic metals crystallize in simple cubic or hexagonal structures. The three common types are shown in Figs. 1 to 3. More complex structures, which occur mainly among the atoms having higher valence, are shown in Figs. 4 to 11. Table II is a tabulation of the crystal parameter values for different

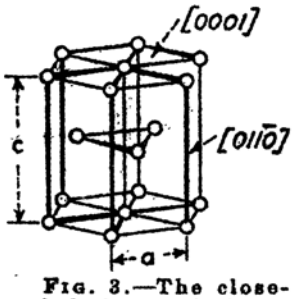

FIG. 3.—The close-packed hexagonal lattice showing two prominent crystallographic planes.

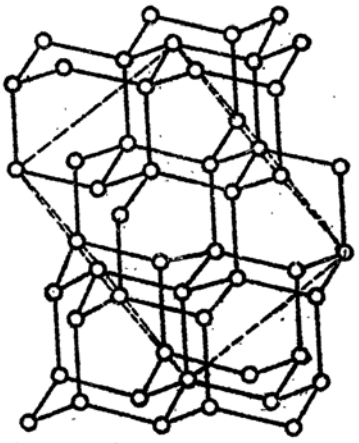

FIG. 4.—The diamond and gray tin lattice.

metals, including various allotropic modifications. These values have been taken from the three editions of *Strukturberichte.*[1]

[1] *Strukturberichte*, Leipzig (1931). Three supplements have appeared since the first volume.

Hume-Rothery[1] has pointed out that many complex structures, such as those of bismuth (Fig. 9), tin (Figs. 4 and 8), mercury (Fig. 6), and gallium (Fig. 7), are strikingly like those met among valence crystals, which are discussed below. For this reason, he would regard these substances as being intermediate between ideal metal and valence types, as we shall see below. This view is supported by observations on the conductivity and magnetic properties of these metals.

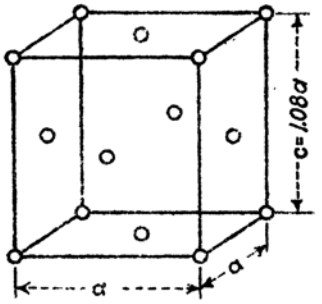

FIG. 5.—The indium lattice. The parameter values are given for indium.

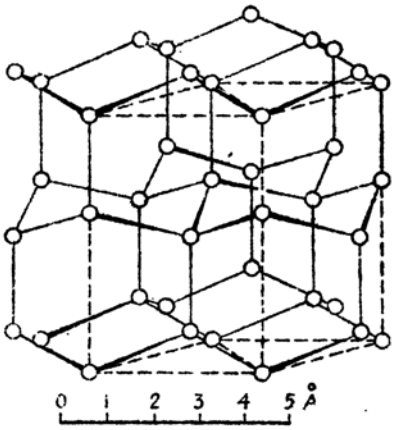

FIG. 6.—The mercury lattice.

c. Allotropy.—Table II shows that at ordinary pressures the transition metals exhibit allotropy more commonly than do simple metals. Bridgman[2] has found, however, that many of the simpler metals change their structure at high pressures. Cesium, for example, has a close-packed modification which appears at 22,000 kg/cm² of pressure; similar changes occur in magnesium. For this reason, it is doubtful whether polymorphism is a particular characteristic of any one group of metals.

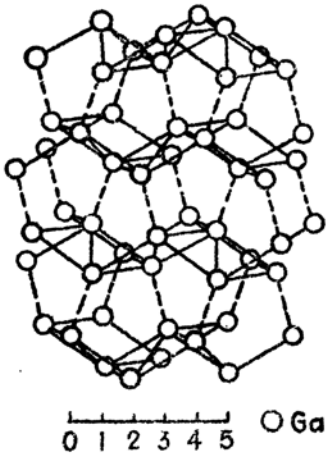

FIG. 7.—The gallium lattice, showing the layer structure in which each atom is surrounded by three neighbors.

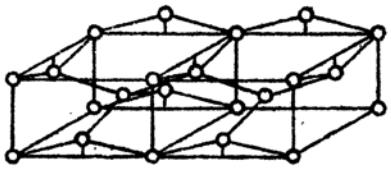

FIG. 8.—The white tin lattice.

[1] W. HUME-ROTHERY, *The Structure of Metals and Alloys* (Institute of Metals Monograph, London, 1936).

[2] P. W. BRIDGMAN, *Phys. Rev.*, **48**, 893 (1935); *Nat. Acad. Sci. Proc.*, **23**, 202 (1937).

TABLE II.—TABULATION OF CRYSTAL STRUCTURES AND PARAMETERS OF MONATOMIC METALS

(*d* is the nearest interatomic distance in angstroms and *a* is the edge length of the fundamental cube for cubic lattices. The parameters *a* and *c* for hexagonal close-packed lattices are shown in Fig. 3.)

The *alkali metals* form body-centered cubic lattices with the following parameter values.

	a	*d*
Li	3.46	3.09
Na	4.24	3.67
K	5.25	4.54
Rb	5.62	4.87
Cs	6.05	5.24

The *monovalent noble metals* have face-centered cubic lattices with the following parameter values.

	a	*d*
Cu	3.609	2.55
Ag	4.078	2.88
Au	4.070	2.87

Divalent Metals

	Type	*a*	*c*	*d*
Be	h.c.p.	2.28	3.59	*a*
Mg	h.c.p.	3.20	5.20	*a*
α Ca	f.c.c.	5.56	....	3.93
β Ca	h.c.p.	3.98	6.52	*a*
Sr	f.c.c.	6.06	....	4.28
Ba	b.c.c.	5.01	....	4.34
Zn	h.c.p.	2.65	4.930	*a*
Cd	h.c.p.	2.97	5.61	*a*
Hg	(see Fig. 6)			

Trivalent Metals

	Type	*a*	*c*	*d*
Al	f.c.c.	4.04	....	2.86
Sc				
Yt	h.c.p.	3.66	5.81	*a*
α La	h.c.p.	3.72	6.06	*a*
β La	f.c.c.	5.30	....	3.74
Ga	(see Fig. 7)			
In	tet.f.c.	4.59	4.94	(see Fig. 5)
α Tl	h.c.p.	3.45	5.52	*a*
β Tl	f.c.c.	4.84	....	3.42

TABLE II.—TABULATION OF CRYSTAL STRUCTURES AND PARAMETERS OF MONATOMIC METALS.—(*Continued*)

Tetravalent Metals

	Type	*a*	*c*	*d*
Ti	h.c.p.	2.953	4.73	2.892
α Zr	h.c.p.	3.23	5.14	*a*
β Zr	b.c.c.	3.61	....	3.13
Hf	h.c.p.	3.20	....	3.14
Ge	diam. str.	5.62	....	2.43
α Sn (gray)	diam. str.	6.46	....	2.80
β Sn (white)	(see Fig. 8)			
Pb	f.c.c.	4.93	....	3.48

Pentavalent Metals

As (see Fig. 9)
Sb As type; (see Fig. 9)
Bi As type; (see Fig. 9)

Transition Metals

	Type	*a*	*c*	*d*
V	b.c.c.	3.01	....	2.61
α Cr	b.c.c.	2.87	....	2.49
β Cr	h.c.p.	2.72	4.42	*a*
α Mn	(see Fig. 11)			
β Mn	.....	12.58		
γ Mn	tet.f.c.	3.77	3.53	2.08
α, β, δ Fe (α is low-temperature magnetic form)	b.c.c.	2.86	....	2.58
γ Fe	f.c.c.	3.56	....	2.57
α Co	h.c.p.	2.51	4.11	*a*
β Co	f.c.c.	3.55	....	2.51
Ni	f.c.c.	3.51	....	2.48
Nb	b.c.c.	3.30	....	2.86
Mo	b.c.c.	3.14	....	2.96
Ma				
Ru	h.c.p.	2.765	4.470	
Rh	f.c.c.	3.78	....	2.67
Pd	f.c.c.	3.88	....	2.74
Ta	b.c.c.	3.29	....	2.72
α W	b.c.c.	3.16	....	2.73
β W	(see Fig. 11)			
Re	h.c.p.	2.76	4.45	*a*
Os	h.c.p.	2.71	4.32	*a*
Ir	f.c.c.	3.83	....	2.71
Pt	f.c.c.	3.92	....	2.71

There is, however, one type of allotropic change that is characteristic of transition metals. This may be illustrated by comparing the cases of tin and iron. The α, or gray, modification of tin is stable at very low temperatures, whereas the β, or white, modification is stable at high

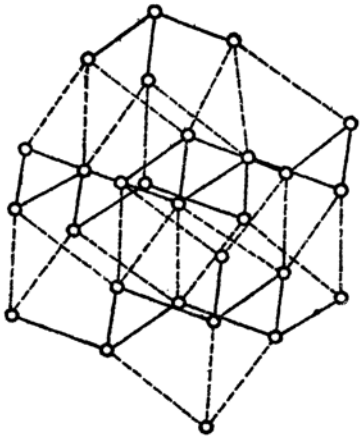

FIG. 9.—The bismuth lattice, showing the layer structure in which each atom is surrounded by three nearest neighbors.

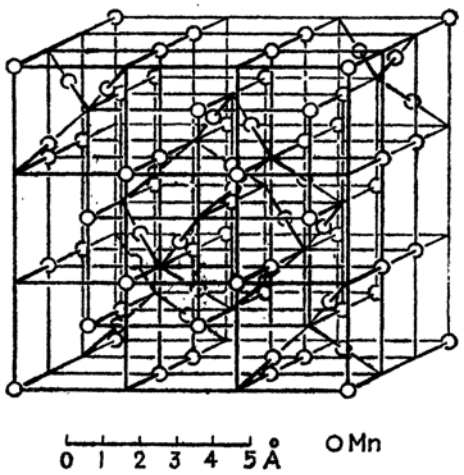

FIG. 10.—The α manganese lattice.

temperatures (*cf.* Table II). The transition temperature, which is 18°C, was determined by Cohen and van Eijk[1] from measurements on the emf of an electrolytic cell that had one gray-tin electrode and one white-tin electrode. The case of tin is typical of the simple metals inasmuch as the α modification is not again stable in any range from 18°C to the melting point of the β phase. On the other hand, the body-centered modification of iron is stable[2] in two temperature ranges, namely, from 0° to 1179°K and from 1674°K to the melting point 1803°K. The face-centered, or γ, modification is stable in the intermediate range from 1179° to 1674°K. This "intrusion" of one phase into the range of another also occurs in cobalt.[3] In this case a face-centered cubic phase splits the stable range of a close-packed hexagonal phase into two parts.

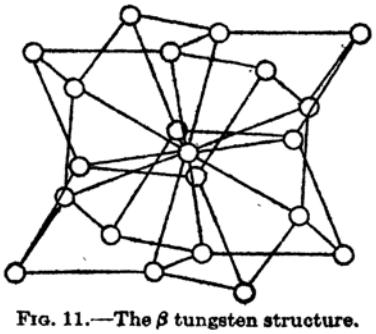

FIG. 11.—The β tungsten structure.

[1] E. COHEN and C. VAN EIJK, *Z. physik. Chem.*, **30**, 601 (1899).

[2] *Cf. Strukturberichte.*

[3] S. B. HENDRICKS, M. E. JEFFERSON, and J. F. SHULTZ, *Z. Krist.*, **73**, 376 (1930).

d. Atomic Radii.—It is often convenient to ascribe to each atom or ion a radius that is determined by the volume which the atom or ion occupies in a given compound. In monatomic metals the radius r is defined as half the distance between the centers of nearest neighbors, this definition being based upon the rigid-sphere concept of atoms, according to which the observed interatomic distance should be the distance for which neighboring spherical atoms come into contact, or twice the atomic radius. Since the nearest-neighbor distance is seldom precisely the same for two allotropic metal phases, it is clear that the rigid-sphere picture cannot be accurate. Nevertheless, the concept can be very valuable for semiquantitative work as will be seen when we discuss the

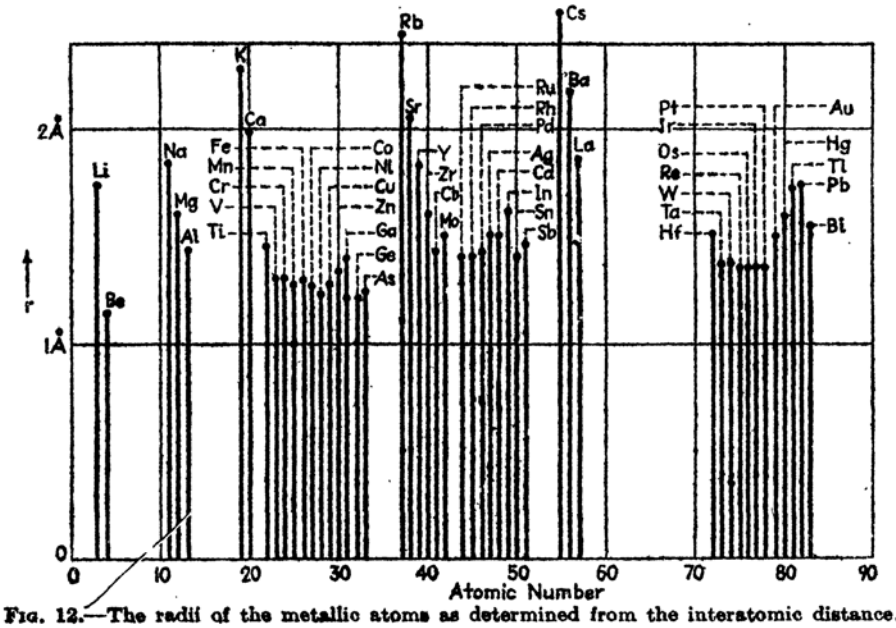

FIG. 12.—The radii of the metallic atoms as determined from the interatomic distance.

Hume-Rothery rules for alloys (*cf.* Sec. 3). Figure 12 shows the atomic radii as determined from the interatomic distances of Table II.

e. Electrical Conductivity.—The electrical resistivity ρ of a substance is a tensor quantity that is defined in terms of the electrostatic field intensity E and the current per unit area J by the relation

$$\mathsf{E} = \rho \cdot \mathsf{J}.$$

ρ is a constant tensor for cubic crystals and may be represented by a single number in these cases. It has two independent values for hexagonal and tetragonal crystals. These may be determined by using fields parallel and perpendicular to the principal axes, since E and J are parallel

to one another in these two cases. The two components of ρ are respectively designated as the $\parallel$ and $\perp$ components.

TABLE III.—THE RESISTIVITIES OF METALS AT ROOM TEMPERATURE

(The resistivity ρ is expressed in units of 10^{-6} ohm-cm. $\parallel$ and $\perp$ designate, respectively, values in directions parallel and perpendicular to the principal axis in hexagonal and tetragonal crystals.)

		Monovalent Metals	
Li	8.75	Cu	1.56
Na	4.35	Ag	1.49
K	6.62	Au	2.04
Rb	12.0		
Cs	19.0		
		Divalent Metals	
Be	5.5	Zn	$\parallel$6.0; $\perp$5.8
Mg	$\parallel$3.50; $\perp$4.22	Cd	$\parallel$8.4; $\perp$6.9
Ca	9.80	Hg	(-45.5°C) $\parallel$17.8; $\perp$23.5
Sr	32		
Ba	60		
Ra			
		Trivalent Metals	
Al	2.50	Ga	52.6
Sc		In	8.4
Yt		Tl	17.2
La	57.6		
		Tetravalent Metals	
Ti	47.5	Ge	89,000
Zr	41.0	Sn	$\parallel$13.1; $\perp$9.1
Hf	32.1	Pb	19.8
Th	18		
		Pentavalent Metals	
		As	$\perp$28
		Sb	39
		Bi	$\parallel$143; $\perp$109

Transition Metals

V	58.8	Nb	21	Ta	14
Cr	2.6	Mo	5.03	W	4.9
Mn α	710(?)	Ma		Re	18.9
Mn β	91(?)	Ru	7.64	Os	8.9
Mn γ	23	Rh	4.58	Ir	5.0
α Fe	8.71	Pd	10.2	Pt	9.81
α Co	6.2				
Ni	12.0				

Table III and Fig. 13 contain tabulations of the electrical resistivities[1] of monatomic metals at temperatures near 0°C. The lowest resistivities are those of the monovalent noble metals copper, silver and gold. In comparison, the alkali metals are only moderately good conductors.

[1] See, for example, the compilations of Landolt-Bornstein and the International Critical Tables.

Usually the resistivity decreases with increasing valence for the lighter elements and increases with increasing valence for the heavier elements, as may be seen by comparing the following two sequences:

Resistivity · 10^6 ohm-cm		Resistivity · 10^6 ohm-cm	
Na	4.35	Cu	1.56
Mg	3.50	Zn	5.8
Al	2.50	Ga	52.6

One of the most striking periodic properties of the resistivity is the large decrease that follows the completion of a d shell. The change in

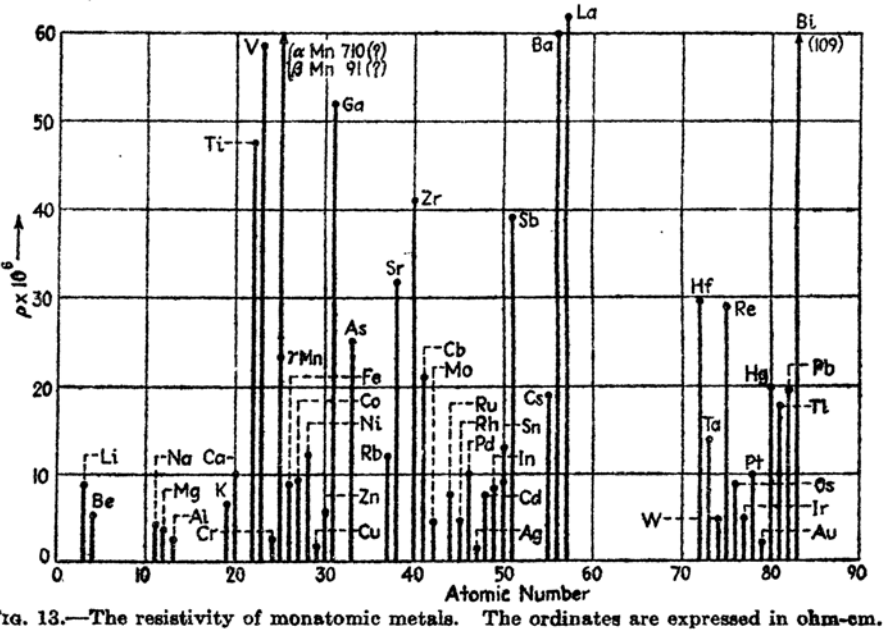

FIG. 13.—The resistivity of monatomic metals. The ordinates are expressed in ohm-cm.

resistance from $12 \cdot 10^{-6}$ ohm-cm for nickel to $1.56 \cdot 10^{-6}$ for copper is the most prominent illustration of this.

The metals with the highest resistivities are those such as arsenic, antimony, bismuth, tin, mercury, and gallium which have complex structures, a fact lending additional support to Hume-Rothery's view that these are intermediate between ideal metals and insulating crystals.

The ratio of the resistance at temperature T to the resistance at zero degrees centigrade is shown in Fig. 14 for a number of metals.[1] The fact that the curves are closely alike justifies our comparison, in the preceding paragraphs, of the room-temperature values. The high-temperature resistivity of most metals varies linearly with temperature, whereas the low-temperature resistance varies with a higher power of T. The

[1] *Ibid.*

most reliable measurements seem to show that the low-temperature variation is as T^5 for metals that are not superconducting. We shall discuss this topic more fully in Chap. XV.

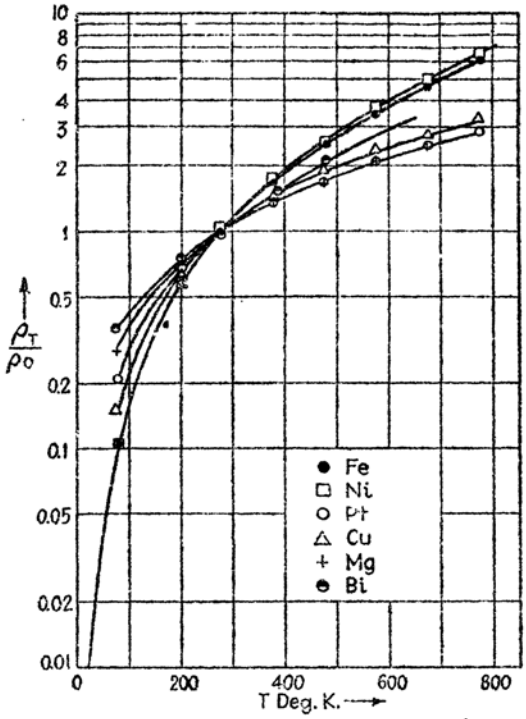

FIG. 14.—Temperature dependence of the relative resistivities of several metals. The ordinate is the ratio of the resistivity at temperature T to that at 0°C.

A large number of metals become superconducting below a temperature characteristic of the substance. These metals and their transition

TABLE IV.—THE SUPERCONDUCTING METALS AND THEIR TRANSITION TEMPERATURES

Metal	T_t, °K	Metal	T_t, °K
Zn	7.86	Ti	1.13
Cd	0.6	Zr	0.7
Hg	4.16	Hf	0.3
		Th	1.33
Al	1.14	Sn	3.72
Ga	1.07	Pb	7.2
In	3.38		
Tl	2.47	V	4.29
		Nb	9.22
		Ta	4.27

temperatures are listed in Table IV. There seems to be no striking regularity beyond the fact that none of the monovalent metals is superconducting. The resistances of nonsuperconducting metals at very low temperatures and the normal resistance of superconducting metals just above the transition temperature usually are dependent upon the previous history of the specimen on which measurements are made. It is believed[1] that, at least in principle, one can divide the resistance into two parts, namely, a part that is characteristic of the pure substance and that extrapolates to zero at the absolute zero, and a part ρ_r, generally termed the residual resistance, that arises from imperfections and that presumably would be zero for a perfectly pure undistorted crystal.

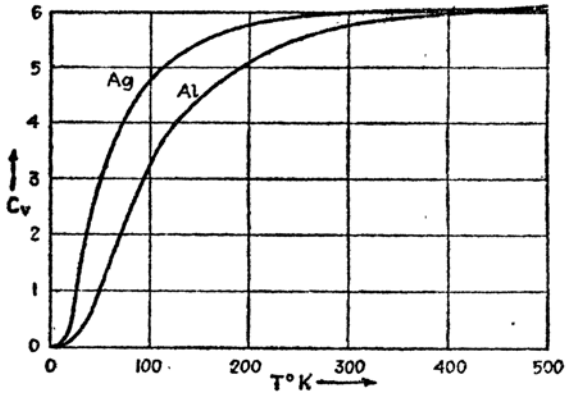

FIG. 15.—The atomic-heat curves of silver and aluminum. The ordinates are cal/mol.

Since ρ_r does not seem to vary reversibly during temperature changes, it usually cannot be separated from a set of measured resistances in a precisely quantitative way. It is known from fluctuations in resistance, however, that the residual resistance is of the same order of magnitude as the total resistance at 5°K. Since the room-temperature value is about a thousand times larger than this, the fluctuations caused by residual resistance are of the order of 0.1 per cent at ordinary temperatures.

f. Specific Heats.—Figure 15 shows the temperature dependence of the atomic heats at constant volume C_v of silver and aluminum.[2] These curves are typical of most of the simpler metals. They are characterized by a monotonic rise from zero at absolute zero to a nearly constant

[1] *Cf.* E. GRÜNEISEN, *Handbuch der Physik*, Vol. XIII (1928). More recent work, such as that of W. J. de Haas and G. J. van den Berg, *Physica*, **4**, 683 *ff.* (1937), seems to show that the residual resistance increases with decreasing temperature in the case of gold.

[2] Silver: A. EUCKEN, K. CLUSIUS, and H. WOLTINEK, *Z. anorg. Chem.*, **203**, 47 (1931). Aluminum: C. G. MAIER and C. T. ANDERSON, *Jour. Chem. Phys.*, **2**, 513 (1934).

value at high temperatures. This constant value should be about $3R$, according to the law of Dulong and Petit, where R is the gas constant. Actually, the measured values are slightly higher and rise with increasing temperature. The part of the curves near 0°K may often be approximated closely by the expression $R(T/\Theta_D)^3$, where Θ_D is a constant, known

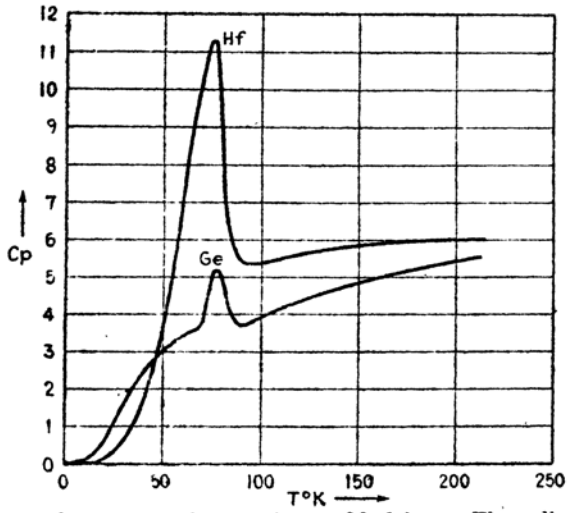

FIG. 16.—The atomic-heat curves of germanium and hafnium. The ordinates are cal/mol. *(After Simon and Cristescu.)*

as the Debye characteristic temperature. Table V contains values of Θ_D for several metals that exhibit this type of specific-heat behavior.

TABLE V.—CHARACTERISTIC TEMPERATURE OF SEVERAL SIMPLE METALS AS DETERMINED FROM THE T^3 LAW

(See Table XXXIV, Chap. III, for additional values.)

Metal	Θ_D
Ag	210
Ca	219
Zn	200
Tl	94
Sn	140
Bi	107

All the nontransition elements resemble silver and aluminum in that they have a limiting high-temperature atomic heat of about $3R$, but a number of them do not behave quite so simply at low temperatures. The differences vary from slight deviations from the T^3 law to greater ones represented by large peaks, such as those shown in the curves[1] for germanium and hafnium (Fig. 16). The metals that exhibit anomalies

[1] S. CRISTESCU and F. SIMON, *Z. physik. Chem.*, **25B**, 273 (1934).

of the extreme type generally have lattices in which more than one atom is contained in the unit cell.[1] For example, hafnium forms a close-packed hexagonal crystal, and germanium has the diamond lattice; both these types contain two atoms per unit cell.

The atomic-heat curves of the transition metals generally rise well above the Dulong and Petit value of $3R$ at high temperatures and increase linearly in this region. Figure 17 shows the behavior[2] for γ iron, which is a typical case. The ferromagnetic metals α iron and nickel show the same behavior but have additional peaks that accompany the decrease in their permanent magnetization. Figures 17 and 18 illustrate[3] these two cases.

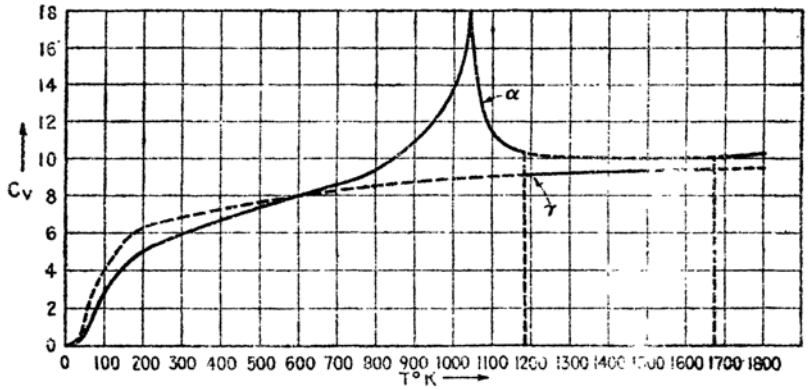

FIG. 17.—The atomic-heat curves of α and γ iron. The ordinates are cal/mol. (*After Austin.*)

The specific heat of nickel does not follow the T^3 law in the region below 10°K but becomes linear in accordance with the equation[4]

$$C_v = 0.001744T \text{ cal/deg-mol (see Fig. 19).} \qquad (1)$$

If the values given by this function are subtracted from the observed values, the residue is found to follow the T^3 law. This fact indicates

[1] If τ_1, τ_2, τ_3 are the primitive translation vectors of the lattice, the unit cell is the unit of the lattice from which the entire lattice may be generated by translations of the type

$$\mathbf{T} = n_1\tau_1 + n_2\tau_2 + n_3\tau_3$$

in which n_1, n_2, and n_3 range over all integer values. The volume of the unit cell is equal to the volume of the parallelepiped the edges of which are equal to τ_1, τ_2, τ_3, namely, $\tau_1 \cdot \tau_2 \times \tau_3$.

[2] *Cf.* the compilation of J. B. Austin, *Industrial Eng. Chem.*, **24**, 1225 (1932); **24**, 1388 (1932).

[3] See *ibid.* for iron and Landolt-Bornstein for a survey of the specific heats of nickel.

[4] W. H. KEESOM and C. W. CLARK, *Physica*, 2, 513 (1935).

that the total specific heat of nickel is composed of two parts, one that has the same source as the specific heats of the nontransition metals and one that has another origin. The first part is believed to arise from thermal excitation of lattice vibrations; there is fairly conclusive evidence,

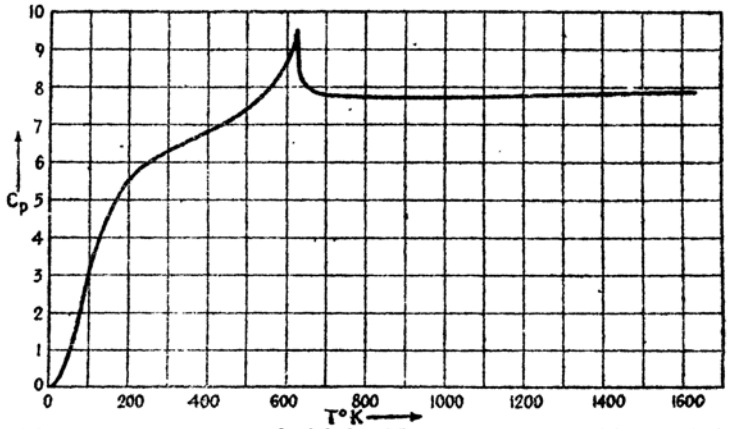

FIG. 18.—The atomic-heat curve of nickel. The ordinates are cal/mol. (*After Lapp.*)

which will be presented in Chap. IV, that the second is related to the excitation of the electrons in the unfilled *d* shells. In this connection, it is worth pointing out that at 1000°K the value of C_v in Eq. (1) is of the same order of magnitude as the difference between C_v and $3R$.

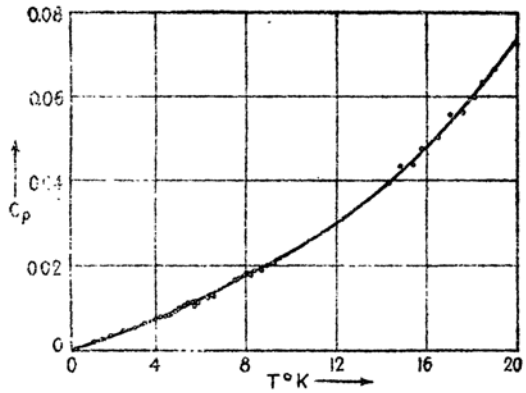

FIG. 19.—The atomic heat of nickel near absolute zero. The ordinates are cal/mol. (*After Keesom and Clark.*)

g. Magnetic Properties.—The magnetic susceptibility per unit volume χ, like the resistivity, is a tensor quantity. It is defined in terms of the magnetic field intensity H and the magnetization per unit volume M by the relation

$$\mathsf{M} = \chi \cdot \mathsf{H}$$

We shall ordinarily use cgs units for these quantities, expressing H in gauss and M in terms of the cgs unit of dipole moment per unit volume. Another important magnetic quantity is the permeability μ which is defined in terms of H and the magnetic induction B, by the equation

$$B = \mu H.$$

μ and χ are related by the equation

$$\mu = 1 + 4\pi\chi.$$

χ is practically independent of both temperature and field strength for a large number of nontransition metals; however, it varies with

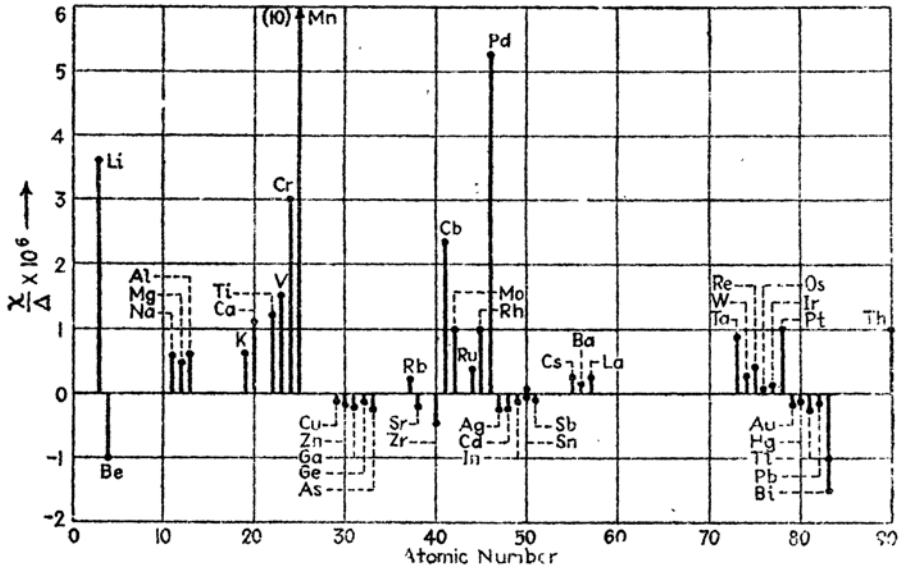

FIG. 20.—The specific susceptibilities of the monatomic metals at room temperature.

temperature in some cases. The room-temperature values for the simple metals are given[1] in Table VI and in Fig. 20. For convenience, the specific susceptibilities χ/Δ, where Δ is the density, rather than the susceptibilities are listed. It may be seen that the metals in the short periods and those which precede the transition elements are paramagnetic, that is, have positive susceptibilities, whereas the metals which follow the transition metals are diamagnetic. The susceptibilities usually are so small that traces of ferromagnetic impurity affect the measured values considerably and cause them to vary from specimen to specimen.

[1] See, for example, the compilations of Landolt-Bornstein and the International Critical Tables. As in the case of the resistivity, the scalar components of the susceptibility are listed.

In addition, the susceptibility of a given specimen depends markedly on the mechanical treatment it receives. For example, Bitter[1] found that the susceptibility of a piece of copper wire could be varied by as much as 40 per cent by stretching. Similarly, Honda and Shimizu[2] found that cold working changed the susceptibility of a sample of copper

TABLE VI.—ROOM-TEMPERATURE SPECIFIC MAGNETIC SUSCEPTIBILITIES OF MONATOMIC SUBSTANCES
(In cgs units)

	$\chi \cdot 10^6$		$\chi \cdot 10^6$
Monovalent Metals			
Li	3.6	Cs	0.2
Na	0.6	Cu	−0.08
K	0.6	Ag	−0.19
Rb	0.2	Au	−0.14
Divalent Metals			
Be	−1.0	Ba	0.15
Mg	0.5	Zn	−0.15
α Ca	1.1	Cd	−0.18
Sr	−0.2	Hg	−0.12
Trivalent Metals			
Al	0.6	Ga	−0.24
Sc		In	−0.10
Yt		α Tl	−0.22
β La	20		
Tetravalent Metals			
Ti	1.2	Ge	−0.12
Zr	−0.45	α Sn	−0.03
Hf		β Sn	0.03
Th	1.0	Pb	−0.12
Pentavalent Metals			
As	−0.25		
Sb	$\parallel -0.497$; $\perp -1.38$		
Bi	$\parallel -1.0$; $\perp -1.5$		
Transition Metals			
V	1.5	Pd	5.2
α Cr	3.0	Ta	0.9
α Mn	10	α W	0.28
Nb	2.3	Re	0.4
Mo	1.0	Os	0.05
Ru	0.4	Ir	0.13
Rh	1.0	Pt	1.0

from negative to positive and that annealing after cold work restored the original diamagnetism. It is probably true that measurements on perfectly pure, unstrained specimens of the same metal would be closely alike. However, ordinary materials do not conform to these conditions.

[1] F. BITTER, *Phys. Rev.*, **36**, 978 (1930); also, *Introduction to Ferromagnetism* (McGraw-Hill Book Company, Inc., New York, 1937).

[2] K. HONDA and Y. SHIMIZU, *Science Repts. Imp. Tôhoku Univ.*, **20**, 460 (1931); **22**, 915 (1933).

Figure 21 shows the temperature dependence of the measured[1] susceptibilities of a few nontransition elements.

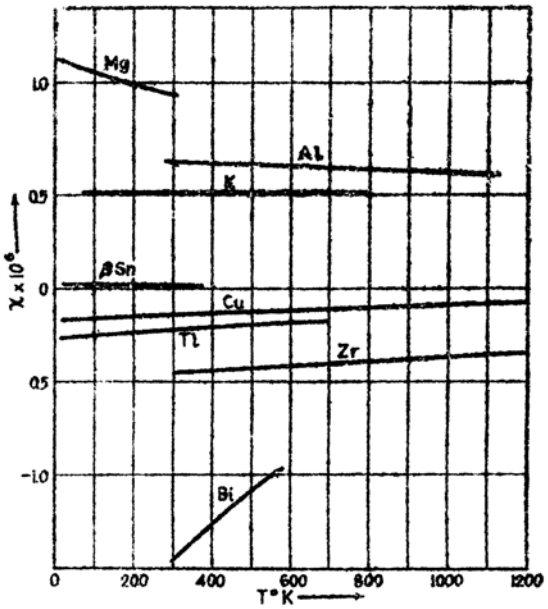

FIG. 21.—The temperature dependence of the specific susceptibilities of several metals.

The susceptibilities of the nonferromagnetic transition elements are all positive and are generally larger than those of the paramagnetic

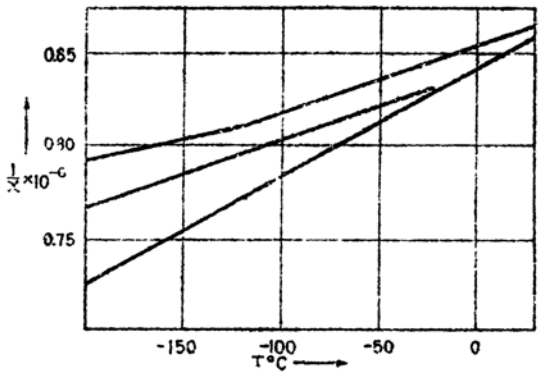

FIG. 22.—The temperature dependence of $1/\chi$ for several specimens of platinum. (*After Collet and Foëx.*)

simple metals. The temperature dependence[2] is shown in Fig. 22 for several specimens of platinum.

[1] See footnote 1, p. 17.

[2] P. COLLET and G. FOËX, *Compt. rend.*, **192**, 1213 (1931).

The susceptibilities of ferromagnetic solids are so strongly dependent upon field strength that the magnetic properties are described most conveniently by giving M as a function of H. Figure 23 shows[1] M versus H curves at room temperature for three directions in a single crystal of nickel. M and H are parallel in each of the three cases. It may be observed that M is small when no field is present and that it rises very rapidly at first as H increases. It approaches a saturation value M_s in the [111] direction[2] at about fifty gauss and then remains practically constant. The intensity curves for the [110] and [100]

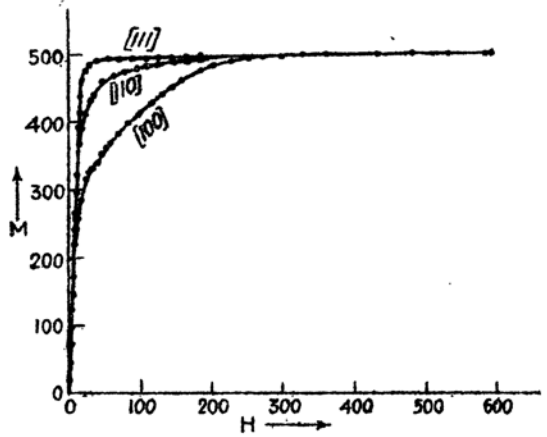

FIG. 23.—The magnetisation curves of a single crystal of nickel. The abscissa is expressed in gauss. (*After Kaya.*)

directions bend over sharply at values of about $M_s \cos 30°$ and $M_s \cos 60°$, respectively. They then rise relatively slowly and approach the value M_s.

This behavior may be interpreted in terms of the domain concept of ferromagnetic materials which was first proposed by Weiss.[3] According

[1] S. KAYA, *Science Repts. Imp. Tôhoku Univ.*, **17**, 639 (1928).

[2] We shall commonly use this notation in which a crystallographic direction is specified by a set of integers (Miller indices) that are proportional to the direction cosines. In cubic crystals, the coordinate axes are usually taken as an ordinary Cartesian system; in a hexagonal crystal such as cobalt, however, one coordinate axis is taken in the direction of the hexagonal axis of the crystal, and three others, separated by 120 deg, are taken in the plane normal to the hexagonal axis. In the second case, directions are specified by four integers, the last being proportional to the direction cosine between the given direction and the hexagonal axis. We shall specify planes in a similar way by giving the integers that are proportional to the direction cosines of the normals to the planes.

[3] P. WEISS, *Jour. phys.*, **6**, 661 (1907). An excellent discussion of the present status of domain theory may be found in the book edited by R. Becker, *Probleme der technischen Magnetisierungskurve* (Julius Springer, Berlin, 1938). See also the more recent review article by W. F. Brown, *Jour. App. Phys.*, **11**, 160 (1940), and the book by R. Becker and W. Döring, *Ferromagnetismus* (Julius Springer, Berlin, 1939).

to this concept, ferromagnetic substances contain a large number of small domains that have an intrinsic value of magnetic intensity equal to M_s even in the absence of an external field. It is assumed that the direction of this intensity lies along one member of a prominent set of equivalent crystallographic directions, this set being the eight directions equivalent to [111] in the case of nickel, for example. The resultant magnetization of the entire crystal is zero when H is zero, for the domains have their magnetization distributed uniformly among the eight [111] directions. If a weak field is applied in the [111] direction, all the domains have their magnetization changed to this orientation and the crystal becomes magnetized to the saturation value M_s. This process of rotation is demonstrated very convincingly by the Barkhausen[1]

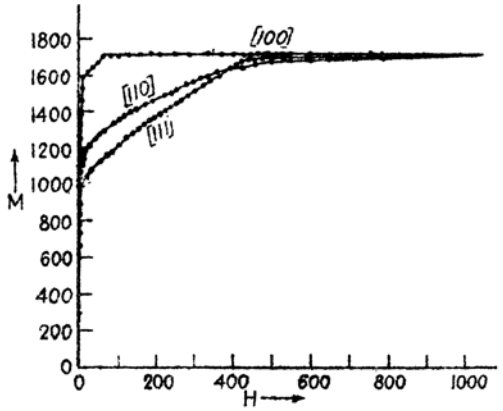

FIG. 24.—The magnetization curves of iron. The abscissa is expressed in gauss. (*After Honda and Kaya.*)

effect, which shows that magnetization takes place in very small discrete stages. The size of the domains may be estimated from the size of these steps if it is assumed that each jump represents the effect of one domain changing its direction. In this way, Bozorth and Dillinger[2] estimated that there are 10^9 domains per cubic centimeter.

It should be emphasized that the domains are not identical with crystalline units of the lattice, such as the grains in polycrystals. It now seems well established that the domain size, which may be larger or smaller than the grain size, is determined primarily by the magnetic interaction of different parts of a specimen and by variations in its internal stress. In addition, it should be mentioned that magnetization may take place by more or less continuous growth of properly oriented domains at their boundaries, much as crystals grow from nuclei.

[1] H. BARKHAUSEN, *Physik. Z.*, **20**, 401 (1919).

[2] R. BOZORTH and J. DILLINGER, *Phys. Rev.*, **35**, 733 (1930).

If a weak field is applied in the [100] direction instead of the [111] direction, the domains are reoriented as nearly parallel to the [100] direction as possible without leaving the eight [111] directions. Since the angle between the [100] direction and the four nearest [111] axes is 55°, it follows that the largest value of |M| that can be obtained in this way is $M_s \cos 55°$. This actually is the value at the bend in the [100] curve of Fig. 23. Since M increases beyond this value as H is increased, it follows that the intensities of the domains can be bent away from the normal directions of magnetization and eventually become parallel to the direction of the applied field. The curve for the [110] direction supports the same picture.

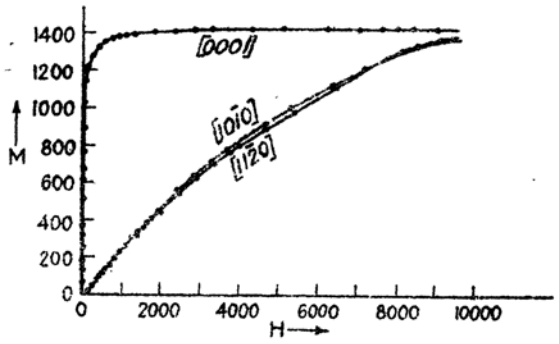

FIG. 25.—The magnetisation curve of cobalt. The abscissa is expressed in gauss. (*After Kaya.*)

Figures 24 and 25 give corresponding curves for iron[1] and cobalt.[2] The [100] direction is the easy direction of magnetization in the first case, whereas the hexagonal axis is in the second. The [1010] and [1120] direction curves for cobalt show no sharply rising portions because they are orthogonal to the easy direction of magnetization.

Akulov[3] has shown that one may account for the form of the magnetization curves of iron by assuming that the energy E_m of the crystal varies with the direction of magnetization of the domain in the manner

$$E_m = K_1(S_1^2S_2^2 + S_2^2S_3^2 + S_3^2S_1^2). \tag{2}$$

K_1 is a constant which ordinarily is determined experimentally, and S_1, S_2, S_3 are the direction cosines of the magnetic intensity. The total energy E_t of the crystal in the presence of a field H is then

$$E_t = E_m - \mathsf{H}\mathsf{M}\cos\varphi$$

[1] K. HONDA and S. KAYA, *Science Repts. Imp. Tôhoku Univ.*, **15**, 721 (1926).
[2] S. KAYA, *Science Repts. Imp. Tôhoku Univ.*, **17**, 1157 (1928).
[3] N. S. AKULOV, *Z. Physik*, **67**, 794 (1931).

TABLE VII.—MAGNETIC DATA OF THE FERROMAGNETIC MONATOMIC METALS
(The values of M_s and σ_s correspond to 0°K.)

	Fe	Co	Ni	Gd
M_s, cgs	1752	1446	512	1560
σ_s Bohr magnetons per atom	2.22	1.71	0.606	
Θ_c, °C	780	1075	365	16 ± 2
Θ_p, °C	774	1231	372	

where φ is the angle between H and M. The value of φ corresponding to equilibrium for a given field intensity is determined by the condition

$$\frac{dE_t}{d\varphi} = 0.$$

This leads to a relationship between φ and H, from which the component of intensity in the field direction may be determined as a function of H that involves the constant K_1. Figure 26 shows the calculated and observed values of M for the [111] and [110] directions of iron; these were obtained by using

$$K_1 = 2.14 \cdot 10^5 \text{ ergs.}$$

This method of correlating experimental measurements with energy expressions of the type of Eq. (2) has been extended by Gans,[1] Bozorth,[2] and others.

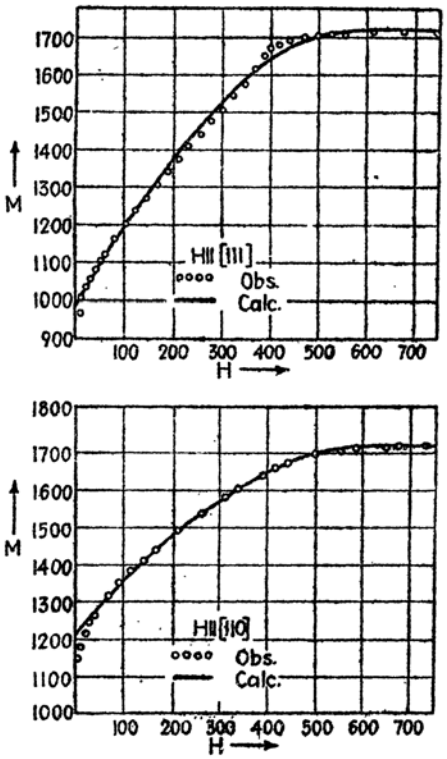

FIG. 26.—Calculated and observed magnetization curves for iron. (*After Akulov.*)

The magnetization curves vary with temperature in two striking ways: (1) The value of M decreases with increasing temperature and eventually approaches zero at the ferromagnetic Curie point Θ_c. (2) The relative values of the magnetization curves for different directions change with temperature. Figure 27 shows the ratio of $M_s(T)$ to $M_s(0)$, the value of M_s at 0°K, as a function of T/Θ_c for iron, cobalt, and nickel.

[1] R. GANS, *Physik. Z.*, **33**, 924 (1932).
[2] R. M. BOZORTH, *Phys. Rev.*, **50**, 1076 (1936).

Table VII contains[1] values of Θ_c for each of these substances as well as values of $M_s(0)$ and σ_s, the saturation moment per atom.

Above the Curie temperature, ferromagnetic crystals exhibit a paramagnetism that is the same order of magnitude as the paramagnet-

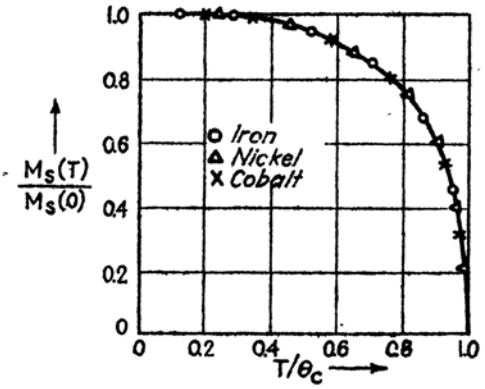

FIG. 27.—The temperature dependence of saturation magnetization for iron, cobalt and nickel. The abscissa is the ratio of the temperature to the Curie temperature θ_c and the ordinate is the ratio of the magnetization at temperature T to that at absolute zero. (*After Tyler.*)

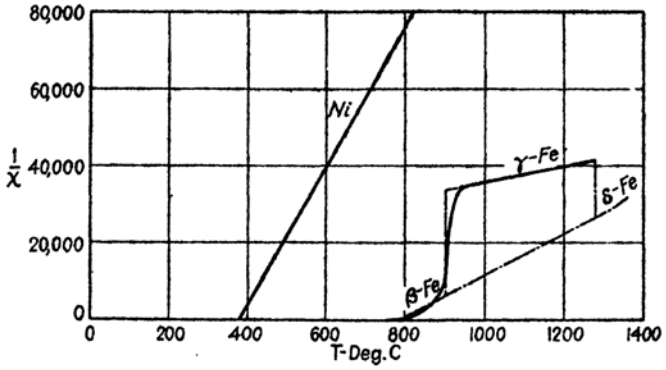

FIG. 28.—Temperature dependence of the paramagnetic susceptibility of nickel and iron above the Curie point.

ism of other transition metals. The susceptibility is highest at $T = \Theta_c$ and decreases with increasing temperature. The temperature dependence of the reciprocals of χ for iron and nickel is shown in Fig. 28. It may be noted that these curves are almost linear, a fact which shows that

$$\chi \cong \frac{C}{T - \Theta_p}$$

[1] After F. TYLER, *Phil. Mag.*, **9**, 1026 (1930); **11**, 596 (1931). See also E. C. STONER, *Magnetism and Matter* (Methuen & Company, Ltd., London, 1934).

where C and Θ_p are constants. Some values of Θ_p, which is called the paramagnetic Curie point, are given in Table VII.

We have remarked in f that the specific heat of a ferromagnetic metal has a sharp peak in the neighborhood of the Curie point. Peaks of this type appear in Figs. 17 and 18. Figure 29 gives a more detailed plot of measured values for nickel. This curve,[1] which is characteristic, also, of iron and cobalt, shows that the specific heat does not return to the normal $3R$ value above the Curie point.

It should be mentioned in passing that Urbain, Weiss, and Trombe[2] have found that metallic gadolinium is ferromagnetic, its ferromagnetic Curie point being 16°C.

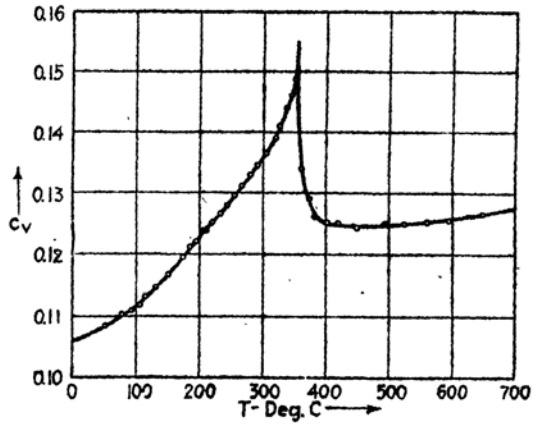

FIG. 29.—The specific heat of nickel in the vicinity of the Curie temperature. (*After Moser.*)

We shall not discuss the many intricate facts about the properties of polycrystalline ferromagnetic materials.

3. Metal Alloys.—The metallurgist defines a metal alloy as a combination of two or more monatomic metals that has metallic properties. This definition does not require that the material should be a homogeneous phase, and, indeed, many useful alloys are not. For simplicity, however, we shall restrict practically all of our discussion to single phases.

Alloys generally may be divided into two distinct classes, namely, substitutional alloys and interstitial alloys. In the first type the different constituent atoms occupy the same type of lattice position. Gold and silver form an alloy of this type in which both atoms occupy at random face-centered lattice positions. As we shall see, one general requisite

[1] H. MOSER, *Physik. Z.*, **37**, 737 (1936).

[2] G. URBAIN, P. WEISS, and F. TROMBE, *Compt. rend.*, **200**, 2132 (1935).

for the formation of an alloy of this type is that the radii of the constituent atoms be nearly equal.

In interstitial alloys, one or more kinds of atom fit into the interstices of the lattice formed by another kind. Low-concentration carbon steels are alloys of this type. In these, carbon atoms probably occupy some of the face-centered positions of the ordinary body-centered structure of iron. The interstitial atom usually is much smaller than the atoms of the lattice into which it fits.

The properties of a large number of substitutional alloys have been investigated extensively, whereas information concerning interstitial alloys seems to be fragmentary. The main reason for this deficiency is

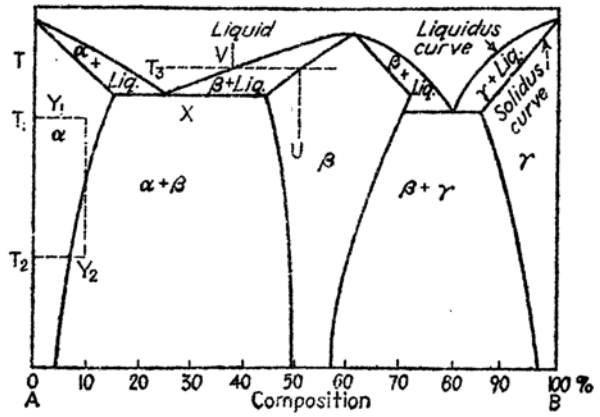

FIG. 30.—A symbolical phase diagram. The regions α, β and γ mark regions in which three different phases are stable. The phases α and β coexist at a point such as X in the region $\alpha + \beta$. The α phase Y_1 formed at the temperature T_1 may, in certain cases, be quenched to temperature T_2, where it is unstable, without actually reverting to the stable mixture of α and β.

that the small atoms, such as hydrogen, carbon, and nitrogen, of clear-cut interstitial cases are not good X-ray scatterers, so that structures of these alloys cannot be determined with certainty. We shall be concerned almost entirely with substitutional alloys in the following discussion.

a. The Phase Diagrams of Binary Alloys.—It is most convenient to discuss binary-alloy systems in terms of the conventional phase diagram. In this diagram the temperature-composition boundaries of the phases of the system are plotted as functions of composition. Figure 30 shows a typical case, the areas α, β, and γ marking regions in which three different phases exist. The composition and temperature may be varied within the limits of any one of these regions without changing the homogeneous structure of the material. If, however, one attempts to make an alloy corresponding to a point, such as X, that is not contained in one of these regions, the result is a mechanical mixture of two phases—the phases α and β in the case corresponding to the point X. It should be

added that this statement is rigorously true only if we imply thermodynamically stable phases, for it is possible to prepare an unstable phase that corresponds to the temperature and composition of a point such as X. For example, it is possible that the α phase, formed at temperature and composition corresponding to the point Y_1 at the temperature T_1, where the α phase is thermodynamically stable, could be cooled to the temperature T_2, where it no longer is stable, without breaking into two phases in a measurable time. This procedure, known as quenching, has great practical importance and depends upon the fact that the time required to attain thermodynamical equilibrium may be very long at sufficiently low temperatures.

It may be proved by means of thermodynamics[1] that the boundaries of different phases usually are not continuous but are separated as in Fig. 30 (*cf.* Sec. 123).

The liquidus curve shown in the figure marks the temperature at which a solid phase begins to separate from the molten solution of two metals. This curve has significance only over a range of composition in which the molten metals are miscible. The solidus curve, on the other hand, marks the temperature at which a solid phase of given composition begins to melt. The two curves coincide only at special points such as at the ends of the diagram. At temperature T_3, the liquid and solid phases that may be in equilibrium with one another are given respectively by the two intercepts that the temperature line makes with the liquidus and solidus curves. Consider, for example, the case of Fig. 30 again. Starting with the solid of composition U, we find that this begins to melt at temperature T_3 and that the composition of the first sample of molten metal corresponds to the point (V,T_3). Conversely, if we start with the liquid of composition V and cool it to temperature T_3, the solid that forms has the composition U. It follows that the composition of the solid and melt changes as the process of melting or freezing proceeds in either of these two cases.

It is possible to derive a number of important and interesting relationships among liquidus, solidus, and solubility-limit curves by use of thermodynamics. We refer the reader to other sources[1] for the development of these topics.

b. Rules of Combination of Binary Substitutional Alloys of Simple Metals.—A systematic investigation of the phase diagrams of metal alloys has led to the formulation of a number of simple rules that correlate many facts. We shall include a brief summary of these rules for

[1] See, for example, G. TAMMANN, *The States of Aggregation* (translation by R. F. Mehl, D. Van Nostrand Company, Inc., New York, 1925); R. VOGEL, *Handbuch der Metallphysik*, Vol. II (Akademische Verlagsgesellschaft, Leipzig, 1937).

reference in later work. They should not be accepted as though rigorous, for many exceptions exist.

1. *The rule of atomic size.*—This rule attempts to make more precise the qualitative notion that atoms must be nearly the same size if they form substitutional alloys over a wide range of composition. It has been developed by many workers, but the most nearly quantitative formulation has been given by Hume-Rothery, Mabbott, and Channel-Evans.[1] These workers find that atoms the radii of which differ by more than about 15 per cent do not form extensive solid solutions. If the difference is less than this, they are soluble over a wide range.. This rule is restricted by the condition that the radii must be derived from monatomic phases that have similar structures and that it should not be applied to systems

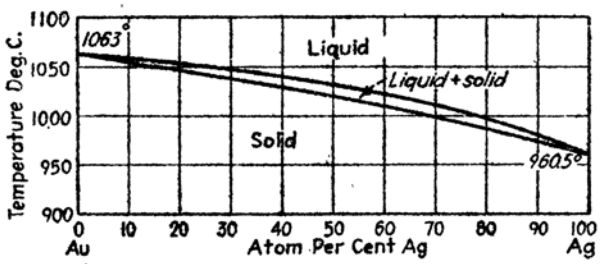

FIG. 31.—The phase diagram of the silver-gold system. This is an example of a case in which the components are completely miscible.

in which one of the atoms has a tendency to exhibit valence characteristics, as do the atoms of arsenic, antimony, and bismuth. These two conditions obviously are interrelated since atoms that have valence characteristics usually have complex lattice structures. In Table VIII, the range of solubility of different metals in copper and the range of solubility of copper in these metals are compared with the atomic radii.

TABLE VIII.—SOLUBILITY LIMITS OF THE PRIMARY PHASES OF SEVERAL COPPER ALLOYS

(The solubilities are expressed in atom percentage of the solute.)

System	Size factor	Solubility in Cu	Solubility of Cu
Cu-Be	Favorable	16.5 Be	2.0 Cu
Cu-Mg	Unfavorable	6.5 Mg	0.01 Cu
Cu-Zn	Favorable	38.4 Zn	2.3 Cu
Cu-Cd	Unfavorable	1.7 Cd	0.12 Cu
Cu-Ga	Favorable	20.3 Ga	Very small
Cu-Tl	Unfavorable	Small	Small
Cu-Ge	Favorable	12.0 Ge	Small

[1] W. HUME-ROTHERY, G. W. MABBOTT, and K. M. CHANNEL-EVANS, *Phil. Trans. Roy. Soc.*, **233**, 1 *ff.* (1934). See also HUME-ROTHERY, *op. cit.*

The silver-gold system is one of the most favorable cases for high solubility, according to this rule, for the lattice of both constituents is face-centered cubic, the valences are the same, and the atomic radii are

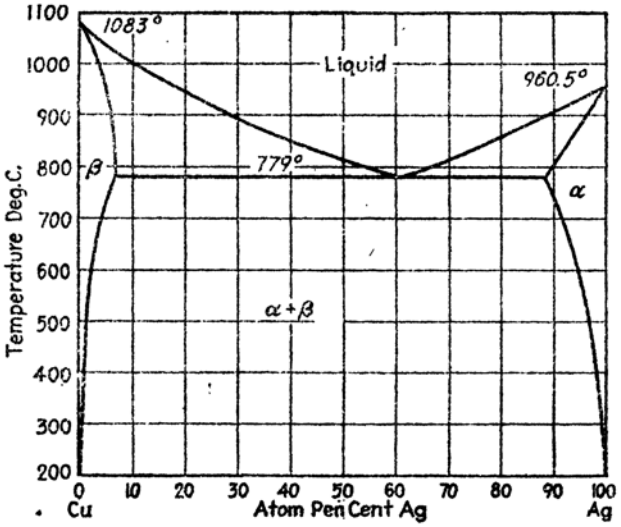

FIG. 32.—The phase diagram of the copper-silver system.

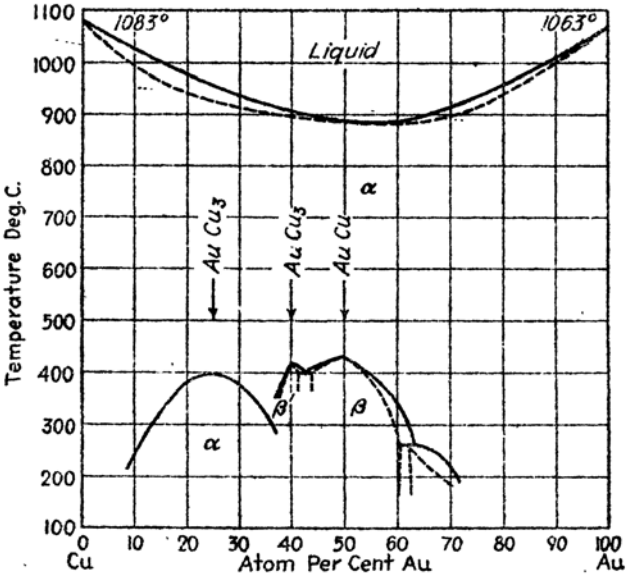

FIG. 33.—The phase diagram of the copper-gold system. These metals are completely miscible at all temperatures. The low-temperature curves correspond to ordered phases.

equal to within 2 per cent. Figure 31, which is the phase diagram of this system, shows that these metals form a single phase for the entire range of composition.

The radius of copper is about 13 per cent less than that of either gold or silver, and therefore the copper-gold and copper-silver systems should be borderline cases. The phase diagrams are shown[1] in Figs. 32 and 33. Whereas copper and silver do not mix, copper and gold are completely miscible except at low temperatures where more complex structures occur. These cases indicate that the rule of atomic sizes does not tell the entire story.

At the opposite extreme are lead and copper the radii of which differ by about 30 per cent and which do not mix in any proportion.

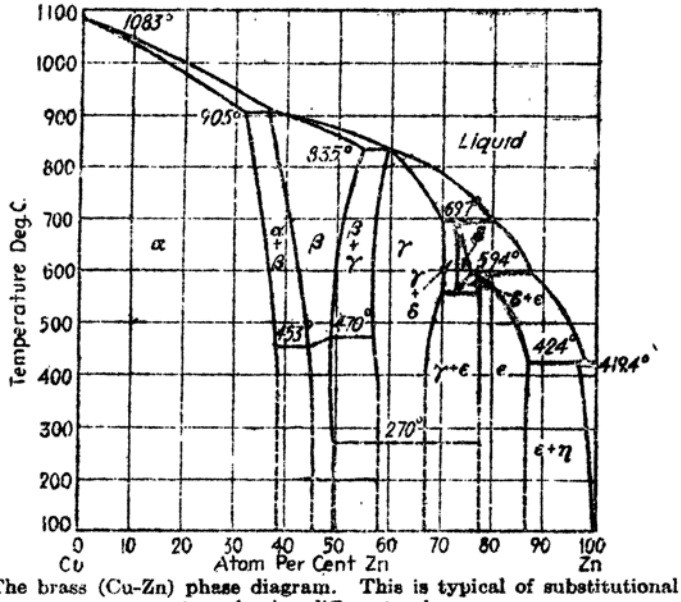

FIG. 34.—The brass (Cu-Zn) phase diagram. This is typical of substitutional alloys of atoms having different valences.

2. *The Hume-Rothery electron-atom ratio rules.*—The copper-rich and silver-rich portions of the phase diagrams of such systems as Cu-Zn, Cu-Cd, Cu-Al, Cu-In, Cu-Sn, Ag-Zn, Ag-Sn, in which the rule of favorable atomic sizes is satisfied, are strikingly similar. Figures 34 and 35 show the cases of Cu-Zn and Cu-Al which are typical examples. The β phase is body-centered cubic in both cases and appears immediately to the right of the primary face-centered phase of pure copper. The γ phase, which has a complex cubic structure, occurs next. The structure of this phase, which is shown in Fig. 36 for brass, is similar though not identical in all alloys; moreover, the γ phase generally has a high resistivity and a negative magnetic susceptibility and is brittle. The

[1] See M. HANSEN, *Aufbau der Zweistofflegierung* (Julius Springer, Berlin, 1936) for references on the alloy systems discussed here.

ϵ phase of the brass system, which is close-packed hexagonal, also occurs in the copper-tin system.

Although these similar phases usually occur for different atomic concentrations, Hume-Rothery[1] has pointed out that they occur for about the same value of the ratio of valence electrons to atoms. In computing the number of valence electrons, he used the usual chemical valences, namely, one for copper, two for zinc,

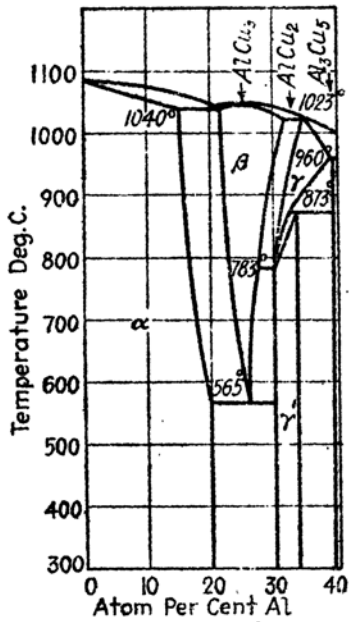

FIG. 35.—The copper-aluminum system.

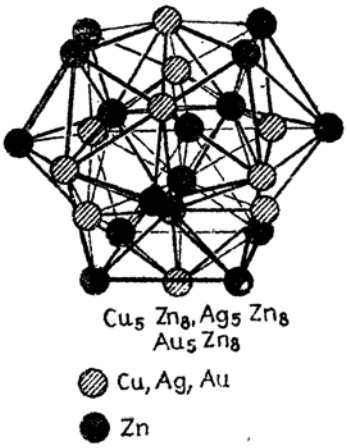

FIG. 36.—The γ brass structure.

three for aluminum, and so on. Table IX gives a compilation of metals that form one or more of the three alloy phases mentioned above and

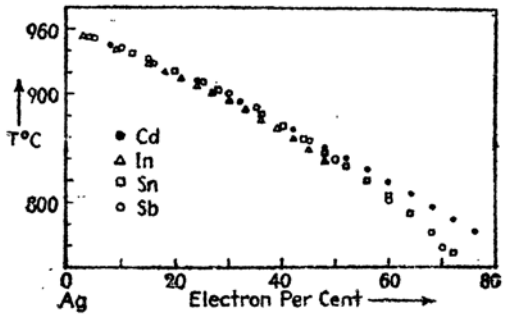

FIG. 37.—The liquidus curves of silver solutions.

that satisfy the Hume-Rothery rule. The electron-atom ratio characteristic of each structure is given at the head of each column.

[1] W. HUME-ROTHERY, *Jour. Inst. Metals*, 35, 309 (1926). See also HUME-ROTHERY, *op. cit.*

TABLE IX.—PHASES THAT CONFORM TO HUME-ROTHERY'S ELECTRON-ATOM RATIO RULE

Electron-atom ratio	1.5	1.61	1.75
Structure	β brass (b.c.c.)	γ brass type	h.c.p.
Nontransition cases	CuZn	Cu_5Zn_8	$CuZn_3$
	CuBe	Cu_5Cd_8	$CuCd_3$
	AgZn	Ag_5Zn_8	$AgZn_3$
	AgCd	.....	$AgCd_3$
	AuZn	Au_5Zn_8	$AuZn_3$
	AuCd	.....	$AuCd_3$
	Cu_3Al	Cu_9Al_4	Cu_3Sn
	Cu_3Ga	Cu_9Ga_4	Cu_3Ge
	Cu_3Sn	Cu_9In_4	Ag_3Sn
		$Cu_{31}Sn_8$	Au_3Al
Transition cases	CoAl	$CoZn_3$	
	NiAl		
	FeAl		

This rule is analogous to the ordinary rule of eight, being valid for substitutional alloys instead of for ionic and valence compounds.

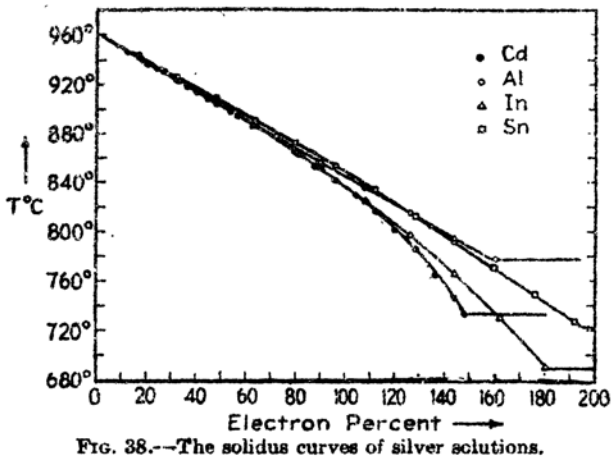

FIG. 38.—The solidus curves of silver solutions.

3. *The Hume-Rothery liquidus- and solidus-curve rules.*[1]—The liquidus curves of the primary solid solutions of elements such as zinc, cadmium, gallium, indium, tin, and antimony in copper and silver are similar in form and coincide if the electron percentage instead of the atomic

[1] W. HUME-ROTHERY, G. W. MABBOTT, and K. M. CHANNEL-EVANS, *Phil. Trans. Roy. Soc.*, **233**, 1 *ff.* (1934). See also HUME-ROTHERY, *op. cit.*

percentage is used as abscissa. In cases in which the solvent is monovalent, the electron percentage is obtained by multiplying the atom percentage with the valence of the solute. The liquidus curves of silver are given in Fig. 37 in order to show the extent to which the rule is valid. Figure 38 shows that the solidus curves obey the same principle.

c. Alloys Involving Metals with Strong Valence Characteristics.—The phase diagrams of systems in which one of the constituents is a metal of low valence, such as copper, silver, zinc, or magnesium, and the other is a less electropositive atom, such as arsenic, antimony, or bismuth, show that these substances do not combine to form extensive solid solutions, even when the size factors are favorable. Consider, for example, the phase diagram[1] of the magnesium-antimony system shown in Fig. 39.

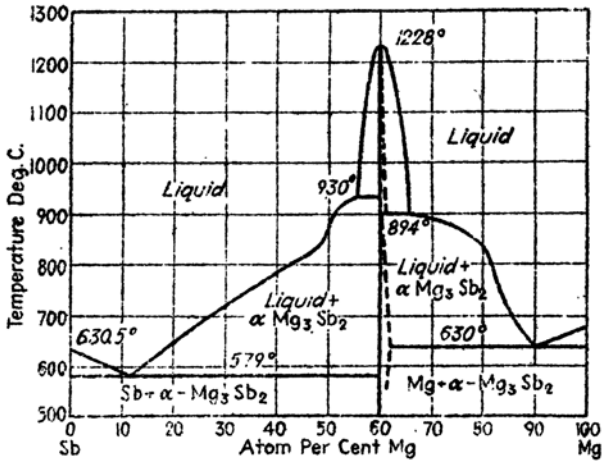

FIG. 39.—The magnesium-antimony system. These atoms are completely immisable except for the compound Mg_3Sb_2.

The fact that the primary phases are narrow shows that neither atom is appreciably soluble in the lattice of the other. The intermediate phases occur over narrow regions of composition and for atomic ratios that are characteristic of ionic or valence compounds rather than of ideal metal alloys; moreover, the structures of these phases are often similar to those of ionic crystals. For example, the only intermediate phase in the magnesium-antimony system is the compound Mg_3Sb_2 in which the constituents are exhibiting their normal electropositive and electronegative valences. This compound exists in two phases, both of which have the structures of rare earth metal oxides.

We may conclude from evidence such as this that these alloys form part of a bridge between ideal metals and ionic crystals.

[1] See footnote 1, p. 30.

It should be added that arsenic and antimony are completely miscible in one another, a fact showing that they do not form ionic-like lattices unless they are combined with strongly electropositive elements.

d. Rules for Combination of Transition Metals.—When transition metals combine with simpler metals, they obey fairly closely the three

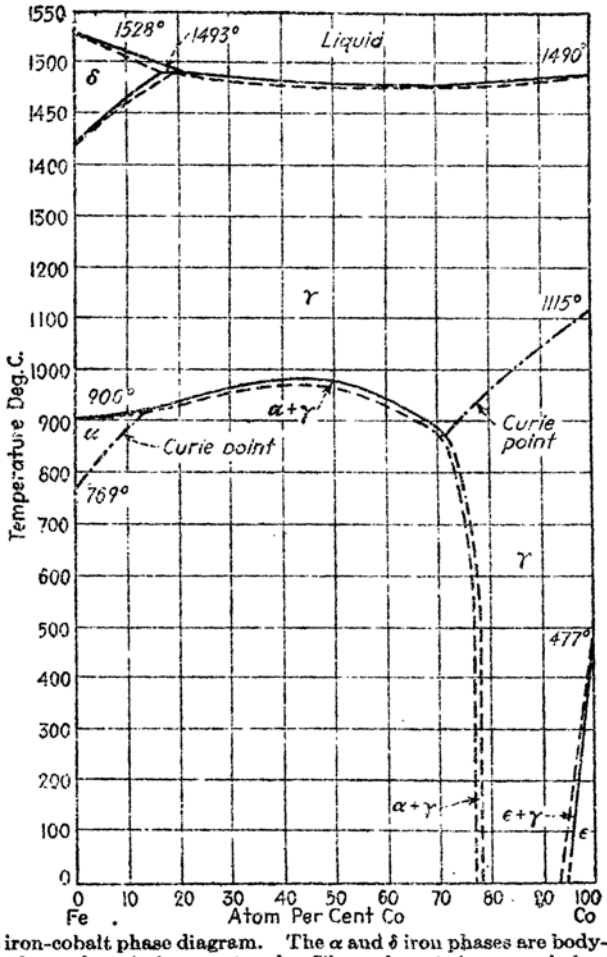

FIG. 40.—The iron-cobalt phase diagram. The α and δ iron phases are body-centered cubic, whereas the γ phase is face-centered. The ϵ phase is hexagonal close-packed.

rules that were presented in part *b*. In applying the rules, however, it is necessary to treat transition metals as though their valences were practically zero. For example, Table IX contains several alloys that have structures compatible with Hume-Rothery's electron-atom ratio rule (2 of part *b*) if this assumption is made.

Transition elements in the same row of the periodic chart have almost identical radii and combine over wide ranges of composition. Figure 40

shows the phase diagram of the iron-cobalt system which is a typical example.

Transition metals form an interesting sequence of interstitial alloys that has been studied extensively by Hägg.[1] He has found that the "metalloid" atoms hydrogen, nitrogen, carbon, and boron enter into the interstices of transition metals forming alloys of composition M_4X, M_2X, MX, and MX_2, where M is the metal atom and X is the metalloid, if the ratio of the radius of the metalloid atom to that of the metal is less than 0.59. This family of alloys usually forms lattices in which the metal atoms are arranged in cubic or hexagonal close-packed structures, although there are a few notable exceptions, such as tungsten carbide, WC, in which the tungsten atoms possess a simple hexagonal

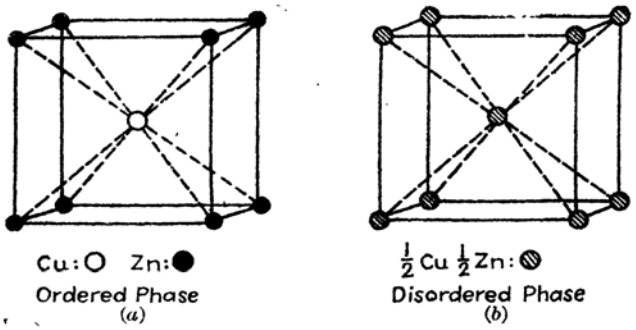

FIG. 41.—The ordered and disordered phases of β brass.

arrangement. The iron-carbon system lies just outside the domain of applicability of Hägg's rules, for the radius ratio is about 0.63 in this case; however, the iron-nitrogen system is a typical one for which they are valid. The nontransition metals do not usually form genuinely metallic interstitial alloys when combined with the metalloid atoms, but rather tend to form more nearly ionic crystals, such as calcium carbide, presumably because they are more electropositive than the transition metals.

e. Ordered and Disordered Phases.—In an ideal substitutional alloy, different kinds of atom occupy a given set of lattice positions at random. Many alloys in which the atoms have this property at high temperatures change as the temperature is lowered. Consider, for example, the case of β brass,[2] which has the composition CuZn and the structure shown in Fig. 41*b*. At high temperatures, each site is occupied with equal prob-

[1] G. HÄGG, *Z. phys. Chem. B*, **6**, 221 (1929); **7**, 339 (1930); **8**, 455 (1930).

[2] The possibility of order and disorder was first suggested by G. Tammann, *Z. anorg. Chem.*, **107**, 1 (1919). The structures of ordered and disordered β brass were first established by X-ray methods by F. W. Jones and C. Sykes, *Proc. Roy. Soc.*, **161**, 440 (1937).

ability by either type of atom. Below 480°C, however, the copper atoms prefer[1] body-centered positions and the zinc atoms prefer cube corners. These preferences increase as the temperature is lowered, and the structure becomes that of Fig. 41*a* at very low temperatures. The ordering process takes place continuously with decreasing temperature in this case and is completely reversible if the system is maintained at equilibrium. If we let p_{Cu} designate the probability that a body-centered position may be occupied by a copper atom and p_{Zn} the probability that it may be occupied by a zinc atom, we may conveniently define an order parameter S by the equation

$$S = p_{Cu} - p_{Zn}.$$

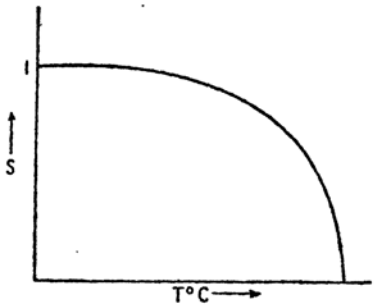

FIG. 42.—Order versus temperature for β brass (schematic).

This parameter varies from 0 to 1 as the lattice passes from the relatively disordered high-temperature phase to the ordered structure. Figure 42 shows schematically the way in which S depends upon temperature in β brass. This curve resembles closely the curve that shows the dependence of saturation magnetization of ferromagnetic materials upon temperature (*cf.* Fig. 27). The analogy with ferromagnetism becomes even more striking when one examines the specific heat curve[2] of β brass, which is shown in Fig. 43. It may be seen that there is a sharp peak, similar to the peak that occurs in nickel at the Curie temperature, at the temperature where ordering begins.

All changes between the ordered and disordered phases do not occur so gradually as that observed in β brass. For example, in the Cu_3Au system, a high degree of order occurs abruptly when the alloy is cooled below 380°C. This abrupt and reversible change is accompanied by the appearance of a latent heat.

Other substitutional phases that exhibit ordering are shown in Fig. 44.

f. Additional Properties of Substitutional Alloys of Nontransition Metals.—The thermal, electrical, and magnetic properties of metal alloys are, on the whole, much the same as those of monatomic metals. There are, however, a few striking differences that make additional discussion worth while. In this section, we shall consider alloys of nontransition metals.

[1] The body-centered and cube-corner positions are completely interchangeable in this case.

[2] MOSER, *op. cit.*

1. *Thermal properties.*—The difference between the heats of formation of alloys and of the metallic phases of their constituents has been

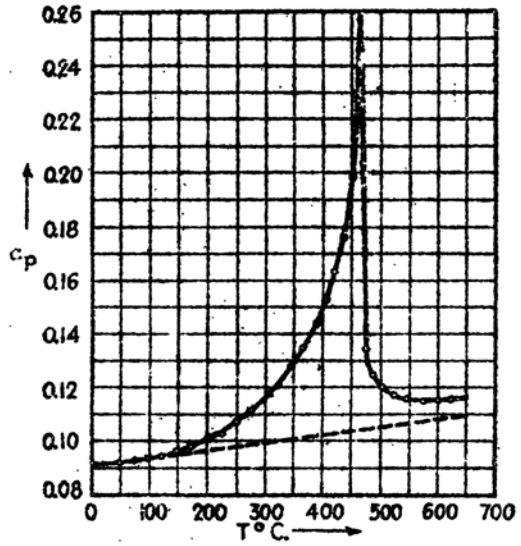

FIG. 43.—The specific heat of β brass during the transition from the ordered to a disordered structure. (*After Moser.*)

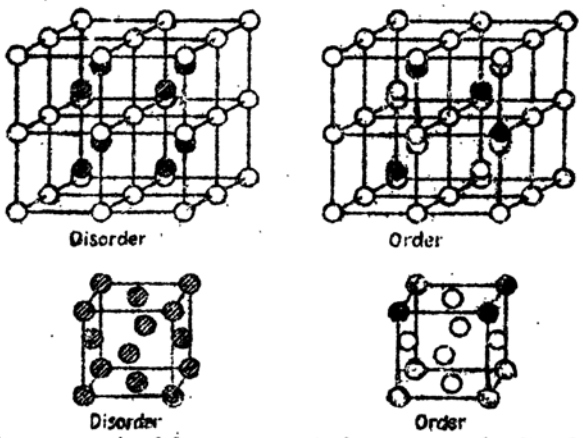

FIG. 44.—The upper pair of figures respectively represent the disordered and ordered arrangements of Fe_3Al ○ = Fe; ◎ = Al; ● = $\frac{1}{2}Fe+\frac{1}{2}Al$. The lower pair represent the disordered and ordered structures of CuPd. ○ = Cu; ● = Pd.

determined in a number of cases. Three methods are commonly used: (1) comparative measurements of the heat evolved when monatomic metals and alloys are dissolved in acids; (2) direct calorimetric measurements of the heats of reaction of monatomic metals; and (3) measurement of emfs of cells in which one of the electrodes is the alloy under investiga-

tion. The room-temperature values[1] of some results of this work, expressed in terms of kilogram calories per gram-atom, appear in Table X. The experimental errors are usually of the order of magnitude of 1 kg cal.

TABLE X.—ROOM-TEMPERATURE VALUES OF THE HEATS OF FORMATION OF ALLOYS
(In kg cal/gram-atom)

Cu-Zn System			
β brass (unordered)	2.2		
γ brass	2.6		
$CuZn_3$	1.8		
Ag-Cd System			
β phase (unordered)	1.31		
γ phase	1.42		
$AgCd_3$	1.23		
Au-Zn System			
Au_3Zn	6.0		
$AuZn$	5.5		
$AuZn_3$	5.6		
Miscellaneous Cases			
Mg_2Sn	20	$MgLa$	2.9
$CaAl_4$	13	$CdSb$	1.8
Ca_2Zn_3	8	$AuSb_2$	1.2
Na_2Sn	7	Tl_7Bi	0.7
$NaHg$	5.4	Hg_5Tl_2	0 06
Cases Involving Transition Elements			
Ni_3Sn	5.8		
Ni_3Sn_2	7.5		
$NiSn$	7.5		
Al_5Co_2	12		
Al_3Fe	6.3		

The nearly ideal substitutional alloys, like the Cu-Zn and Ag-Cd systems, are not bound so tightly as compounds such as Mg_2Sn that evidently are transition cases. On the whole, however, it does not seem to be possible to draw any striking conclusions from this table.

The specific heats of nontransition metal alloys usually resemble those of monatomic metals in approaching zero at 0°K, in obeying Dulong and Petit's law at high temperature, and in increasing monotonically in the intervening region. The exceptions are those phases which undergo allotropic changes or which become ordered as described in part *e*. In the first case, there is a discontinuity in the specific heat curve and a latent heat, just as for any allotropic change. The behavior of the specific heat during ordering was discussed in part *e*.

Early investigation of the specific heats of alloys led to the formulation of the Kopp-Neumann law, which states that the molecular heat of any alloy is equal to the sum of the atomic heats of its constituent

[1] These are taken from the compilation of W. Biltz, *Z. Metallkunde*, **29**, 73 (1937). See also W. Seith and O. Kubaschewski, *Z. f. Elektrochem.*, **43**, 743 (1937).

TABLE XI.—THE MOLECULAR HEAT OF Ag_3Au

Temperature, °C	Observed molar heat, cal	Sum of atomic heats, cal
100	24.926	24.942
200	25.415	25.475
300	26.005	26.000
400	26.599	26.513
500	27.195	27.012
600	27.789	27.500
700	28.384	27.979
800	28.937	28.463

monatomic metals. More modern work has shown that this law is never precisely correct, although it is often correct to within 10 per cent. Table

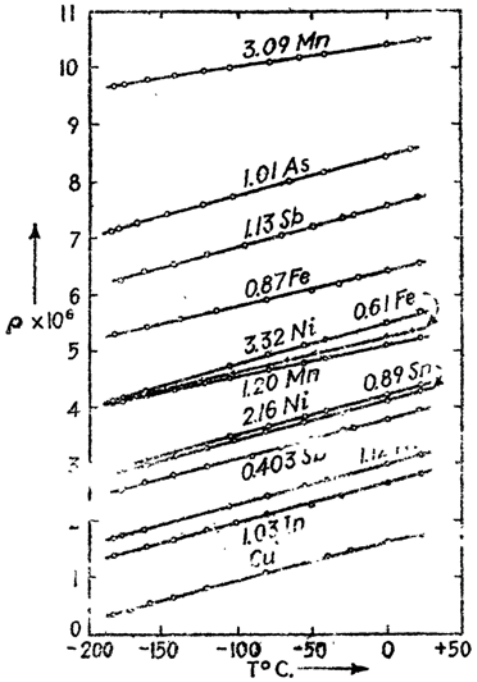

FIG. 45.—The resistivity versus temperature curves of a number of copper alloys in the range in which the resistivity of copper varies linearly with temperature. It should be observed that the resistivities of the alloys are much higher at low temperatures, implying very large residual resistivities. The resistivities are expressed in ohms-cm. The numbers indicate the atomic per cent of the alloying metal. (*After Linde.*)

XI gives a comparison[1] of the molecular specific heat of Ag_3Au and the sum of the atomic heats of the constituents over a 700° temperature

[1] J. A. BOTTEMA and F. M. JAEGER, *Proc. Roy. Soc. Amsterdam*, **35**, 928 (1932).

range. The close agreement in this case is partly a consequence of Dulong and Petit's law, since both the pure metals and the alloy obey it closely. It is evident that the Kopp-Neumann rule will fail badly in any temperature range in which the alloy becomes ordered.

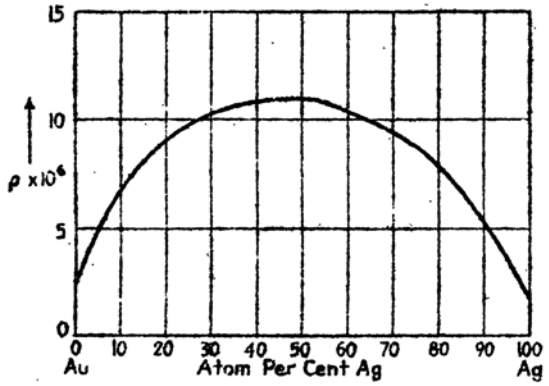

FIG. 46.—The resistivity of the silver-gold system at room temperature.

2. *Electrical resistivity of alloys.*—One of the most striking characteristics of the temperature-resistance curves of alloys is the fact that they do not extrapolate to zero at absolute zero so closely as those of monatomic metals do. In other words, their residual resistance usually is

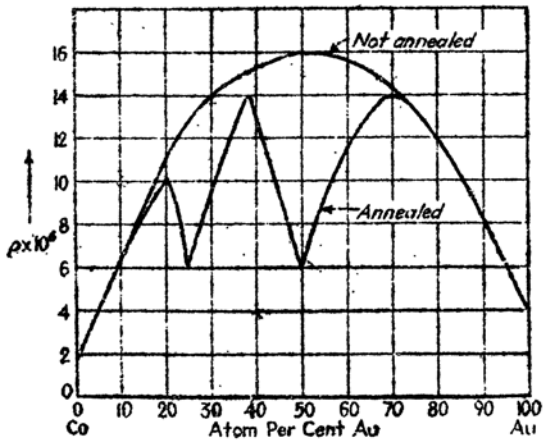

FIG. 47.—The room-temperature resistivity of quenched and annealed specimens of copper-gold alloys. The quenched specimens, which are not ordered, have the typical resistivity versus composition curve of perfect solid solutions. The annealed alloys show resistance minima at the compositions at which ordering occurs.

very large. This fact is shown by the curves[1] of Fig. 45 which are temperature-resistance plots for a number of copper alloys. One of the

[1] J. O. LINDE, *Ann. Physik*, 15, 219 (1932).

natural consequences of this large residual resistance is the fact that the low-temperature resistance of dilute primary solid solutions rises with increasing concentration of the solute. This behavior[1] is shown in

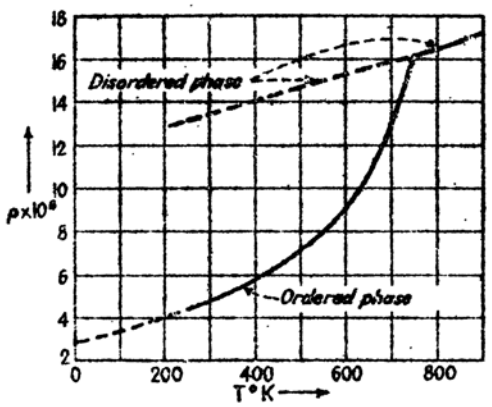

FIG. 48.—The temperature-resistivity curve of β brass during the ordering transition. The residual resistivity of the ordered alloy is much lower than that of the disordered phase.

Fig. 46 by the curve of resistance versus concentration for the Au-Ag system.

The resistance of an ordered phase invariably is lower than that of the disordered one, and there usually is a sharp kink in the curve of resistivity

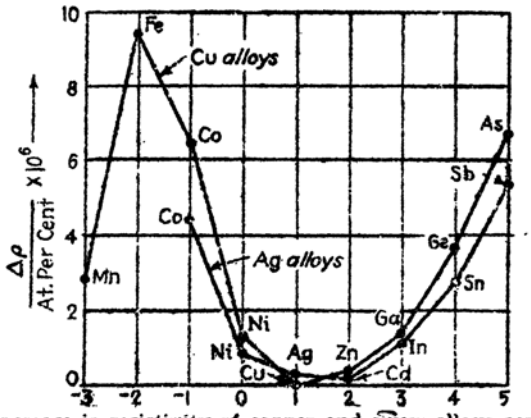

FIG. 49.—The increase in resistivity of copper and silver alloys per atom per cent of solute. The abscissa is the number of valence electrons of the solute atom relative to the closed *d* shell.

versus temperature at the temperature where ordering begins. These facts are illustrated[2] in Figs. 47 and 48, which show the dependence of

[1] See, for example, the references in Landolt-Bornstein.

[2] C. H. JOHANSSON and J. O. LINDE, *Ann. Physik*, **5**, 762 (1930); **25**, 1 (1936). G. BORELIUS, *Proc. Phys. Soc.* (*Sup.*), **49**, 77 (1937).

resistivity on composition and on temperature for some alloys that were well enough annealed to allow ordering to take place.

Different atoms, dissolved in a given solvent metal, affect the resistivity in different ways. If, for example, one plots the increase in resistance per atom percentage of solute for different copper and silver solutions, one obtains the curves[1] of Fig. 49, which show that the resistance increases with the difference in valence of the solvent and solute atoms.

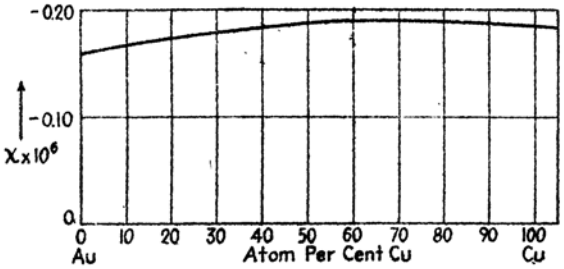

FIG. 50.—The dependence of magnetic susceptibility upon composition in the copper-gold system. (*After Shimizu.*)

3. *Magnetic susceptibilities.*—The specific magnetic susceptibilities of alloys of the nontransition metals are of the order of magnitude of 10^{-6}, just as are those of pure nontransition metals. If the constituents are soluble in all proportions and if the alloys do not form ordered phases, the variation of susceptibility with composition usually is uniform. Examples[2] of such cases are shown in Figs. 50 and 51. In other cases, particular phases have their own magnetic properties which may be

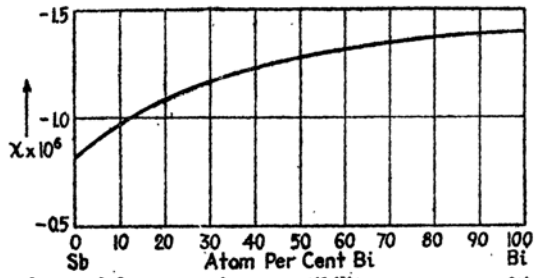

FIG. 51.—The dependence of the magnetic susceptibility upon composition in the antimony-bismuth system. (*After Shimizu.*)

considerably different from those of the pure metal. Thus, the γ brass type of phase usually is strongly diamagnetic, whereas the β brass type is usually normal. The susceptibility of the brass system[3] is shown in Fig. 52.

[1] See footnote 1, p. 40.

[2] Y. SHIMIZU, *Science Repts. Imp. Tôhoku Univ.*, **21**, 826 (1932).

[3] H. ENDO, *Science Repts. Imp. Tôhoku Univ.*, **14**, 479 (1925).

g. Additional Properties of Transition-metal Alloys. 1. *Thermal properties.*—The heats of formation of only a few alloys that contain transition metals have been measured. Several of these values are contained in Table X. Generally speaking, they are of the same order of magnitude as those of phases of nontransition metals.

The specific heats of the transition-metal alloys generally show the same types of behavior as the specific heats of monatomic transition metals. Thus, they do not obey Dulong and Petit's law at high temperatures, if they are strongly paramagnetic or ferromagnetic; and if ferromagnetic, they exhibit "anomalous" peaks near the Curie point. There is a close correlation between the height of such peaks and the saturation moment of the ferromagnetic substance. For example, the addition of copper to nickel quenches the magnetization of the latter,

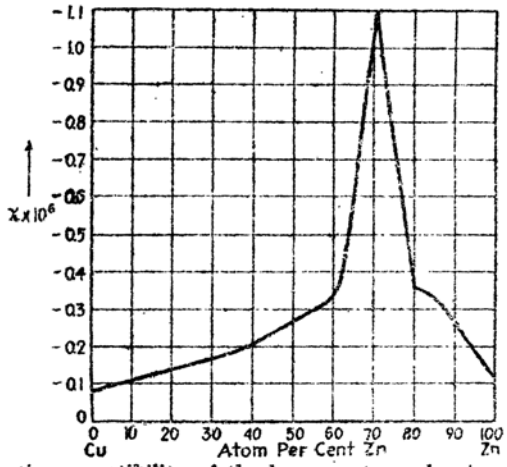

FIG. 52.—The magnetic susceptibility of the brass system, showing the large peak associated with the γ phase. (*After Endo.*)

and Fig. 53 shows[1] that the peak in the specific-heat curve disappears as the percentage of copper increases. Similar behavior has been observed in the chromium-nickel system.

2. *Electrical conductivity.*—The effect of temperature upon the electrical conductivity of alloys that contain transition metals has not been investigated so widely as has that upon the electrical conductivity of simpler alloys. The general facts, however, probably are about the same.[2] In disordered alloys, for example, there is a large residual resistance that decreases with increasing order. Typical composition-

[1] K. E. GREW, *Proc. Roy. Soc.*, **145**, 509 (1934).

[2] The resistance of several transition-metal alloys, such as constantan (Cu30 Ni43) and manganin (Cu84 Mn12 Ni4), are nearly temperature-independent over a temperature range of several hundred degrees.

resistance curves are enough like those of nontransition cases to require no additional comment.

3. *Magnetic properties.*—The magnetic properties of this group of alloys form a large and interesting body of material which we have only limited space to discuss. Most of the alloys are paramagnetic and have

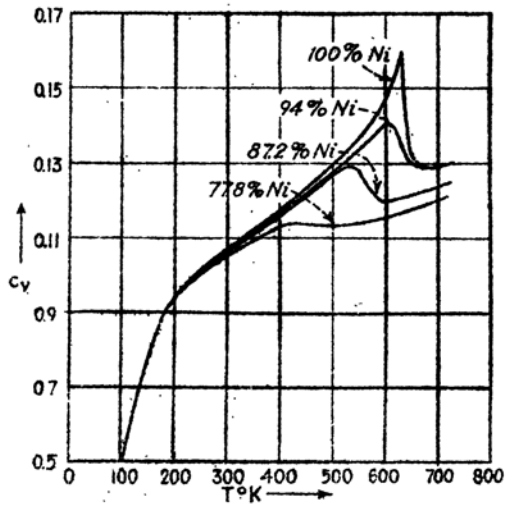

FIG. 53.—The quenching of the magnetic specific heat of nickel by addition of copper. (*After Grew.*)

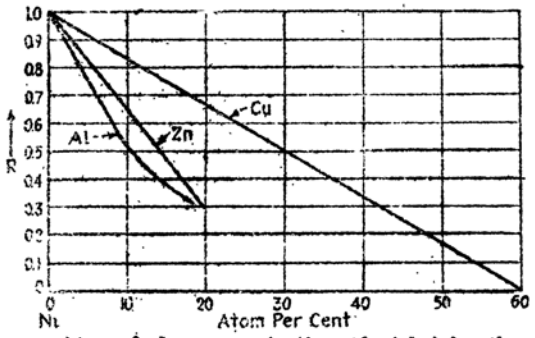

FIG. 54.—The quenching of the magnetization of nickel by the addition of simple metals. R is the ratio of the saturation magnetization at a given composition to that of pure nickel. (*After Sadron.*)

susceptibilities that decrease with increasing temperature, just as do the susceptibilities of monatomic transition metals. The alloys that contain one of the ferromagnetic metals, however, are ferromagnetic, at least for large concentrations of the ferromagnetic metal. This ferromagnetism usually decreases with increasing dilution of the ferromagnetic constituent if the other constituent is not ferromagnetic. For example,[1]

[1] C. SADRON, *Ann. Physik*, **17**, 371 (1932).

Fig. 54 shows the decrease in the saturation magnetization of a number of nickel alloys as their composition is varied. These curves show the customary behavior, namely, the magnetization decreases uniformly as the concentration of solute is increased. The manganese-nickel system is an exception to this rule, for the magnetization passes through two peaks when manganese is added to nickel in gradually increasing amounts[1] (*cf.* Fig. 55).

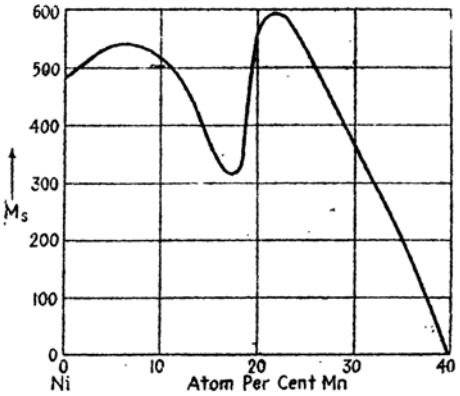

FIG. 55.—Dependence of the saturation magnetization of nickel-manganese alloys upon composition. (*After Kaya and Kussmann.*)

The alloys of ferromagnetic elements are all ferromagnetic. Figure 56 shows the behavior[2] of the saturation moment and the Curie point in the iron-cobalt system.

The alloys of copper and manganese[3] have the susceptibility curves shown in Fig. 57 at room and liquid-air temperatures. The susceptibility of the phase that contains about 23 per cent of manganese is very high, indicating a strong

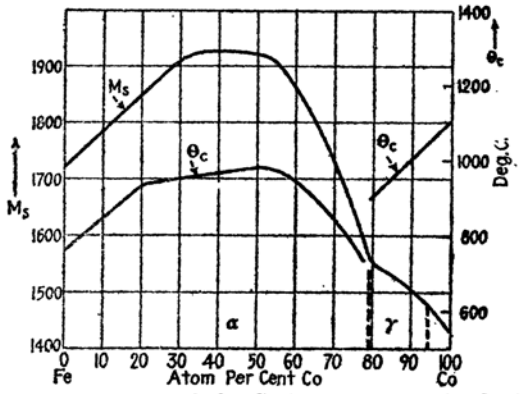

FIG. 56.—The saturation moment and the Curie temperature in the iron-cobalt system.

tendency toward ferromagnetism. By adding aluminum or tin to this system, one obtains the Heusler alloys, of which some, such as the phase[4] of composition Cu_2AlMn, are ferromagnetic. Ferromagnetic alloys also

[1] S. KAYA and A. KUSSMANN, *Z. Physik*, **72**, 293 (1931).

[2] Magnetization: A. KUSSMANN, B. SCHARNOW, and A. SCHULZE, *Z. tech. Physik*, **13**, 449 (1932); P. WEISS and R. FORRER, *Compt. rend.*, **189**, 663 (1929).

[3] S. VALENTINER and G. BECKER, *Z. Physik*, **80**, 735 (1933).

[4] F. HEUSLER, *Verh. deut. physik. Ges.*, **5**, 219 (1903).

occur in the chromium-tellurium, manganese-arsenic, and platinum-chromium systems.

4. Ionic Crystals.—The salts produced by combining highly electropositive metals and highly electronegative elements such as the halogens, oxygen, and sulfur are the ideal ionic crystals. Other, more complex salts, such as metal carbonates and nitrates and ammonium halides, also may be classified as ionic crystals. We shall be interested principally in the diatomic types, however, since they are the easiest to handle theoretically.

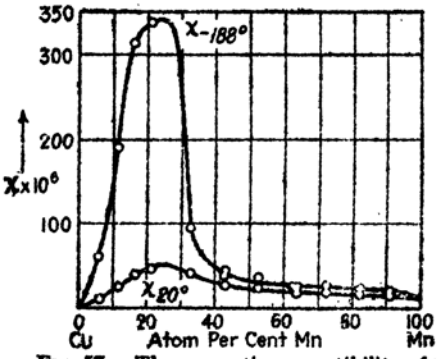

FIG. 57.—The magnetic susceptibility of the copper-manganese system. The abscissa is the atom per cent of manganese. (*After Valentiner and Becker.*)

Ionic crystals closely obey the ordinary rules of classical valency; in fact, most valence numbers are derived from investigations of the combining ratios of atoms in ionic compounds.

a. Cohesion.—The heats of formation[1] of a number of diatomic ionic crystals are given in Table XII. The standard state to which these values are referred is that of the monatomic gases of the constituents. It is noteworthy that the cohesive energy generally is larger for components containing atoms of higher valency than for compounds containing atoms of lower valency.

In many instances, it is convenient to refer the cohesive energies to a standard state of free ions rather than of free atoms. Thus, we shall have occasion to use the energy required to sublime sodium chloride into free Na^+ and Cl^- ions. These energies may be obtained from those of Table XII by adding the energy required to transfer valence electrons from the metal atoms to the electronegative atoms. In compounds of formula MX, this additional term obviously is a multiple of the difference between the ionization energy of the metal atoms and the electron affinity of the electronegative atoms. The first of these quantities has been determined very accurately, by spectroscopic means, for practically all metals. The second, however, has been measured only for the halogens. The most direct method of determining electron affinities has been developed by Mayer[2] and is based upon measurement of the equilibrium

[1] BICHOWSKY and ROSSINI, *op. cit.*

[2] J. E. MAYER, *Z. Physik*, 61, 798 (1930). L. HELMHOLZ and J. E. MAYER, *Jour. Chem. Phys.*, 2, 245 (1934). P. P. SUTTON and J. E. MAYER, *Jour. Chem. Phys.*, 2, 146 (1934); 3, 20 (1935).

density of atomic ions in heated alkali halide vapor. From this quantity, it is possible to determine the heat of the reaction

$$MX \text{ (molecule)} \rightleftarrows M^+ \text{ (atom)} + X^- \text{ (atom)}$$

where M is the metal atom and X is the halogen atom. By subtracting from this the heat of the reaction

$$MX \text{ (molecule)} \rightleftarrows M \text{ (atom)} + X \text{ (atom)}$$

and the ionization energy of the metal atom, one obtains the energy of the process

$$\text{Electron} + X \rightleftarrows X^-$$

TABLE XII.—COHESIVE ENERGIES OF IONIC CRYSTALS RELATIVE TO THE MONATOMIC GASES OF THE CONSTITUENTS

(In kg cal/mol at room temperature)

Alkali Hydrides			
		LiH	112.5
		NaH	91.8
		KH	81.7
		RbH	82.8
		CsH	82.7
Alkali Halides			
LiF	216.4	LiCl	162.6
NaF	193.7	NaCl	153.1
KF	186.1	KCl	153.1
RbF	183.9	RbCl	152.9
CsF	182.2	CsCl	154.0
LiBr	149.7	LiI	129.7
NaBr	139.5	NaI	120.6
KBr	140.8	KI	124.3
RbBr	141.6	RbI	125.5
CsBr	143.4	CsI	128.3
Other Monovalent Metal Halides			
		CuCl	144.4
		$CuCl_2$	192.4
AgF	148.5	AgCl	127.2
		AuCl	129.2
		TlCl	117.5
CuBr	134.8	CuI	124.6
$CuBr_2$	169.0	CuI_2	137.2
AgBr	118.7	AgI	108.5
AuBr	122.3	AuI	117.4
Alkaline Earth Halides			
		$BeCl_2$	245.4
MgF_2	363.7	$MgCl_2$	247.4
CaF_2	401.6	$CaCl_2$	286.2
SrF_2	389.6	$SrCl_2$	302.7
BaF_2	400.5	$BaCl_2$	312.1

TABLE XII.—COHESIVE ENERGIES OF IONIC CRYSTALS RELATIVE TO THE MONATOMIC GASES OF THE CONSTITUENTS.—(*Continued*)

$BeBr_2$	208.2	BeI_2	165.6
$MgBr_2$	214.0	MgI_2	174.3
$CaBr_2$	263.8	CaI_2	227.5
$SrBr_2$	271.8	SrI_2	234.3
$BaBr_2$	283.2	BaI_2	244.8
	Other Divalent Metal Halides		
$ZnCl_2$	184.8	$ZnBr_2$	159.6
$CdCl_2$	177.6	$CdBr_2$	156.4
$HgCl_2$	125.8	$HgBr_2$	109.1
$PbCl_2$	191.0	$PbBr_2$	167.6
ZnI_2	128.4		
CdI_2	126.4		
HgI_2	91.1		
PbI_2	140.5		
	Miscellaneous Cases		
AlF_3	510	$AlCl_3$	308
SbF_3	352	$TlCl_3$	209
		$SbCl_3$	218
		$SbCl_5$	292
		$SnCl_2$	217
		$FeCl_2$	234
		$FeCl_3$	277
$AlBr_3$	262	AlI_3	209
$SbBr_3$	180	SbI_3	140
$SnBr_2$	193	SnI_2	168
	Alkali Metal Oxides, Sulfides, and Selenides		
Li_2O	279	Na_2S	208
Na_2O	210	K_2S	227
K_2O	185	Rb_2S	192
Cs_2O	179	Cs_2S	191
Li_2Se	224		
Na_2Se	172		
K_2Se	175		
	Other Cases		
BeO	269		
MgO	242	MgS	185
CaO	258	CaS	228
SrO	247	SrS	226
BaO	241	BaS	226
Al_2O_3	667	Al_2S_3	449
TiO_2	436		
CaSe	191		
SrSe	191		
BaSe	191		

Another scheme of the same type, also devised by Mayer, involves the use of halogen vapor instead of alkali halide vapor. Table XIII contains some values of the electron affinities of the halogens that were determined by Mayer, Helmholz, and Sutton, using these methods.

TABLE XIII.—ELECTRON AFFINITY OF HALOGEN ATOMS AS DETERMINED DIRECTLY BY MAYER'S METHODS

(In kg cal/mol)

Cl	88.3
Br	84.2
I	72.4

The electron affinities of the doubly charged negative ions O^{--}, S^{--}, and Se^{--}, etc. probably are negative; hence, these ions are unstable, and therefore their affinities cannot be determined by direct methods.

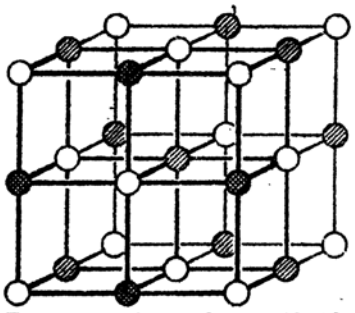

FIG. 58.—The sodium chloride lattice.

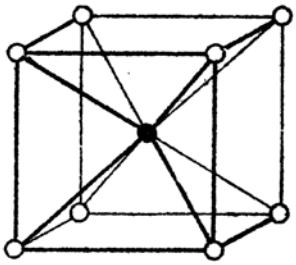

FIG. 59.—The cesium chloride lattice.

b. Crystal Structure.—Ideal ionic compounds usually crystallize in one of several simple structures.[1] The sodium chloride structure of Fig. 58 is characteristic of all the alkali halides[2] except the low-temperature modifications of cesium chloride, bromide, and iodide, which have the simple cubic structure of Fig. 59. Many divalent metal oxides, sulfides, selenides, and tellurides also have the sodium chloride lattice, although many others crystallize in the zincblende and wurtzite structures of Figs. 60 and 61. On the whole, it may be said that all four of these structures are characteristic of ionic crystals in which the constituent atoms have equal positive and negative valences. An interesting feature of these lattices is that they remain the same when metal and electronegative atoms are interchanged.

The fluorite lattice of Fig. 62 is typical of ionic compounds that have the formula M_2X or MX_2. This structure occurs among the alkaline earth halides such as calcium fluoride and barium fluoride and among the

[1] See *Strukturberichte.*

[2] Many of the alkali halides possess the cesium chloride structure at high pressures. See R. B. Jacobs, *Phys. Rev.*, **54**, 468 (1938).

alkali metal oxides and sulfides such as lithium oxide and sulfide. Many other compounds that have these formulae possess structures in which the chemical molecule shows a tendency to form an "island" in the lattice. This is evident, for example, in the rutile or titanium oxide lattice of Fig. 63 which may be regarded as being built of a body-centered arrangement of TiO_2 molecules. Since this behavior is typical of molecular crystals, it may be said that ionic crystals such as titanium oxide are mild transition cases between ionic and molecular types. Another

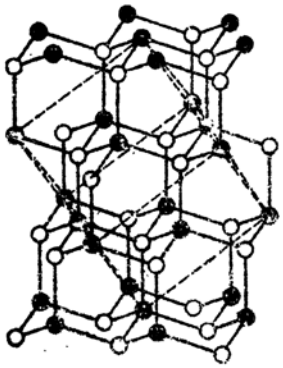

FIG. 60.—The zincblende lattice.

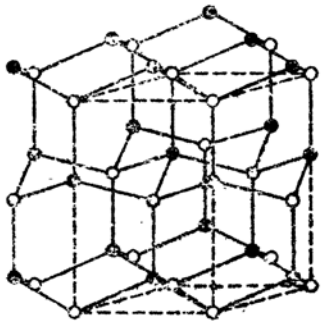

FIG. 61.—The wurtzite lattice.

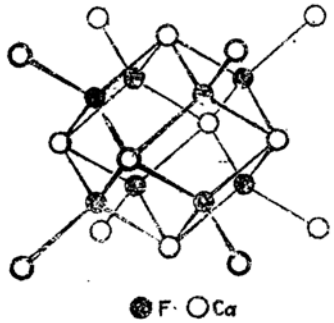

FIG. 62.—The calcium fluoride lattice.

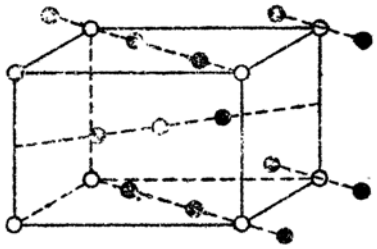

FIG. 63.—The rutile lattice.

simple lattice that shows the same behavior is the pyrites or FeS_2 structure of Fig. 64. This is equivalent to a face-centered cubic arrangement of FeS_2 molecules. Carbon dioxide, which is a typical molecular compound, has a similar lattice.

The α-corundum structure of Fig. 65 is formed by several ionic compounds that have the form A_2B_3, such as Al_2O_3 and Fe_2O_3.

The structures of a number of ionic crystals are listed in Table XIV. This collection[1] also contains a few substances such as AlN and GaP,

[1] See *Strukturberichte*.

which are transition cases but which are of interest because they have typical ionic structures. Several ammonium salts also have been included in order to show that a radical may play the same role as an atomic ion.

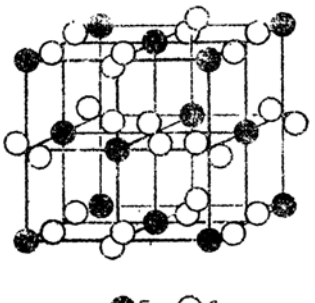

FIG. 64.—The pyrites lattice. Both ferric sulfide and carbon dioxide possess this structure.

FIG. 65.—The α-corundum lattice.

The interatomic distances d of the alkali halides, show a regularity, namely, that the differences between the d values of NaF and KF, NaCl and KCl, NaBr and KBr, and NaI and KI are equal to within a few hundredths of an angstrom unit:

NaF-KF	NaCl-KCl	NaBr-KBr	NaI-KI
0.36 Å	0.33 Å	0.31 Å	0.30 Å

The same type of relationship is valid for the differences between the d values of the halides of other pairs of alkali metals and the d values of the alkali metal salts of pairs of halogens. It follows from this rule that we may associate with each ion a definite radius, the interatomic distance of each substance being given closely by the sum of the radii of the constituent ions. Thus, the crystals behave as though they were composed of rigid spherical ions that are in contact with one another. It obviously is not possible to determine the absolute values of the ionic radii from the d values alone, although it is possible to determine the differences of all the alkali metal ion radii and all the halogen-ion radii. Thus, it is necessary to know the absolute value of only one radius in order to determine all radii.

This additivity principle also is roughly valid for many other classes of ionic crystal, such as the copper and silver halides and the alkaline earth oxides, sulfides, and selenides.

TABLE XIV.—SOME CRYSTAL CONSTANTS OF IONIC SOLIDS

Monovalent Metal Halides

(See Table II, for notation)

The alkali metal hydrides probably have the NaCl type of lattice.

	Type	Parameters, Å		
		a	*c*	*d*
LiF	f.c.c.	4.02	....	2.01
NaF	f.c.c.	4.62	....	2.31
KF	f.c.c.	5.33	....	2.67
RbF	f.c.c.	5.63	....	2.82
CsF	f.c.c.	6.01	....	3.00
LiCl	f.c.c.	5.14	....	2.57
NaCl	f.c.c.	5.63	....	2.81
KCl	f.c.c.	6.28	....	3.14
RbCl	f.c.c.	6.54	....	3.27
CsCl	f.c.c.			
CsCl	s.c.	4.11	....	3.56
LiBr	f.c.c.	5.49	....	2.75
NaBr	f.c.c.	5.96	....	2.98
KBr	f.c.c.	6.58	....	3.29
RbBr	f.c.c.	6.85	....	3.43
CsBr	s.c.	4.29	....	3.71
LiI	f.c.c.	6.00	....	3.00
NaI	f.c.c.	6.46	....	3.23
KI	f.c.c.	7.05	....	3.53
RbI	f.c.c.	7.33	....	3.66
CsI	s.c.	4.56	....	3.95
AgF	f.c.c.	4.92	....	2.46
CuCl	Zincblende	5.41	....	2.34
AgCl	f.c.c.	5.54	....	2.77
CuBr	Zincblende	5.68	....	2.46
AgBr	f.c.c.	5.76	....	2.88
CuI	Zincblende	6.05	....	2.62
α AgI	Wurtzite	4.59	7.53	
β AgI	Zincblende	6.49	....	2.81
TlCl	s.c.	3.84	....	3.33
TlBr	s.c.	3.97	....	3.44
TlI	s.c.	4.18	....	3.62

TABLE XIV.—SOME CRYSTAL CONSTANTS OF IONIC SOLIDS.—(*Continued*)

Alkaline Earth Halides

(Many of these have complex structures which we shall not discuss. The following are a few simple cases.)

	Type	Parameters, Å		
		a	*c*	*d*
MgF_2	Rutile (see Fig. 63)	4.64	3.06	2.05
CaF_2	Fluorite	5.45	....	2.36
BaF_2	Fluorite	6.19	....	2.68
ZnF_2	Rutile	4.72	3.14	2.10
CdF_2	Fluorite	5.40	....	2.34
	Alkali Metal Oxides, Sulfides, and Selenides			
Li_2O	Fluorite	4.61	....	2.00
Li_2S	Fluorite	5.70	....	2.47
Na_2S	Fluorite	6.53	....	2.83
	Monovalent Meta Oxides, Sulfides, Selenides			
Cu_2O	Complex cubic lattice	2.46	....	1.84
Cu_2S	Fluorite	5.59	....	2.42
Cu_2Se	Fluorite	5.75	....	2.49
Ag_2O	Same as Cu_2O	4.70	....	2.05
	Bivalent Metal Oxides, Sulfides, Selenides			
BeO	Wurtzite	2.69	4.37	1.64
MgO	f.c.c.	4.21	....	2.10
CaO	f.c.c.	4.80	....	2.40
SrO	f.c.c.	5.15	....	2.58
BaO	f.c.c.	5.53	....	2.77
ZnO	Wurtzite	3.24	5.18	1.94
CdO	f.c.c.	4.70	....	2.35
BeS	Zincblende	4.86	....	2.10
MgS	f.c.c.	5.19	....	2.60
CaS	f.c.c.	5.68	....	2.84
SrS	f.c.c.	6.01	....	3.01
BaS	f.c.c.	6.37	....	3.19
α ZnS	Zincblende	5.42	....	2.35
β ZnS	Wurtzite	3.84	6.28	2.36
α CdS	Zincblende	5.82	....	2.52
β CdS	Wurtzite	4.14	6.72	2.52
HgS	Zincblende	5.84	....	2.53
BeSe	Zincblende	5.18	....	2.18
MgSe	f.c.c.	5.45	....	2.73
CaSe	f.c.c.	5.91	....	2.96
SrSe	f.c.c.	6.24	....	3.12
BaSe	f.c.c.	6.59	....	3.30

TABLE XIV.—SOME CRYSTAL CONSTANTS OF IONIC SOLIDS.—(*Continued*)

	Type	Parameters, Å		
		a	*c*	*d*
Bivalent Metal Oxides, Sulfides, Selenides				
ZnSe	Zincblende	5.66	....	2.45
CdSe	Zincblende	6.05	....	2.62
HgSe	Zincblende	6.07	....	2.63
MgTe	Wurtzite	4.52	7.33	2.75
CaTe	f.c.c.	5.91	....	2.96
SrTe	f.c.c.	6.55	....	3.33
BaTe	f.c.c.	6.99	....	3.50
ZnTe	Zincblende	6.09	....	2.64
CdTe	Zincblende	6.46	....	2.79
HgTe	Zincblende	6.44	....	2.80

Oxides of Trivalent Metals

α Al_2O_3, Fe_2O_3, α Ga_2O_3 form crystals with the corundum structure of Fig. 65.

Miscellaneous Other Cases

Face-centered Cubic Type

	a	*d*
FeO	4.28	2.14
CoO	4.25	2.13
NiO	4.17	2.08
ScN	4.14	2.20

Simple Cubic Type

	a	*d*
NH_4Cl	3.86	3.34
NH_4Br	4.05	3.51
NH_4I	4.37	3.78

Zincblende Type

	a	*d*
AlP	5.45	2.36
GaP	5.44	2.36

Wurtzite Type

	a	*c*	*d*
NH_4F	4.39	7.02	2.63
AlN	3.11	4.98	1.87

c. Conductivity.—The halides and oxides of the simpler metals generally have an electrolytic conductivity that increases with increasing temperature. Figure 66 shows the specific conductivity of a number of very pure alkali halide crystals as determined by Lehfeldt.[1] The scale of abscissae is adjusted in order to be proportional to $1/T$, and the scale of ordinates is logarithmic. It may be observed that the low-temperature portions of the conductivity curves depend upon the history of the specimen, whereas the high-temperature portions are reproducible straight lines in this

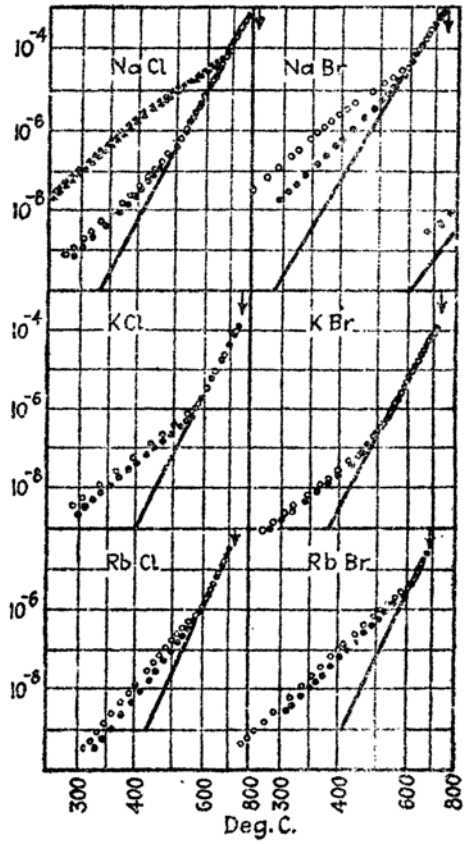

FIG. 66.

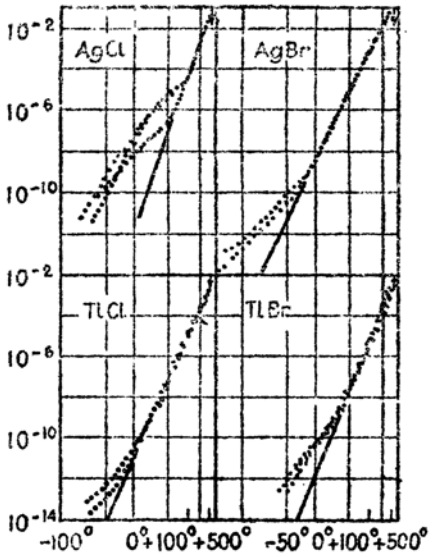

FIG. 67.

FIG. 66.—The ionic conductivity of the alkali halide crystals as a function of temperature. In all cases except that of NaCl the different sets of points refer to two artificial crystals. Additional measurements were made on natural crystals of sodium chloride. The ordinates are expressed in ohm^{-1} cm^{-1}, the abscissas in degrees centigrade. (*After Lehfeldt.*)

FIG. 67.—The conductivity of silver and thallium halides. The ordinates are expressed in ohm^{-1} cm^{-1}. (*After Lehfeldt.*)

type of plot. This fact shows that the high-temperature conductivity σ satisfies the relation

$$\sigma = Ae^{-\frac{\alpha}{T}}$$

where A and α are practically constant.

[1] W. LEHFELDT, *Z. Physik*, **85**, 717 (1933).

Similar curves are shown in Fig. 67 for silver and thallium halides.

It has been demonstrated[1] fairly conclusively that the conductivity of the halides is completely ionic. This fact is by no means obvious from the temperature dependence, for the electronic conductivity of semi-conductors usually follows the same law. Generally speaking, it is possible to establish the existence of ionic conductivity only by performing a number of indirect experiments, which will be discussed in connection with semi-conductors. Oxides, sulfides, and selenides usually exhibit some electronic conductivity.

Table XV gives the fractions n_+ and n_- of the current carried, respectively, by positive and negative ions in a number of halide crystals. These fractions are called the transport numbers of the corresponding ions.

TABLE XV.—THE TRANSPORT NUMBERS OF IONIC SOLIDS AT DIFFERENT TEMPERATURES

Compound	Temperature, °C	n_+	n_-
NaF	500	1.000	0.000
	550	0.996	0.004
	600	0.916	0.084
	625	0.861	0.139
NaCl	400	1.000	0.000
	510	0.981	0.019
	600	0.946	0.054
	625	0.929	0.071
KCl	435	0.956	0.044
	500	0.941	0.059
	550	0.917	0.083
	600	0.884	0.166
AgCl	20–350	1.00	0.00
AgBr	20–200	1.00	0.00
BaF_2	500	0.00	1.00
$BaCl_2$	400–700	0.00	1.00
$BeBr_2$	350–450	0.00	1.00
PbF_2	200	0.00	1.00
$PbCl_2$	200–450	0.00	1.00
$PbBr_2$	250–365	0.00	1.00
PbI_2	255	0.39	0.61
	290	0.67	0.33

[1] See the survey article by C. Tubandt, *Handbuch der Experimental Physik*, Vol. XII, part 1.

d. Specific Heats.—The specific-heat curves of the diatomic salts of simple metals are normal in the sense that they obey Dulong and Petit's law at high temperatures and decrease monotonically with decreasing temperature. Figure 3, Chap. III, shows the atomic-heat curves of

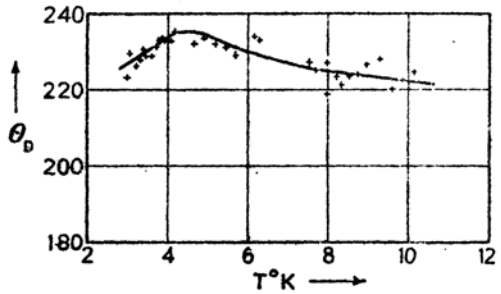

FIG. 68.—Variation of the characteristic temperature of potassium chloride with temperature near absolute zero. The characteristic temperature is defined by the T^3 law in cases of this type (see Sec. 19). (*After Keesom and Clark.*)

several alkali halides above 20°K. Keesom and Clark[1] have found that the specific heat of potassium chloride shows slight deviations from Debye's T^3 law at temperatures below 10°K (*cf.* Fig. 68). We shall discuss this effect in Chap. III.

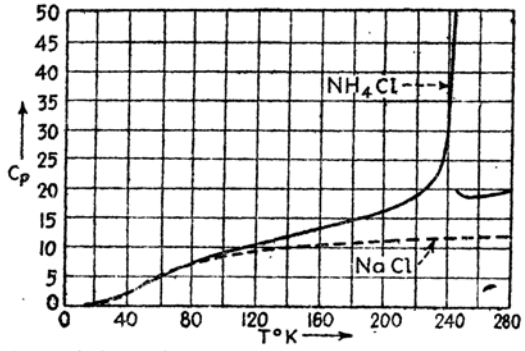

FIG. 69.—A comparison of the molar heats of ammonium chloride and sodium chloride. (*After Simon, v. Simson, and Ruhemann.*)

The specific heats of ammonium halides resemble those of the alkali halides at very low temperatures, but they have large anomalies in the region just below room temperature. Figure 69 gives a comparison of the specific-heat curves of ammonium chloride and sodium chloride[2] and shows the high peak that occurs at 250°K in the first case. As we

[1] KEESOM and CLARK, *Physica*, **2**, 698 (1935).

[2] F. SIMON, O. v. SIMSON, and M. RUHEMANN, *Z. physik. Chem.*, **129**, 339 (1927).

shall see later, this peak is connected with the reorientation of the NH_4^+ radical.[1]

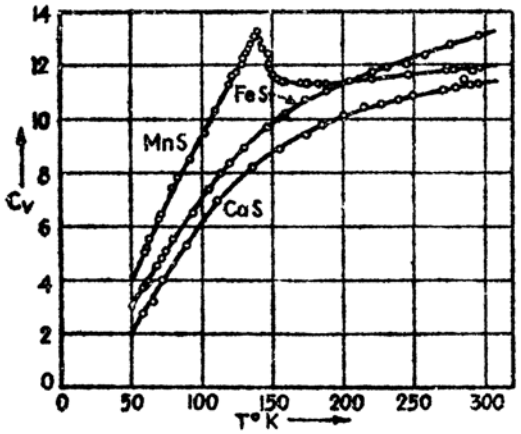

FIG. 70.—The specific-heat curves of manganous sulfide, ferrous sulfide, and calcium sulfide. The ordinate is cal/deg-mol. (*After Anderson.*)

Salts of transition metals usually do not obey Dulong and Petit's law at high temperatures; they show the same type of excess specific heat that is observed in metals that have unfilled *d* shells. Figure 70 shows the curves[2] for manganous sulfide and ferrous sulfide and, for comparison, the "normal" curve of calcium sulfide.

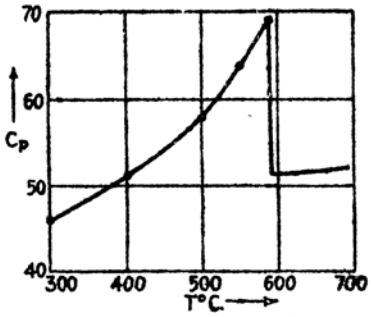

FIG. 71.—The specific-heat curve of magnetite, showing the peak that occurs at the Curie point. The ordinates are cal/deg-mol. (*After Weiss, Piccard and Carrard.*)

Several transition salts are ferromagnetic. Consequently, one might expect their specific-heat curves to have peaks near the ferromagnetic Curie point. Figure 71 shows the peak for magnetite,[3] which seems to be the only case that has been examined.

e. Magnetic Properties.—Most ideal ionic salts are diamagnetic, the exceptions being salts of transition metals, which usually are paramagnetic and sometimes are ferromagnetic.

[1] As a result of careful thermodynamical work in the vicinity of the transition temperature in ammonium chloride, A. Lawson [*Phys. Rev.*, **57**, 417 (1940)] has shown that the anomaly originates in a reorientation of the ammonium radicals rather than in onset of free rotation.

[2] C. T. ANDERSON, *Jour. Am. Chem. Soc.*, **58**, 476 (1931).

[3] P. WEISS, A. PICCARD, and A. CARRARD, *Arch. sci. phys. nat.*, **43**, 113 (1917).

The measured molar susceptibilities of a number of halides,[1] oxides, and sulfides of simple metals are given in Table XVI. The experimental values fluctuate from specimen to specimen, as in simple metals; therefore, these values are not accurate to within more than a few per cent.

It may be verified that the susceptibilities of the alkali halides are additive to a comparatively high degree of accuracy. This fact indicates that each ion preserves a characteristic diamagnetic susceptibility in each compound. However, we cannot determine the absolute susceptibility of the ions from these data alone, just as we could not determine their ionic radii from crystallographic data alone.

The susceptibilities of other halides also seem to be additive, although the results fluctuate so widely in different experimental results tha this statement cannot be made with certainty.

The copper salts furnish an interesting example. The cuprous salts invariably are diamagnetic; however, cupric chloride and iodide are

TABLE XVI
Molar Susceptibilities of Salts of Nontransition Elements
(The unit used is 10^6 times the cgs unit.)

	F	Cl	Br	I
Li	.....	− 25.4	− 37.3	− 55.8
Na	−19.6	− 30.8	− 43.2	− 60.3
K	−25.7	− 36.2	− 49.2	− 67.2
Rb	−31.9	− 46.4	− 56.7	− 67.1
Cs	.....	.....	.....	− 92.5
CuX	.....	− 38.1		
CuX_2	.....	+243.	629.	
AgX	.....	− 54.8	− 48.6	
AuX	.....	− 67.0	− 60.8	− 91.0
TeX	.....	− 58.2		
MgX_2	.....	− 49.7	− 72.2	−111.4
CaX_2	−23.4	− 54.5		
SnX_2	.....	− 63.0	− 85.3	−131.1
BaX_2	−22.8	− 74.0	−103.6	−160
ZnX_2	.....	− 58.2	.....	−112.7
CdX_2	.....	.....	.....	−134.1
HgX_2	.....	− 80.9		
SnX_4	.....	− 114.1		

[1] See, for example, the compilations in Landolt-Bornstein and the International Critical Tables.

TABLE XVI.—(*Continued*)
Molar Susceptibilities of Transition-metal Salts
(Room-temperature values)

	$\chi_m \cdot 10^2$
$FeCl_2$	1.32
$FeBr_2$	1.36
FeI_2	1.36
$FeSO_4$	1.24
$CoCl_2$	1.22
$CoBr_2$	1.27
CoI_2	1.07
$CoSe_4$	1.05
CoO	10^5
$NiCl_2$	0.62
$NiBr_2$	0.55
NiI_2	0.38
NiO	10^4
$PtCl_3$	0.0136
CeF_3	0.220
$CeCl_3$	0.192
Ce_2S_3	0.492
Sm_2S_3	0.322
$Sm_2(SO_4)_3$	0.211

paramagnetic. This fact shows that the cuprous ion is a simple ion whereas cupric ion is similar to the ions of transition metals.

The molar susceptibilities of a number of transition-metal salts also are listed in Table XVI. We shall describe the magnetic properties of some of these salts more fully in Chap. XVI.

The large susceptibilities of cobaltous oxide and nickel oxide suggest that these compounds are ferromagnetic. However, there do not seem to be any measurements on the magnetization curves of these substances. Magnetite, Fe_3O_4, and pyrrhotite Fe_7S_8, are the only salts that are ferromagnetic at room temperature and that have had their ferromagnetic properties measured.[1] The specific-heat curve of magnetite is shown in Fig. 71.

5. Valence Crystals. General Description.—Ideal valence crystals are monatomic nonconducting substances that have high cohesive energies and great hardness. Diamond is the prototype of this class, just as the alkali halides are the prototypes of ionic crystals. A characteristic of the diamond structure, which is shown in Fig. 4, is that the number of

[1] P. WEISS, *Jour. de Phys.*, 6, 661 (1907). M. ZIEGLER, Thesis (Zurich, 1915). See also D. R. INGLIS, *Phys. Rev.*, 45, 119 (1934).

nearest neighbors of each atom, namely, four, is equal to the ordinary valence of carbon. Table XVII contains some data[1] for diamond and other valence crystals. Boron is possibly another ideal valence crystal since it is also a very hard insulator; however, its structure does not seem to be known.

TABLE XVII.—SOME PROPERTIES OF SOLIDS HAVING VALENCE CHARACTERISTICS

Substance	Structure	Cohesive energy, kg cal/mol	Hardness (relative)	Resistivity, ohm-cm	Magnetic susceptibility $\cdot 10^6$
Diamond.....	Fig. 4 $d = 1.54$ Å	$(170.0 + a)$*	10	10^{14}	−0.50
Graphite.....	Fig. 72 $d = 1.42$	$(170.49 + a)$*	0.5	$2 \cdot 10^{-5}$ at 0°C Decreases with decreasing temperature	−3.5
Boron.....	.....	115	9.5	10^{18}	−0.7
Silicon.....	Diamond type $d = 2.35$	85	7	$8 \cdot 10^{-2}$	−0.13
Germanium...	Diamond type $d = 2.43$	85	....	$9 \cdot 10^{-2}$	−0.10
Gray tin.....	Diamond type $d = 2.80$	78.6	....	.....	−0.35
Silicon carbide	ZnS type $d = 1.89$, etc.	283	9		
Silicon dioxide	See text	405.7	7	10^{14}	−0.45
Boron nitride.	Fig. 72 $d = 1.45$	.....	....	.....	0.0

* Two different methods of determining the heat of sublimation of graphite lead to values that differ by about 50 kg cal/mol.

Many substances may be classified between the valence type and one of the other types. For example, silicon, germanium, and gray tin crystallize in the diamond structure although they ordinarily have a much higher conductivity than diamond. Silicon and germanium may also be classified among semi-conductors, whereas gray tin may be classified among metals. Similarly, silicon carbide and silicon dioxide have some valence properties, such as great hardness, and some ionic characteristics, such as the ability to absorb infrared radiation strongly.

Silicon carbide crystallizes in several different lattice structures, which have in common the property that each atom is surrounded by four atoms of opposite type that are situated at the corners of a tetrahedron. One

[1] See, for example, the compilations of Landolt-Bornstein and the International Critical Tables.

form, for example, corresponds to the zincblende structure of Fig. 60. Similarly, silicon dioxide has several crystalline forms. In each of these, a silicon atom is surrounded tetrahedrally by four oxygen atoms and each oxygen atom is joined to two silicon atoms.

The most stable solid form of carbon probably is graphite, which has a layer lattice structure similar to that of boron nitride, shown in Fig. 72. This substance is not hard in spite of its high cohesive energy (*cf.* Table XVII), presumably because the planes of carbon atoms slide easily over one another. Nevertheless, we shall classify graphite among the valence compounds since the forces between carbon atoms in the planes of graphite are believed to be similar to the forces between carbon atoms in diamond. It may be seen from Table XVII, that graphite has a large conductivity which increases with decreasing temperature. Since this conductivity is electronic, graphite may also be classified among metals.

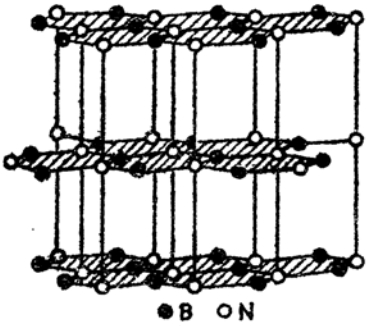

FIG. 72.—The lattice of boron nitride.

Let us consider the following two sequences of compounds:

LiF	NaF
BeO	MgO
BN	AlN.

In each of these sequences the valences of the electropositive and electronegative elements increase by unity as we move down a given column. The top members are ideal ionic compounds. The second members exhibit ionic conductivity and crystallize in typical ionic structure but are very hard. Boron nitride has the structure shown in Fig. 72, which is similar to that of graphite, whereas aluminum nitride has the wurtzite structure, which is similar to that of diamond. Evidently the properties of a compound of light elements become more nearly like those of typical valence crystals the nearer the center of the periodic chart its constituents lie. This general rule is obeyed by many compounds of the lighter elements, the principal exceptions being molecular compounds.

In this connection, it may be recalled that the metals arsenic, antimony, and bismuth crystallize in a layer lattice structure in which the number of nearest neighbors of a given atom is equal to the electronegative valency of these elements. This behavior, being analogous to that of ideal valence crystals, indicates that these substances should be classified between the metallic and the valence types.

6. Semi-conductors. *a. General Properties.*—A number of solids have a small electronic conductivity that is negligible at very low tem-

peratures and increases with increasing temperature. The experiments that distinguish this conductivity from ionic conductivity will be described in the next section. The electrical properties of these semi-conductors are so unique that it is convenient to introduce a separate classification for the group, even though it would be possible to place them among the other types. It will become apparent that this class does not possess the same degree of unity as the other four classes of solids.

Table XVIII contains a list of established semi-conductors and several substances that probably are semi-conductors. Most of these solids crystallize in the ionic type of structure and were discussed with ideal ionic crystals. Carborundum, on the other hand, was previously included under valence types. The semi-conducting specimens of most monatomic substances, such as silicon and tellurium, are usually impure.

TABLE XVIII

Semi-conductors

Monatomic Substances

Si(impure)

Te

Halides

AgI

CuI

Oxides

CuO	NiO	Cr_2O_3
ZnO	FeO	Fe_2O_3
BaO	WO_2	Fe_3O_4
CoO	UO_2	CU_2O

Sulfides and Selenides

PbS	
Ag_2S	Ag_2Se
CdS	
MoS_2	

Probable Semi-conductors

SiC(impure)

Ag_2Te

There are two methods of measuring the conductivities of semi-conductors. The first of these, which is used more commonly,[1] consists in placing a single crystal or a pressed powder specimen between two metal electrodes and measuring its resistance by some ordinary means such as a Wheatstone bridge. This direct method has a number of advantages; for example, the specimen may be heated or cooled easily, and it may be placed in any kind of atmosphere. Its main disadvantage is that contact resistance between the electrodes and the specimen, or

[1] TUBANDT, *op. cit.* See also the survey article on semi-conductors by B. Gudden, *Ergebnisse exakt. Natur.*, **13**, 223 (1934).

between granules of the powder, may affect the current-resistance curve. In some cases, these effects cause an apparent deviation from Ohm's law; in other cases, they simply give rise to a spurious value of the conductivity. For these reasons, it is always difficult to be certain that the conductivities obtained by the method actually are constants of the material under investigation.

An alternative method has been developed and employed by Gudden[1] and his coworkers Völkl and Guillery. They mix a quantity of the powdered semi-conducting material with a nonconducting dielectric, such as a heavy oil, and use the mixture as a dielectric medium in a condenser. The electrical conductivity of the semi-conductor is determined from an

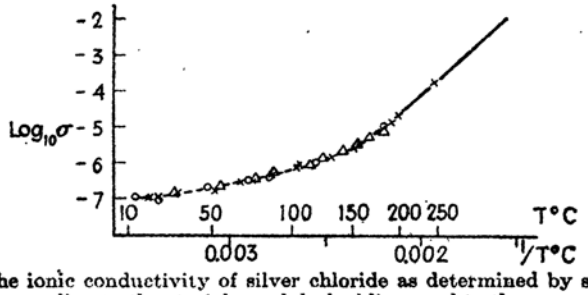

FIG. 73.—The ionic conductivity of silver chloride as determined by several methods. The values corresponding to the straight and dashed lines and to the crosses were measured by direct means. The values corresponding to the circles and triangles were measured by Völkl using the method described in the text.

investigation of the effective resistance of the condenser when it is part of a resonating high-frequency circuit. This procedure has the advantage that it eliminates contact resistance, for the current simply surges back and forth within the granules during the experiment. The principal disadvantages of the method are: (1) It does not allow a very wide choice of conditions under which measurements may be made. (2) It does not lead to very accurate results, since the experimental error usually is of the order of 10 per cent.

Results obtained by the two different methods agree in some cases and disagree widely in others. Guillery found, for example, that the pressed-powder and condenser methods give nearly identical results for cuprous oxides, but widely different ones for stannic oxide and silicon carbide.

Figure 73 shows the conductivity of silver chloride, which is an ionic conductor, as determined by a number of workers,[2] using different methods (see legend to Fig. 73). In this case, the logarithm of the low-temperature conductivity is not linear when plotted as a function of $1/T$ (*cf.* Sec. 4).

[1] A. VÖLKL, *Ann. Physik*, **14**, 193 (1932); P. GUILLERY, *ibid.*, **14**, 216 (1932).
[2] VÖLKL, *op. cit.*

Figure 74 shows the temperature dependence of the conductivity of cuprous oxide[1] as measured by the direct method. The conductivities of a number of semi-conductors give linear plots of this type, a fact showing that the conductivity varies with temperature in the manner

$$\sigma = Ae^{-\frac{E}{kT}}$$

where E and A are practically constant for a given specimen. Meyer[2] has shown that A is about 1 ohm^{-1} cm^{-1} for many semi-conductors, whereas E varies considerably for different substances and for different specimens of the same substance. For example, values of E between 0.06 and 0.6 ev have been quoted for cuprous oxide.

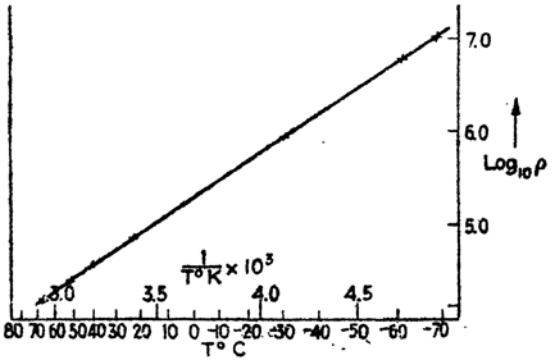

FIG. 74.—The resistivity of cuprous oxide as a function of temperature. The ordinates are ohm-cm. (*After Vogt.*)

b. Methods of Determining the Type of Conductivity.—One or more of the following three quantities are commonly measured in trying to determine whether or not the conductivity of a substance is electronic:

1. Transport numbers.
2. The Hall constant.
3. The thermoelectric coefficient.

We shall discuss the guiding principles that are used in each case.

1. *Measurement of transport numbers.*—An important characteristic of ionic conductivity is that electrolysis accompanies the flow of current, since this electrolysis should be absent in a substance the conductivity of which is entirely electronic. There is some electrolysis if the conductivity is partly ionic and partly electronic, but Faraday's transport law should not be valid in this case. Hence, in principle at least, one should be able to determine the fraction of electronic current by measur-

[1] W. VOGT, *Ann. Physik*, **7**, 190 (1930).
[2] W. MEYER, Thesis (Berlin, 1936); *Z. Physik*, **85**, 278 (1933).

ing the deviation from Faraday's law. Tubandt[1] and his coworkers have employed this method of determining semi-conductors with a great deal of success. Their procedure varies somewhat from case to case, but the underlying principles may be understood by considering the following hypothetical example.

Suppose it is suspected that a substance of formula MX, in which M is the metallic constituent and X is the electronegative constituent, conducts by both positive ions M^+ and electrons. In this case, Tubandt would place three pressed-powder or crystalline disks of the material in series with three disks of a substance MY, which is known to conduct only by means of M^+ ions, and would place these between two electrodes of the metal M (see Fig. 75). The disks adjoining the electrodes and the electrodes are weighed accurately. A current then is sent through the system, and the total quantity of electricity that passes is measured by means of a coulometer. If there is any electrolysis, the disks in contact with the electrodes usually become fastened to the electrodes during this procedure. The electrodes and the disks attached to them are weighed together in order to determine the amount of material lost by the anode and gained by the cathode. The positive-ion transport number may then be computed from all these measured quantities.

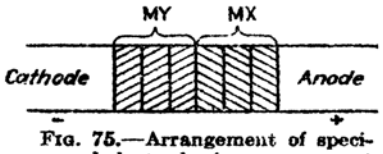

FIG. 75.—Arrangement of specimens and electrodes in measurement of transport numbers. In this case the specimen MX is a positive-ion conductor whose transport number is unknown, whereas MY is a conductor in which all of the current is carried by positive ions.

Tubandt uses the known ionic conductor MY in the circuit, partly to check the coulometer measurement and partly to make certain that negative ions do not leave the disk nearest the cathode. Three disks of each substance are employed in order that one disk may be in contact only with chemically similar substances. If the experiment is flawless, the weight of this disk should not change. A negative transport number can be determined in a similar way by placing three disks of a substance NX, which conducts only by negative ions X^-, between the anode and the three disks of the substance under test.

In practically all cases, this direct method yields results that agree with those determined by other methods. A set of transport numbers that Tubandt and other workers have obtained in this way and that are generally accepted at the present time are given in Table XIX. It should be noted that several semi-conductors have a small, but finite ionic conductivity.

[1] See footnote 1, p. 56.

Tubandt's procedure seems to lead to incorrect results in the case of α silver sulfide, which we shall discuss briefly. In this case, Tubandt employed the scheme symbolized by Fig. 75, making the disks MX of silver sulfide, the disks MY of silver chloride, and the electrodes of silver. Since the α phase is stable only above 180°C, the system was maintained at an elevated temperature during the experiment. Tubandt found that the silver anode lost as much weight as the cathode gained, and he concluded that positive ions carry all of the current. This conclusion was contradicted by a large amount of subsequent work that indicated that the conductivity is mainly electronic. For example, the conductivity of silver sulfide at 180°C is about fifty times larger than that of any

TABLE XIX.—TRANSPORT NUMBERS OF SEMI-CONDUCTORS
(n_e = electronic transport number; n_+ = positive-ion transport number)

Substance	Temperature, °C	n_e	n_+
PbS	240	1.00	
α Ag_2S	Above 180	~0.99	~0.01
β Ag_2S	20	0.015	0.985
β Ag_2Se	20	~1.00	<0.01
β Ag_2Te	20	~1.00	<0.01
γ CuI	200	~1.00	$\sim 2.7 \cdot 10^{-6}$
	325	0.50	0.50
	400	0.00	1.00
Cu_2O	800	~1.00	$\sim 2 \cdot 10^{-4}$
	1000	~1.00	$\sim 5 \cdot 10^{-4}$

ionic conductor, and the Hall coefficient (*cf.* 2 below) is of the magnitude ordinarily associated with electronic conductivity.

At present, this contradiction is explained[1] in the following way. It is assumed that most of the conductivity of α silver sulfide is electronic, although a very small fraction is ionic. The halogen atoms that are ased at the boundary between the silver chloride and silver sulfide s by electrolysis of silver chloride reduce some of the sulfide and ? an equivalent amount of sulfur. This sulfur in turn reacts with an alent number of silver atoms that diffuse from the anode through sulfide. This picture has many direct supports. For example, it is d that the anode does not lose an equivalent of weight if a very large nt is passed through the silver sulfide. In this case, the diffusion utral silver atoms is not rapid enough to change all the liberated to sulfide.

[1] work on silver sulfide is discussed by C. Wagner, *Z. physik. Chem.*, B, 21, 3, 469 (1933).

2. *Measurement of the Hall constant.*—In 1879, Hall[1] found that an emf may be produced across a strip of metal which is carrying a current by placing the strip in a magnetic field. For a cubic crystal, the direction and magnitude of the induced electrostatic field are given by the vector relationship

$$\mathsf{E} = R\mathsf{J} \times \mathsf{H}. \tag{1}$$

Here, E is the electrostatic field vector, J is the current per unit area, H is the magnetic field intensity, and R is the Hall constant of the material. The effect usually is measured by passing a large current I through a thin strip of thickness t that is placed in the magnetic field in such a way that H is normal to the surface. In this case, the emf E_H is

$$\mathsf{E}_H = R\frac{I\mathsf{H}}{t}.$$

In Sec. 37, we shall discuss a simple theory of the Hall effect in which the conductor is treated like a gas of free electrons. The results of this theory may be summarized by saying that they relate the Hall constant to the number of electrons per unit volume n and the value of the electronic charge $-e$ by means of the equation

$$R = -\frac{3\pi}{8nec}$$

in which c is the velocity of light. The sign of the charge on the carriers is thus the same as the sign of R. Strangely enough, this sign is positive for a number of metals, such as zinc and antimony, although their conductivity is undoubtedly electronic. The interpretation of this anomaly is one of the striking successes of the zone theory of solids, which will be developed in later chapters. In order of magnitude, the mobility[2] of the current carriers is given by the ratio of the Hall constant to the specific resistivity, which, in the vicinity of room temperature, is about 100 cm²/volt-sec for most metals. These mobilities are about one hundred times larger than the mobilities of ions in the best solid ionic conductors. Incidentally, the Hall effect in ionic conductors is too small to be measured.

[1] E. H. HALL, *Am. Jour. Math.*, **2**, 287 (1879). A survey of early literature is given by L. L. Campbell in *Galvanomagnetic and Thermomagnetic Effects*, (Longman: Green & Company, New York, 1952).

[2] The mobility μ of a current-carrying particle in a conductor is defined as the velocity with which the particle moves when placed in a unit electrostatic field. Thus, the conductivity σ is equal to $ne\mu$ where, as in the equation for the Hall constant, n is the number of conducting particles per unit volume and e is their charge.

Semi-conductors possess a measurable Hall effect, as Baedeker[1] first found in 1909, and the mobility of the carriers turns out to be of the same magnitude as the mobilities of electrons in metals. At the present time, the existence of a measurable Hall effect is accepted as proof that a given substance is a semi-conductor. The apparent sign of the charge of the current carriers is positive for about as many semi-conductors as it is negative. This anomaly is explained by the modern theory of solids in the same way as it is explained for metals.

The sign of the Hall coefficient apparently changes from positive to negative in some specimens of cuprous oxide[2] as the temperature is raised, being zero at about 500°C. The oxide remains an electronic conductor during the transition, however, a fact showing that the absence of a Hall effect does not furnish proof that the conductivity is ionic.

3. *Measurement of the thermoelectric effect.*—A current flows in two wires of different metals, which are joined together to form a closed circuit, if the two junctions are kept at different temperatures. This thermoelectric effect is usually described by giving the emf dE/dT that is developed for each degree difference in the temperatures of the junctions. These *thermoelectric coefficients* are additive in the sense that the value of dE/dT for two metals A and C is equal to the algebraic sum of the values for the metals A and B and the metals B and C. Hence, it is possible to find the value for any pair of metals if the value of each, relative to a standard, is known.

The thermoelectric effect is observed in semi-conductors but not in ionic conductors. For this reason, the effect is used as a test for electronic conductivity in the same way that the Hall effect is used.

c. Factors That Influence the Conductivity.—In addition to temperature, there are three factors that strongly influence the conductivity of a semi-conductor, namely, its impurity content, the mechanical treatment it has received, and the vapor pressures of the gases of its constituent atoms that are maintained in the surrounding atmosphere.

The first two of these influences have not been investigated in a systematic way. It is known, however, that different specimens of most monatomic semi-conductors, such as silicon, selenium, and tellurium, do not have the same conductivity at the same temperature, and it is concluded from this fact that impurities play an important role in determining the conductivity. Similarly, it is found that the conductivity of powders that are prepared from the same material by grinding varies with the amount of grinding, and the conclusion is drawn that

[1] K. BAEDEKER, *Ann. Physik*, **29**, 566 (1909); *Physik. Z.*, **13**, 1080 (1912).

[2] W. SCHOTTKY and F. WAIBEL, *Physik. Z.*, **34**, 858 (1934). See also *Physik. Z.*, **36** (1935) for correction in sign.

mechanical treatment influences the conductivity. It is possible that these phenomena are closely related, that is, that the distribution impurity atoms in the semi-conductor may be partly connected with its mechanical history. Problems such as these await further investigati

A large amount of work has been done on the effect of vapor press upon the conductivity. These investigations have thrown a great d of light upon the origin of electronic conductivity in many of those se conductors which also may be catalogued under the heading of io crystals. The classical example of this type of work is Baedeke experiment[1] on cuprous iodide, in which it was found that the conducti

TABLE XX

Substances in Which the Conductivity Increases with Increasing Vapor Pressure of t Electronegative Atom

Substance	Sign of Hall Coefficient
CuI	+
Cu_2S	
Cu_2O	+
NiO	
UO_2	
FeO	
CoO	

Substances in Which the Conductivity Decreases with Increasing Vapor Pressure of Electronegative Atom

Substance	Sign of Hall Coefficient
Ag_2S	−
ZnO	−
CdO	−

Substances in Which the Conductivity Is Practically Unchanged with Increasing Vc Pressure of the Electronegative Atom

CuO
Co_2O_4
Fe_3O_4
Fe_2O_3

ity of this substance increases with increasing iodine vapor pressu The Hall coefficient, which is positive in this case, decreases at the sa time, a fact indicating that the number of conducting particles increas with increasing iodine vapor pressure. Similar work[2] has been do on the effect of oxygen and sulfur vapor pressure upon the conductiv of oxide and sulfide semi-conductors. In some of these cases, the c ductivity increases with increasing pressure of the electronegati element, and in others it decreases. Table XX contains a list of s stances upon which investigations have been made and summa

[1] BAEDEKER, *op. cit.*

[2] A survey of work prior to 1935 is given by B. Gudden, *Ergebnisse exakt. Na* **13**, 223 (1934). Additional references may be found in *Physik. Z.*, **36**, 717 (19 **36**, 721 (1936).

qualitatively the results of this work. Figures 76 and 77 show the dependence of conductivity[1] upon vapor pressure in the cases of cuprous oxide and cadmium oxide. It may be observed that the conductivity is proportional to a power of the oxygen pressure in both these cases. This type of relationship occurs commonly. It should be mentioned that we have chosen some of the "best" curves that are available in the

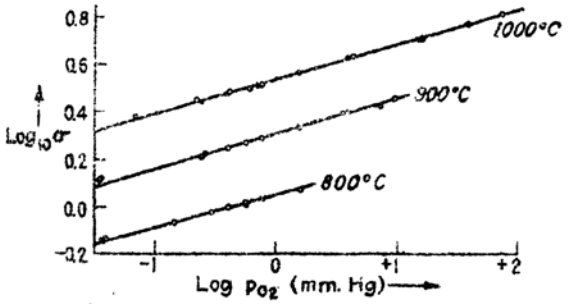

FIG. 76.—The dependence of the conductivity of cuprous oxide upon oxygen pressure. (*After Dünwald and Wagner.*)

literature and that all experimenters do not agree precisely upon the values of the conductivity of a given substance at a given temperature and vapor pressure.

The vapor pressure of the metallic constituent of a semi-conductor usually is not varied in these experiments. Hilsch, Pohl,[2] and their coworkers, however, have placed alkali halide crystals in an atmosphere of alkali vapor for a long enough period of time to observe changes in conductivity. The crystals become colored and exhibit a feeble electronic

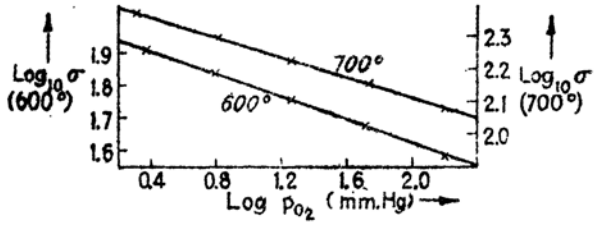

FIG. 77.—The dependence of the conductivity of cadmium oxide upon vapor pressure of oxygen. (*After Baumbach and Wagner.*)

conductivity at the end of this treatment. This shows that even the most ideal ionic crystals may become semi-conductors under suitable conditions.

It should be added that the ionic conductivity of some semi-conductors (*cf.* Table XIX) seems to vary with the pressure of the electro-

[1] H. H. v. BAUMBACH and C. WAGNER, *Z. physik. Chem.*, **B22**, 208 (1933); H. DÜNWALD and C. WAGNER, *ibid.*, **22**, 214 *ff.* (1933).

[2] See the survey article by R. Pohl, *Proc. Phys. Soc.* (*Sup.*), **49**, 3 (1937).

negative gas in the same way as the electronic conductivity. For example, Nagel and Wagner[1] have found evidence to support this in the case of cuprous oxide.

Schottky and Wagner have suggested that there is a correlation between the conductivity of semi-conductors that are sensitive to vapor pressure and the amount by which their composition deviates from ideal stoichiometric proportions. On this basis, they have developed a theory of semi-conductors that will be presented at appropriate places in the following chapters.

7. Molecular Crystals.—Of all the five solid types, we shall be least interested in molecular crystals. Since they are loosely bound aggregates of saturated atoms or molecules, many of their properties are determined primarily by the internal molecular structure, rather than by the solid binding, and thus they are outside the scope of this book. A number of substances that form molecular crystals are listed in Table XXI.[2] The prototypes of the class are the solids of the gaseous elements and of organic compounds that have low boiling points and low heats of sublimation. Several substances, such as sulfur, selenium, tellurium, phosphorus, and iodine, which are transition cases between molecular and valence types, are included as illustrations.

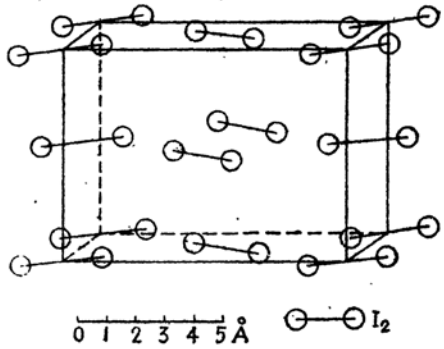

FIG. 78.—The iodine lattice.

The rare gas solids crystallize in the face-centered cubic lattice, as may be seen in Table XXI. Helium forms a true solid only under a pressure of at least 25 atmospheres. The structure of this has not been determined.

Hydrogen, nitrogen, and oxygen have several phases which probably correspond to states in which the diatomic molecules have different relative orientations. All the high-temperature phases apparently are close-packed hexagonal structures of the diatomic molecules.

The solid phases of hydrochloric and hydrobromic acid are face-centered cubic lattices of the diatomic molecules above 98° and 110°K, respectively. Below these temperatures the structures are face-centered tetragonal. It is believed that the molecules are more randomly oriented in the high-temperature forms than in the low-temperature forms.

[1] See footnote 2, p. 70.

[2] See, for example, the compilations of data by Bichowsky and Rossini, *op. cit.*, and *Strukturberichte*.

The halogens, chlorine and iodine, have more complex structures in which the diatomic molecule behaves as a unit. The lattice of iodine is shown in Fig. 78.

Sulfur forms a lattice in which the units are S_8 molecules that have the ring structure shown in Fig. 79. The heat of sublimation relative to free S_8 molecules is about 20 kg cal/mol, a fact indicating that there are fairly large intermolecular forces. Selenium, tellurium, and phosphorus do not have typical molecular structures.

The crystals of organic molecules show a strong tendency to crystallize in simple structures. For example, carbon dioxide and methane

TABLE XXI.—DATA FOR MOLECULAR CRYSTALS

	Heat of sublimation, kg cal/mol	Structure
He	0.052	f.c.c.
Ne	0.52	
A	1.77	
Kr	2.67	
Xe	3.76	
H_2	2.44	There are several phases in each case. Apparently, the high-temperature phases are close-packed hexagonal arrangements of molecules
O_2	1.74	
N_2	1.50	
HCl	4.34	Two phases. The high-temperature phase is f.c.c. —the low-temperature, face-centered tetragonal
HBr	4.79	
Cl_2	6.0	Similar to Fig. 78
I_2	18.9	See Fig. 78
S_8	20	Lattice of S_8 molecules (see Fig. 79)
Se	30	Complex structures similar to those of arsenic, antimony, and bismuth
Te	25	
P	17.7	Complex valencelike structure
NH_3	6.29	Slightly distorted f.c.c. lattice
CO_2	8.24	Same as pyrites (Fig. 64)
CH_4	2.40	f.c.c. lattice

form face-centered cubic lattices of the constituent molecules. The ammonia lattice is very nearly the same, the difference being that the atoms at cube-corner and face-centered positions are slightly displaced relative to one another This tendency toward comparatively simple arrangements extends even to very large molecules.

Practically all molecular crystals are diamagnetic. This is in accord with the fact that the constituent molecules, being saturated, have no resultant spin. Oxygen is an exception, since the normal state of the molecule is triplet. Solid oxygen is strongly paramagnetic.

The specific-heat curves of many molecular crystals have large peaks. The typical cases of methane[1] and hydrochloric[2] acid are shown in Figs. 80 and 81. It is believed that these peaks are associated with changes in degree of molecular orientation.

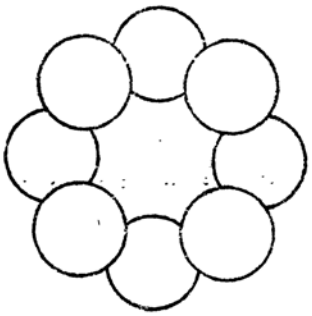

FIG. 79.—The unit ring molecule of sulfur. (*After Warren.*)

8. The Transition between the Solid Types. Figure 82 represents an attempt to show the interrelation of the solid types. The metals are at the left, the two main classifications of these being monatomic metals and alloys. There is necessarily an abrupt transition between these two classes. Valence and ionic types stand next to the right and are in one-to-one correspondence with the monatomic metals and the alloys. The poorly conducting metals, such as bismuth, are transition cases between the ideal monatomic metals and the monatomic valence crystals diamond and boron. Similarly, alloy

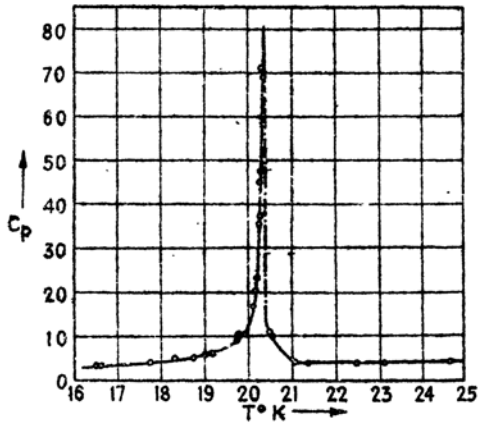

FIG. 80.—The specific heat of methane, showing the large peak at 20.4°K. The ordinates are cal/deg-mol. (*After Clusius and Perlick.*)

systems which have narrow phase boundaries, such as the antimony-magnesium system, are transition substances between alloys and ionic crystals. In the same way, valence crystals, such as quartz

[1] K. CLUSIUS and A. PERLICK, *Z. physik. Chem.*, **24**, 313 (1924).

[2] A. EUCKEN and C. KARWAT, *Z. physik. Chem.*, **112**, 467 (1924). W. F. GIAUQUE and R. WIEBE, *Jour. Am. Chem. Soc.*, **50**, 2193 (1928); **51**, 1441 (1929).

and carborundum, that have strong polar characteristics should be classed between ideal valence and ionic types. The molecular crystals are on the far right. The transition cases between these and valence and ionic crystals are substances such as sulfur and titanium oxide

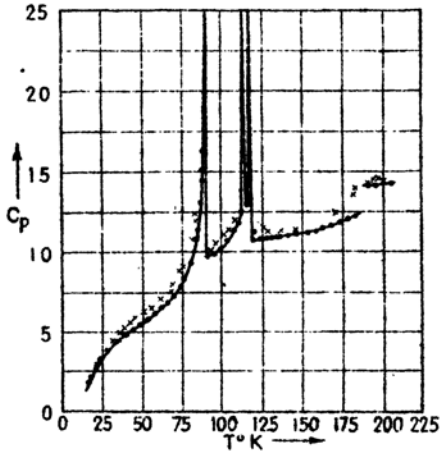

FIG. 81.—The specific-heat curve of hydrogen chloride. The ordinates are cal/deg-mol. (*After Giauque and Wiebe, and Eucken and Karwat.*)

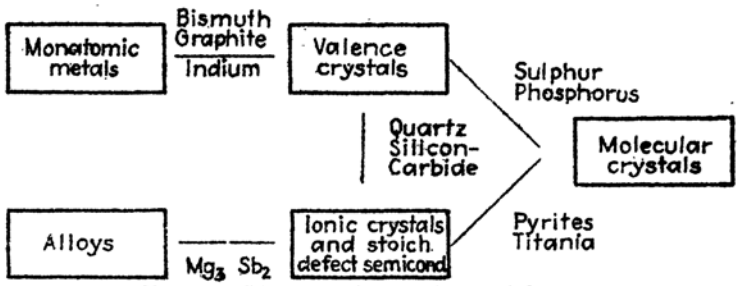

FIG. 82.—The interrelation of the solid types.

or pyrites, respectively, which are bound more tightly than molecular crystals but which show molecular coordination between atoms.

The several different types of semi-conductor cannot be fitted into this chart as a unit. Stoichiometrical defect or excess semi-conductors, such as zinc oxide, may be classed as ionic crystals. On the other hand, impurity semi-conductors, such as selenium, are transition solids between valence and molecular types that contain foreign atoms.

CHAPTER II

THE CLASSICAL THEORY OF IONIC CRYSTALS

9. Introduction.—The foundations of the classical theory of ionic crystals were laid about a quarter of a century ago by Madelung[1] and Born.[2] The basic concept of the theory is that the constituents of ionic crystals are positively charged metal atom ions and negatively charged electronegative atom ions. It is assumed that these ions are spherically symmetrical and that they interact with each other according to simple central force laws. The main interaction, according to the theory, is the ordinary electrostatic, or coulomb, force between the ions, which accounts for the large cohesive energies of the crystals. The electrostatic forces, which tend to contract the dimensions of the crystal, are balanced by repulsive forces which, from the classical viewpoint, have uncertain origin and which vary much more rapidly with interionic distance than do the coulomb forces between charges. Additional interactions are considered in the process of refinement and will be discussed later.

The repulsive term usually is chosen as a function of interionic distance that contains two adjustable parameters which are usually determined empirically by making the expression for the total energy satisfy the following two relations:

$$\left(\frac{dE}{dV}\right)_{V=V_0} = 0 \qquad \text{and} \qquad \left(\frac{d^2E}{dV^2}\right)_{V=V_0} = \frac{1}{V\beta}. \tag{1}$$

Here E is the energy of the crystal, V is its volume, and β is its compressibility.[3] These equations evidently express the conditions that the crystal should be at equilibrium under all forces and that the theoretical compressibility should be equal to the observed value.

As we shall see, the theory is remarkably successful in correlating many of the properties of ionic crystals. From a historical point of view, it may be said to form the basis for a quantitative understanding

[1] E. MADELUNG, *Gött. Nach.*, 100 (1909), 43 (1910); *Physik. Z.*, **11**, 898 (1910).

[2] A survey of Born's work may be found in the *Handbuch der Physik*, Vol. XXIV/2.

[3] The compressibility is defined by the relation

$$\beta = -\frac{1}{V}\left(\frac{\partial V}{\partial p}\right)_T.$$

Since $p = -\partial E/\partial V$, it follows that $1/V\beta = \partial^2 E/\partial V^2$.

of all solids, since it helped to distinguish between those facts which can be understood in terms of classical theory and those which cannot. We shall consider the theory here partly for this reason, and partly for the reason that the mathematical technique employed in it is of great value in more modern developments.

We shall now discuss in detail the various interaction terms employed in the classical theory. These will be considered under three headings: (1) electrostatic interaction, (2) repulsive interaction, and (3) multipole interaction.

10. Electrostatic Interaction Energy.—It is assumed in the classical ionic theory that the ions have charges corresponding to their normal chemical valence. Thus, sodium ions and chlorine ions in sodium chloride have, respectively, one electronic unit of positive and of negative charge, whereas magnesium ions and oxygen ions in magnesium oxide have two electronic units of positive and negative charge. According to electrostatic theory, the interaction energy of two nonoverlapping spherically symmetric charge distributions is

$$\frac{e_1 e_2}{r_{12}} \tag{1}$$

where e_1 and e_2 are the total charges on the distributions and r_{12} is the distance between their centers. Similarly, the total electrostatic energy E_e of n such charges of magnitude e_i ($i = 1, \cdots, n$) is

$$E_e = \sum_{\text{pairs}} \frac{e_i e_j}{r_{ij}} \tag{2}$$

in which the summation extends over all pairs of charges, each pair being considered once. This also may be written in the form

$$E_e = \frac{1}{2} \sum_{i,j}{}' \frac{e_i e_j}{r_{ij}}, \tag{3}$$

where the summation is now a double sum over all charges and the superscript prime indicates that the cases $i = j$ are to be excluded. This convention will be used throughout this volume.

A detailed discussion of Ewald's[1] method of evaluating sums of type (3) for charges that are distributed in crystalline array may be found in the references of the footnote below. For two-atom crystals such as sodium chloride, cesium chloride, zinc sulfide, calcium fluoride, and aluminum oxide, the results may always be expressed in the simple form

$$E_e = -N_A A_i \frac{e_+ e_-}{l}. \tag{4}$$

[1] P. P. EWALD, Thesis (Munich, 1912); see also footnote 2, p. 76.

Here, E_e is the energy per mol; e_+ and e_- are the absolute values of the charges on the positive and negative ions, respectively; l is one of the characteristic crystal dimensions, such as the length of the cube edge, in cubic crystals, or the cation-anion distance; N_A is the Avogadro number; and A_l is the Madelung constant, which is characteristic of the type of crystal structure and is independent of the dimensions of the lattice. The numerical value of the Madelung constant evidently depends on the nature of the parameter l and on the units of charge, length, and energy that are used. Values of A_l are quoted in Table XXII for the types of ionic crystals that were discussed in Sec. 4. These values are given for three different choices of l: (1) when l is the closest cation-anion distance r_0; (2) when l is the cube root of the molecular volume δ_0; (3) when l is the length a_0 of the cube edge. The last case is significant only for cubic crystals.

TABLE XXII.—MADELUNG CONSTANTS FOR SOME TYPICAL IONIC CRYSTALS

(The values in the three columns correspond to three different choices of lattice parameter:
r_0 cation-anion distance
δ_0 cube root of molecular volume
a_0 cube-edge length in cubic crystals.)

	A_{r_0}	A_{δ_0}	A_{a_0}
Sodium chloride	1.7476	2.2018	3.4951
Cesium chloride	1.7627	2.0354	2.0354
Zincblende (ZnS)	1.6381	2.3831	3.7829
Wurtzite (ZnS)	1.641	2.386	
Fluorite (CaF_2)	5.0388	7.3306	11.6366
Cuprite (CuO_2)	4.1155	6.5436	9.5044
Rutile (TiO_2)	4.816	7.70	
Anatase (TiO_2)	4.800	8.04	
Corundum (Al_2O_3)	25.0312	45.825	

It follows from Eq. (2) that the zero of energy is chosen in such a way that E_e vanishes when the ions are infinitely separated. Thus, the standard state is one of free ions. In the case of the alkali halides, $-E_e$ is about 10 per cent larger than the observed values of the cohesive energies U, referred to this standard state. For example, we have in the cases of sodium and cesium chloride

	E_e, kg cal/mol	U (obs.)
NaCl	205	182.8
CsCl	164	155.0

According to the classical picture, the energy associated with repulsive forces should account for the difference.

1.. The Repulsive Term.—The repulsive force between ions is very small until the ions come into contact, whereupon it increases more rapidly than the electrostatic force. Two different forms of the repulsive interaction term have been considered in the course of the development of the theory. The first, which has no independent foundation, was introduced long before the advent of modern quantum theory; the second was introduced as a result of quantum theoretical treatments of ionic interaction. We shall now discuss both of these terms.

a. The b/r^n Interaction.—In his earlier work on ionic crystals, Born[1] assumed that the repulsive forces between ions gave rise to an interaction energy of the type

$$E_r = \frac{b}{r^n} \tag{1}$$

for the entire crystal. Here, b and n are constants and r is the distance between nearest unlike ions. If we assume that only nearest neighboring ions contribute to (1), this term implies that ions repel each other with a central force that varies as $1/r^{n+1}$.

In addition, Born assumed that the total energy of the crystal is simply the sum of the term (1) and an electrostatic term of the type (5) of the previous section. He then determined n and b by use of Eqs. (1) of the preceding section. We shall let E designate the energy of a mol of substance and shall express the molar volume V in terms of r by means of the equation

$$V = N_A \alpha r^3 \tag{2}$$

in which α is a constant that is characteristic of the type of lattice. Thus, the first two derivatives of E with respect to V are

$$\frac{dE}{dV} = \frac{1}{3N_A\alpha r^2}\frac{dE}{dr}, \tag{3}$$

$$\frac{d^2E}{dV^2} = \frac{1}{9N_A^2\alpha^2 r^2}\frac{d}{dr}\left(\frac{1}{r^2}\frac{dE}{dr}\right). \tag{4}$$

Combining Eq. (1) with Eq. (4) of the previous section and solving Eq. (1) of Sec. 9 for b and n with the aid of Eqs. (3) and (4), we obtain

$$b = \frac{N_A A_{r_0} e_+ e_-}{n} r_0^{n-1}, \tag{5}$$

$$n = 1 + \frac{9\alpha r_0^4}{\beta e_+ e_- A_{r_0}}, \tag{6}$$

[1] See footnote 2, p. 76.

$$E = -\frac{N_4 A}{r_0}\frac{n e_+ e_-}{}\left(1 - \frac{1}{n}\right). \tag{7}$$

The values of n that are determined from Eq. (6) by use of experimental values of β generally lie in the vicinity of 9. They are somewhat smaller

TABLE XXIII.—VALUES OF n DERIVED FROM COMPRESSIBILITY MEASUREMENTS

Substance	n
LiF	5.9
LiCl	8.0
LiBr	8.7
NaCl	9.1
NaBr	9.5

TABLE XXIV.—THEORETICAL ENERGIES OF ALKALI HALIDE CRYSTALS DETERMINED BY USE OF THE b/r^n REPULSIVE POTENTIAL

Salt	n	U (calc.), kg cal/mol	U (exp.), kg cal/mol	Affinity, kg cal/mol
LiF	6.0	240.1	.....	100.1
NaF	7.0	215.0	.....	96.7
KF	8.0	190.4	.....	95.3
RbF	8.5	181.8	.....	98.0
CsF	9.5	172.8	.....	98.8
LiCl	7.0	193.3	198.1	
NaCl	8.0	180.4	182.8	
KCl	9.0	164.4	164.4	
RbCl	9.5	158.9	160.5	
CsCl	10.5	148.9	155.1	
LiBr	7.5	183.1	189.3	
NaBr	8.5	171.7	173.3	
KBr	9.5	157.8	156.2	
RbBr	10.0	152.5	153.3	
CsBr	11.0	143.5	148.6	
LiI	8.5	170.7	181.1	
NaI	9.5	160.8	166.4	
KI	10.5	149.0	151.5	
RbI	11.0	144.2	149.0	
CsI	12.0	136.1	145.3	

than 9 for salts involving light ions and somewhat larger for salts involving heavy ions. Table XXIII contains several values that Slater[1] derived, using compressibilities that he obtained from his own measurements by extrapolation to absolute zero of temperature.

[1] J. C. SLATER, *Phys. Rev.*, **23**, 488 (1924).

Various workers have employed values of n that differ slightly from those derived from compressibilities. Thus, Lennard-Jones[1] has determined values from an investigation of the equation of state of rare gases, and Pauling[2] has used values obtained by rules that he derived from an approximate theoretical treatment of the interaction of closed-shell electronic configurations. It is clear from Eq. (7) that a change of n by unity affects E by 1 or 2 per cent. Since the various choices of n for a given substance differ by no more than 1, their differences are of minor consequence as far as the total energy is concerned.

Table XXIV contains a list of observed and calculated values of cohesive energies of alkali halide crystals. The theoretical values, which

TABLE XXV.—THEORETICAL ENERGIES OF OXIDES, SULFIDES, AND SELENIDES DETERMINED BY USE OF THE b/r^n REPULSIVE POTENTIAL

(The electron affinities of O, S, and Se have not been determined by direct experiment. The last column lists the values of these affinities which must be assumed if the calculated and observed results are to agree.)

Salt	n	U (calc.), kg cal/mol	Affinity, kg cal/mol
MgO	7.0	940.1	−175
CaO	8.0	842.1	−171
SrO	8.5	790.9	−160
BaO	9.5	747.0	−157
MgS	8.0	778.3	− 72
CaS	9.0	721.8	− 71
SrS	9.5	687.2	− 77
BaS	10.5	655.9	− 80
CaSe	9.5	695.8	− 94
SrSe	10.0	667.1	− 92
BaSe	11.0	637.1	− 97

were computed by Sherman,[3] include small corrections that allow them to be compared with room-temperature experimental values. Sherman used the values of n appearing in the second column that were determined by use of Pauling's rules. As we mentioned in the preceding paragraph, these differ only slightly from those of Table XXIII. Experimental values for the fluorides are not listed because the electron affinity of fluorine is not known. The last column, however, contains a list of

[1] J. E. LENNARD-JONES, *Proc. Roy. Soc.*, **106**, 441, 463, 709 (1924); **109**, 476 (1925); **109**, 584 (1925).

[2] L. PAULING, *Proc. Roy. Soc.*, **114**, 181 (1927); *Jour. Am. Chem. Soc.*, **49**, 765 (1927); *Z. Krist.*, **67**, 377 (1928).

[3] J. SHERMAN, *Chem. Rev.*, **11**, 93 (1932).

electron affinities of fluorine that must be assumed in order to make the calculated and observed values agree. A similar compilation for several oxides, sulfides, and selenides is given in Table XXV. It should be noted that the values of affinities that must be assumed are negative in these cases.

Sherman used a set of affinities of negative ions that he derived from comparisons of theoretical and experimental cohesive energies of alkali halides and alkaline earth oxides, sulfides, and selenides in order to compare observed and computed energies of a number of other halides, oxides, sulfides, and selenides. Some of these cases are listed in Table XXVI. The agreement is very good for ideal ionic substances and is poor for substances such as cupric oxide.

b. The $ae^{-\frac{r}{\rho}}$ Interaction.—Investigations of interionic forces that have been carried out on the basis of quantum mechanics and that will be discussed in Chap. VII indicate that a repulsive term of the type b/r^n cannot be rigorously correct, although it may be a fair approximation for a short range of r. Born and Mayer[1] attempted to bring the older theory more nearly into accord with the quantum mechanical results by modifying it in two important ways: (1) They replaced the repulsion term (1) by an exponential expression of the type

$$\epsilon(r) = ae^{-\frac{r}{\rho}} \tag{8}$$

for the repulsive interaction of pairs of ions, in which a and ρ are constants. (2) They added another attractive term which is known as the van der Waals interaction. Mayer[2] subsequently modified this term and added another, as will be seen in the next section.

Born and Mayer found that they could take ρ in Eq. (8) as 0.345×10^{-8} cm for all types of ion if they determined a from the equation

$$a = b\left(1 + \frac{z_i}{n_i} + \frac{z_j}{n_j}\right)e^{\frac{r_i + r_j}{\rho}}. \tag{9}$$

Here, b is another fixed constant, z_i and z_j are the valences of two interacting ions, n_i and n_j are the numbers of valence electrons in the outer shells of the ions, and r_i and r_j are the ionic radii. The valences have negative signs for electronegative ions. n is equal to 8 for all simple ions except those which are isoelectronic with lithium, in which case n is 2. Born and Mayer used the ionic radii given in Table XXXII (page 93),

[1] M. BORN and J. E. MAYER, *Z. Physik*, **75**, 1 (1932); see paper following this as well.

[2] J. E. MAYER, *Jour. Chem. Phys.*, **1**, 270 (1933); see also J. E. MAYER and M. G. MAYER, *Phys. Rev.*, **43**, 605 (1933).

which were derived by Goldschmidt in the manner described in Sec. 14. Although Eq. (9) has some theoretical justification, its chief merit lies

TABLE XXVI.—EXPERIMENTAL AND THEORETICAL COHESIVE ENERGIES OF HALIDES OXIDES, SULFIDES, AND SELENIDES DETERMINED BY USE OF THE b/r^n REPULSIVE POTENTIAL

(The experimental values are referred to a standard state of free ions. They involve the following electron affinities, obtained by Sherman in the manner described in the text:

F	98.5 kg cal/mol	O	−168 kg cal/mol
Cl	92.5 kg cal/mol	S	− 79.4 kg cal/mol
Br	87.1 kg cal/mol	Se	− 97 kg cal/mol
I	79.2 kg cal/mol		

The theoretical values are determined from the Born theory by use of the values of n which appear in the third column.)

Substance	Structure	n	U (exp.), kg cal/mol	U (theor.), kg cal/mol	Difference, kg cal/mol
AgF	NaCl	8.5	223.0	207.9	15.1
MgF_2	Rutile	7.0	688.3	696.8	− 8.5
CaF_2	Fluorite	8.0	618.0	617.7	+ 0.3
NiF_2	Rutile	8.0	713.2	697.1	16.1
CuCl	Zincblende	9.0	226.3	206.1	20.2
AgCl	NaCl	9.5	207.3	187.3	20.0
TlCl	CsCl	10.5	170.9	159.3	11.6
Li_2O	Fluorite	6.0	693	695	− 2
Cu_2O	Cuprite	8.0	788	682	106
Ag_2O	Cuprite	8.5	715	585	130
NiO	NaCl	8.0	966	968	− 2
ZnO	Wurtzite	8.0	972	977	− 5
PbO_2	Rutile	9.5	2,831	2,620	−211
Al_2O_3	Corundum	7.0	3,617	3,708	− 91
Na_2S	Fluorite	8.0	524	516	8
Cu_2S	Fluorite	9.0	683	612	71
ZnS	Wurtzite	9.0	851	816	35
ZnS	Sphalerite	9.0	851	819	32
PbS	NaCl	10.5	731	705	26
ZnSe	Sphalerite	9.5	845	790	55
PbSe	NaCl	11.0	735	684	51
Cu_2Se	Fluorite	9.5	685	599	86

in the fact that it leads to a good correlation between observed and calculated results without additional assumptions.

As an illustrative example, let us consider a crystal that contains two types of ion, each of which is surrounded by M like ions and M' unlike

ions. If we consider only nearest like and unlike neighbors, the total repulsive interaction energy per mol is

$$E_r = N_A b\left[MC_{12}e^{\frac{r_1+r_2-r}{\rho}} + \frac{M'}{2}\left(C_{11}e^{\frac{2r_1}{\rho}} + C_{22}e^{\frac{2r_2}{\rho}}\right)e^{-\frac{a'r}{\rho}}\right]_{r_0} \tag{10}$$

where C_{12}, C_{11}, and C_{22} are appropriate values of the quantity $[1 + (z_i/n_i) + (z_j/n_j)]$ in Eq. (9), r is the distance between like ions, and $a'r$ is the distance between unlike ones. If we desire to use an approximation corresponding to that of part (*a*), we should add the electrostatic term to this and apply Eqs. (1) of Sec. 9 to the sum. Instead, Born and Mayer included two additional terms which we shall discuss before proceeding further.

12. The Multipole Interactions and the Zero-point Energy. *a. Van der Waals Terms.*—It may be concluded[1] from the existence of solid and liquid phases of the rare gases that the constituent atoms have a small but finite attraction for one another. London investigated this attraction on the basis of quantum mechanics and was able to derive a simple approximate expression for the interaction energy. We shall use the results here and discuss the derivation in Sec. 58. It turns out that any two atoms or ions have, in addition to the terms of the type discussed in Secs. 10 and 11, an attractive energy term of the approximate form.

$$\epsilon_v = -\frac{3h}{2r^6}\cdot\frac{\nu_1\nu_2}{\nu_1+\nu_2}\alpha_1\alpha_2 \tag{1}$$

where ν_1 and ν_2 are the series limit frequencies of the discrete spectra of the two atoms or ions, and α_1 and α_2 are their polarizabilities. More accurate expressions have been developed for particular cases, but this expression is very useful for a general discussion. It will be shown in Sec. 58 that this attraction is connected with a synchronization of the motion of the electrons in the two atoms or ions.

The total van der Waals energy for a lattice of the same type as that which we considered in deriving Eq. (10) of the preceding section is

$$E_v = -\frac{N_A}{r^6}\left[S_v a_{12} + S_v'\frac{a_{11}+a_{22}}{2}\right] \tag{2}$$

where a_{12}, a_{11}, and a_{22} are, respectively, the coefficients of $1/r^6$ in Eq. (1) for the positive-negative, positive-positive, and negative-negative ion pairs. S_v and S_v' are, respectively, the sums over all unlike and like ions of $1/R_i^6$ where R_i is the distance between a given ion and the ith ion in a lattice in which $r_0 = 1$. Numerical values of these rapidly convergent sums are given in Table XXVII for several types of structure.

[1] LENNARD-JONES, *op. cit.*

In their original work, Born and Mayer evaluated the a that occur in (2) by means of London's equation (1), using values of α and ν that were determined from data on free ions. The results represent only about 1 per cent of the total energy and do not appreciably affect the order of magnitude of agreement between theoretical and experimental

TABLE XXVII.—COEFFICIENTS FOR THE VAN DER WAALS AND FOR THE DIPOLE-QUADRUPOLE ENERGIES

Lattice type	S_n	S'_n	S_M	S'_M
Sodium chloride	6.5952	1.8067	6.1457	0.8002
Cesium chloride	8.7088	3.5445	8.2007	2.1476
Zincblende	4.354	0.762	4.104	0.253

values. These terms, however, are of considerable importance in determining the relative stability of two different lattice types, for they are comparable with the difference in energy of possible modifications. From a study of allotropy, Mayer[1] concluded that the α and ν values for free ions, when used in Eq. (1), do not account for the fact that CsCl, CsBr and CsI form simple cubic lattices instead of face-centered ones. He proposed that "better" values of these quantities should be determined from measurements of the optical properties of crystals by a method which we shall describe in Sec. 150. These values nearly double

TABLE XXVIII.—COMPARISON OF VALUES OF THE CONSTANT C APPEARING IN THE DIPOLE-DIPOLE TERM FOR THE VAN DER WAALS ENERGY C/r^6
(In 10^{-60} erg-cm^6)

	Atomic data	Mayer's method
LiF	6	18
LiBr	102	133
NaCl	100	130
NaI	247	482
RbCl	384	691
CsBr	1,258	2,070

the van der Waals energy. Table XXVIII contains several values of the coefficient of $1/r^6$ that were determined by using the two types of data. We shall employ the energies given by Mayer's method in the following sections of this chapter. It should be emphasized, however, that in doing so we are departing somewhat from the original concepts of the ionic theory. It probably is true that the treatment of ionic crystals that would reproduce Mayer's values of α and ν from the funda-

[1] J. E. MAYER, *Jour. Chem. Phys.*, **1**, 270 (1933); **1**, 327 (1933).

mental equations of quantum mechanics would not substantiate all details of the classical picture.

b. The Dipole-quadrupole Term.—The energy term (2), which varies with the inverse sixth power of the distance, actually is the first term in an infinite series of terms that successively decrease in importance. The next term, which was first investigated by Margenau,[1] varies as $1/r^8$ and usually is about one-tenth as large as the van der Waals term at the observed interionic distance. We shall discuss the derivation of an expression for this interaction that is analogous to the derivation of (1) in Sec. 58. For present purposes, it is sufficient to point out that the contribution to the total energy of the crystal from this term is

$$E_M = -\frac{N_A}{r^8}\left(S_M d_{12} + S'_M \frac{d_{11} + d_{22}}{2}\right) \tag{3}$$

in which the d are analogous to the a in Eq. (2) and S_M and S'_M are sums of $1/R^8$ analogous to the sums in Eq. (2). Table XXVII contains values of S_M for several types of crystal. We shall use values of d that were derived by Mayer from empirical measurements [see part (*a*) of this section].

The energy term (3) is called the dipole-quadrupole interaction because it may be interpreted as arising from the interaction of a dipole moment on one atom with a quadrupole moment on the other. For the same reason, the van der Waals term is sometimes called the dipole-dipole term. The next term in the sequence, which varies as $1/r^{10}$, is called the quadrupole-quadrupole term and is negligible in all cases in which we are interested.

c. Zero-point Energy.—The mechanical motion of the atoms or ions in a crystal containing N atoms may be treated as though the crystal were an assembly of $3N$ oscillators of various frequencies (*cf.* Sec. 22). According to quantum mechanics, an oscillator of frequency ν retains an energy of $h\nu/2$ at the absolute zero of temperature, which must be included along with the other energy terms when a comparison of experimental and theoretical cohesive energies is made. For present purposes, it is sufficiently accurate to use the Debye theory, according to which the frequencies are distributed between zero and a maximum ν_m in the manner described by the density function

$$f(\nu) = \frac{9N}{\nu_m^3}\nu^2. \tag{4}$$

Here, $f(\nu)d\nu$ is the number of oscillators having frequencies in the range from ν to $\nu + d\nu$. Thus, the total zero-point energy is

[1] H. MARGENAU, *Phys. Rev.*, **38**, 747 (1931). See also *Rev. Modern Phys.*, **11**, 1 (1939).

$$\int_0^{\nu_m} f(\nu)\frac{h\nu}{2}d\nu = \frac{9}{8}Nh\nu_m, \tag{5}$$

and the energy per mol of a diatomic crystal is

$$E_{A\nu} = N_A\frac{9}{4}h\nu_m. \tag{6}$$

The maximum frequency may be estimated to a sufficient degree of accuracy by using infrared absorption data, since the zero-point energy actually is small.

Following Mayer, we shall assume that the total energy of the crystal at absolute zero of temperature is given by the sum of the terms (2), (3), (6) of this section and Eqs. (4) and (10) of Secs. 10 and 11, respectively. This sum is

$$E_t = N_A\left\{-A\frac{e^2}{r} + b\left[MC_{12}e^{\frac{r_1+r_2-r}{\rho}} + \frac{M'}{2}(C_{11}e^{\frac{2r_1}{\rho}} + C_{22}e^{\frac{2r_2}{\rho}})e^{-\frac{a'r}{\rho}}\right] - \frac{1}{r^6}\left[S_v a_{12} + S'_v\frac{a_{11}+a_{22}}{2}\right] - \frac{1}{r^8}\left[S_M d_{12} + S'_M\frac{d_{11}+d_{22}}{2}\right] + \frac{9}{4}h\nu_m\right\}. \tag{7}$$

If we eliminate b by means of the first of Eqs. (1) of Sec. 9, we obtain

$$E_t = N_A\left[-A_{r_0}\frac{e^2}{r_0} + (1-k)\frac{\rho}{r_0}\left(A_{r_0}\frac{e^2}{r_0} + \frac{6C}{r_0^6} + \frac{8D}{r_0^8}\right) - \frac{C}{r_0^6} - \frac{D}{r_0^8} + \frac{9}{4}h\nu_m\right] \tag{8}$$

where

$$k = \frac{(a'-1)M'C_{22}\left(1+\frac{C_{11}}{C_{22}}e^{-\frac{2\delta}{\rho}}\right)e^{\left(\frac{\delta}{r_0}+1-a'\right)\frac{r_0}{\rho}}}{2MC_{12} + a'M'C_{22}\left(1+\frac{C_{11}}{C_{22}}e^{-\frac{2\delta}{\rho}}\right)e^{\left(\frac{\delta}{r_0}+1-a'\right)\frac{r_0}{\rho}}},$$

$$C = S_v a_{12} + S'_v\frac{a_{11}+a_{22}}{2},$$

$$D = S_M d_{12} + S'_M\frac{d_{11}+d_{22}}{2},$$

and δ is the difference of the ionic radii. The parameter ρ may be determined from E_t by means of the second of Eqs. (1), Sec. 9. Born and Mayer actually employed a slightly modified form of this equation which allowed them to use room-temperature values of β.

Table XXIX contains a compilation of the values of the five constituent terms of E_t for a number of halides. In addition, the sum of these terms is compared with the energies computed by Sherman (*cf.* sec. 11). Born and Mayer did not use the dipole-quadrupole term E_M

and the empirical van der Waals term when they derived the values of the repulsive term for the alkali halides that appear in this table. They completely neglected the first term and used a smaller value for the second (*cf.* part *b*, Sec. 11). We have subsequently added the E_M and empirical E_v terms, but we have not corrected for the change that this induces in the repulsive term. This change is practically negligible as long as we are interested only in a comparison of E_t with Sherman's values.

TABLE XXIX.—CONTRIBUTIONS TO THE COHESIVE ENERGIES OF THE ALKALI HALIDES AS GIVEN BY THE BORN-MAYER EQUATION
(In kg cal/mol)

	Madelung	Repulsive	Dipole-dipole	Dipole-quadrupole	Zero point	Total	Sherman
LiF	285.5	−44.1	3.9	0.6	−3.9	242.0	240.1
LiCl	223.5	−26.8	5.8	0.1	−2.4	200.2	193.3
LiBr	207.8	−22.5	5.9	0.1	−1.6	189.7	183.1
LiI	188.8	−18.3	6.8	0.1	−1.2	176.2	170.7
NaF	248.1	−35.3	4.5	0.1	−2.9	214.5	215.0
NaCl	204.3	−23.5	5.2	0.1	−1.7	184.4	180.4
NaBr	192.9	−20.6	5.5	0.1	−1.4	176.5	171.7
NaI	178.0	−17.1	6.3	0.1	−1.2	166.1	160.8
KF	215.1	−28.1	6.9	0.1	−2.2	191.8	190.4
KCl	183.2	−21.5	7.1	0.1	−1.4	167.5	164.4
KBr	174.5	−18.6	6.9	0.1	−1.2	161.7	157.8
KI	162.8	−15.9	7.1	0.1	−1.0	153.1	149.0
RbF	203.8	−26.2	7.9	0.1	−1.4	184.2	181.8
RbCl	175.8	−19.9	7.9	0.1	−1.2	162.7	158.9
RbBr	167.2	−17.6	7.9	0.1	−0.9	156.7	152.5
RbI	156.5	−15.4	7.9	0.1	−0.7	148.4	144.2
CsF	191.1	−23.9	9.7	0.1	−1.2	175.8	172.8
CsCl	162.5	−17.7	11.7	0.1	−1.0	155.6	148.9
CsBr	155.8	−16.4	11.4	0.1	−0.9	150.0	143.5
CsI	146.8	−14.6	11.1	0.1	−0.7	142.7	136.1

In computing the energy of the alkaline earth oxides and sulfides by means of Eq. (8), Mayer and Maltbie[1] used the mean of the values of ρ that were determined from the alkali halides. They estimated the van der Waals term from London's formula and omitted the dipole-quadrupole term. The results of this work are given in Table XXX and may

[1] J. E. MAYER and M. McC. MALTBIE, *Z. Physik*, **75**, 748 (1932).

be compared with Sherman's values in Table XXV. The affinities of oxygen and sulfur, determined from these newer computations, show a wider spread than those determined by Sherman.

TABLE XXX.—CRYSTAL ENERGIES DERIVED FROM THE BORN-MAYER EQUATION

Substance	Structure	$-E_0$, kg cal/mol	Affinity, kg cal/mol
MgO	NaCl	939	−190
CaO	NaCl	831	−165
SrO	NaCl	766	−144
BaO	NaCl	727	−147
MgS	NaCl	800	− 98
CaS	NaCl	737	−102
SrS	NaCl	686	− 85
BaS	NaCl	647	− 83

Bleick[1] has determined the lattice energy of the ammonium halides by the use of the Born-Mayer equation, treating the ammonium radical ion NH_4^+ as though it were a centrally symmetric unit. He obtained good agreement with measured cohesive energies in this way.

13. The Relative Stability of Different Lattice Types.—One of the important problems of crystal physics is that of determining the relative stability of different types of structure. This problem has a simple formal answer if questions of unstable equilibrium are neglected, for the stable lattice is that which has the lowest free energy. Unfortunately, the differences in free energies of different possible modifications may be extremely small; indeed, they are often less than the computational accuracy of the free energies. This fact is usually true, for example, of the differences in free energy of the four characteristic types of ionic structure (*cf.* Sec. 4). The free energies of different structures often are computed and compared in spite of this, the hope in such cases being that the computational errors lie in the same direction in each instance and cancel.

Mayer,[2] and Mayer and Levy,[3] using the Born-Mayer equation, have examined the relative stability of several halides in each of the structures listed in Table XXXI. They made no attempt to include temperature dependence, and so their discussion is valid only for the absolute zero of temperature. As they changed from the observed lattice type to a hypothetical one, they evaluated the differences of each of the constituent energy terms in Eq. (7) of Sec. 12, keeping r_0 fixed and equal to the

[1] W. E. BLEICK, *Jour. Chem. Phys.*, **2**, 160 (1934).
[2] J. E. MAYER, *Jour. Chem. Phys.*, **1**, 270 (1933); **1**, 327 (1933).
[3] J. E. MAYER and R. B. LEVY, *Jour. Chem. Phys.*, **1**, 647 (1933).

value for the observed stable lattice. They then added to the sum of these terms the change in energy of the hypothetical lattice as its interionic parameter changed from r_0 to the observed equilibrium value r_0'. This second term is equal to

$$\left(\frac{\partial^2 E_t'}{\partial r^2}\right)_{r=r_0'} \cdot \frac{(r_0 - r_0')^2}{2}$$

where E_t' is the energy of the hypothetical lattice. The zero-point energy was assumed to be the same for the different modifications. Table XXXI contains a list of the differences of the energy for the parameter r_0, the correction for change in density, and the sum of these two quantities. The sum would always be positive if the results agreed with experiment.

TABLE XXXI.—THE RELATIVE ENERGIES OF DIFFERENT LATTICE TYPES FOR SEVERAL HALIDES

(The first four columns of numbers are the changes in each of the contributions to the energy when the nearest-neighbor distance is held constant. The fifth column is the energy change δE obtained when the lattice expands or contracts to the true equilibrium distance. In kg cal/mol)

	Stable lattice	Hypothetical lattice	Change for fixed r_0				δE	Total change
			Madelung	Dipole-dipole	Dipole-quadrupole	Repulsive		
AgF	Sodium chloride	Zincblende	+14.7	+ 7.9	+1.3	−13.7	−1.9	+ 8.3
AgCl	Sodium chloride	Zincblende	+13.0	+ 9.7	+1.9	−13.8	−1.9	+ 8.9
AgBr	Sodium chloride	Zincblende	+12.5	+ 9.7	+1.6	−13.4	−1.9	+ 8.4
AgI	Zincblende	Sodium chloride	−12.8	−17.7	−3.7	+30.0	−1.9	− 6.1
AgF	Sodium chloride	Cesium chloride	− 2.0	− 9.7	−1.7	+16.7	−0.9	+ 2.4
AgCl	Sodium chloride	Cesium chloride	− 1.7	−11.8	−2.3	+24.7	−0.9	+ 8.0
AgBr	Sodium chloride	Cesium chloride	− 1.7	−11.6	−2.2	+28.7	−0.9	+12.3
TlCl	Cesium chloride	Sodium chloride	+ 1.4	+ 8.1	+1.3	−10.5	−0.6	− 0.3
TlBr	Cesium chloride	Sodium chloride	+ 1.4	+ 7.9	+1.2	−10.1	−0.6	− 0.2
TlI	Cesium chloride	Sodium chloride	+ 1.3	+ 8.2	+1.6	−10.1	−0.6	+ 0.4
CuCl	Zincblende	Sodium chloride	−15.4	−10.5	−1.9	+33.4	−1.9	+ 3.7
CuBr	Zincblende	Sodium chloride	−14.7	−11.0	−2.0	+34.3	−1.9	+ 4.7
CuI	Zincblende	Sodium chloride	−13.6	−13.6	−2.6	+34.0	−1.9	+ 2.3

It is not positive for AgI, TlCl, and TlBr, for in these cases, the sodium chloride type of structure, instead of the observed one, has the lowest energy. It may be seen from the table that the attractive terms favor the stability of the following three types of structure in the order given: CsCl, NaCl, and ZnS. Conversely, the repulsive term favors them in the reverse order. It would require only a small increase in one of the attractive terms, such as the van der Waals term, to shift the calculated stable lattice of TlCl and TlBr from the sodium chloride type to the cesium chloride type. On the other hand, it would require a compara-

tively large change of an opposite kind to account for the stability of the zincblende structure in AgI. Mayer concluded from these facts that thallium salts and most silver salts conform closely to the assumptions of the ionic theory and that AgI has some nonionic characteristics.

May[1] has made a very thorough investigation of the problem of the relative stability of the cesium chloride and sodium chloride lattice for cesium chloride on the basis of the Born-Mayer equation. He found that this equation does not account for the stability of the cesium chloride lattice at absolute zero of temperature if a two-parameter repulsive term and Mayer's value of E_0 are employed. In order to generalize the equation, he introduced two additional parameters. One of these was a multiplicative factor in the van der Waals term and the other was a factor in M' of Eq. (7), Sec. 12. In effect, the second parameter makes the constant b in the repulsive term different for like and unlike ions. These parameters were adjusted in order to make the cesium chloride lattice stable at absolute zero. The additional measured quantities that were used in doing this are the measured heat of the phase transition (about 1.34 kg cal/mol), and the observed change in lattice constant. The multiplicative constant of the attractive terms turns out to be 3.6, and the coefficient of M' to be 0.70. At the same time, the constant b is doubled, and ρ changes from 0.290 to 0.365Å. May suggested that part of the required increase in the attractive terms should be associated with a change in the purely electrostatic energy caused by distortion of the ions from spherical form. It is easy to show that the distortion of the charge on an ion in a cubic crystal may be described in the first approximation by fourth-order surface harmonics and that the associated potential, which is not centrally symmetrical, varies with r as $1/r^9$. There is, however, no conclusive quantitative evidence to support May's suggestion.

May also carried through a similar treatment of the stability of ammonium chloride and found further support for the conclusions drawn from the case of cesium chloride. In addition to this, he estimated the transition temperature for both cesium chloride and ammonium chloride and obtained qualitative agreement with the observed facts.

In contrast with May's work is that of Jacobs,[2] who correlated the phase transitions that occur in the alkali halides at high pressures with empirically determined constants that are more nearly like those used in the earlier work of Born and Mayer.

14. Ionic Radii.—We saw, in Sec. 4, Chap. I, that the nearest ion distances of the alkali halide crystals are additive. This fact, which was

[1] A. MAY, *Phys. Rev.*, **52**, 339 (1937).
[2] R. B. JACOBS, *Phys. Rev.*, **54**, 468 (1938).

first emphasized by Goldschmidt,[1] suggests that the ions of these crystals may be regarded as rigid spheres and that the equilibrium interionic distance is simply the distance at which these spheres come into contact. This view is justified in a rough way by the Born theory, which shows that the repulsive forces between ions vary about ten times more rapidly with interionic distance than do the electrostatic forces. The rigid sphere concept, however, cannot be used very widely for quantitative purposes; we shall discuss some of its limitations in a later paragraph of this section.

It is not possible to obtain the values of ionic radii from crystallographic data without knowing at least one radius at the start. Gold-

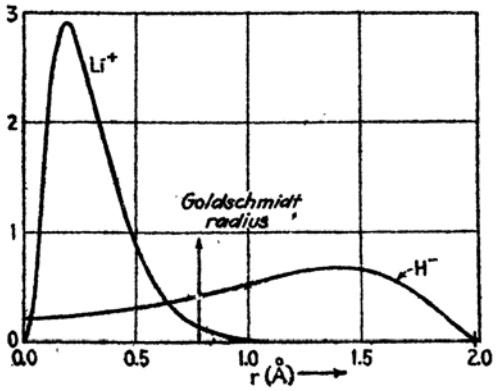

FIG. 1.—The charge distributions of Li^+ and H^- ions as determined by Hylleraas and Bethe. The vertical line marks the point of contact of the Goldschmidt radii.

schmidt began his extensive work on the tabulation of radii with the following values of the radii of F^- and O^{--}:

$$F^-: 1.33 \text{ A}; \qquad O^{--}: 1.32 \text{ Å}.$$

These values were determined by Wasastjernas[2] who correlated measurements on the optical properties of fluorides and oxides with a classical quantum theory of dispersion and thereby determined the radii of the outermost orbits of these ions. The theoretical basis of this work need not concern us here. We may, however, compare the value of the radius of Li^+ that is deduced from the value for F^- with the electronic charge distribution of Li^+ as determined by modern quantum mechanics,[3] which is shown in Fig. 1. It is obvious that there is no precisely defined radius; however, the vertical line corresponding to Wasastjernas' radius is a reasonable point at which to say that the distribution stops. The right-

[1] V. M. GOLDSCHMIDT, *Skrifter det Norske Videnskaps* (1926, 1927).
[2] J. WASASTJERNAS, *Comm. Fenn.*, **1**, (1923); **6**, (1932).
[3] E. HYLLERAAS, *Z. Physik*, **54**, 347 (1929); H. BETHE, *Z. Physik*, **57**, 815 (1929).

hand part of Fig. 1 shows the distribution in H^-. The two curves are plotted in such a way that the distance between origins corresponds to the Li^+–H^- distance in lithium hydride. Table XXXII contains a

TABLE XXXII.—GOLDSCHMIDT'S IONIC RADII
(IN Å)

H^-	1.27	O^{--}	1.32
F^-	1.33	S^{--}	1.74
Cl^-	1.81	Se^{--}	1.91
Br^-	1.96	Te^{--}	2.03
I^-	2.20		
Li^+	0.78	Mg^{++}	0.78
Na^+	0.98	Ca^{++}	1.06
K^+	1.33	Sr^{++}	1.27
Rb^+	1.49	Ba^{++}	1.43
Cs^+	1.65	Be^{++}	0.34
Tl^+	1.49	Zn^{++}	0.83
Cu^+	0.53	Cd^{++}	1.03
Ag^+	1.0	Hg^{++}	1.12
Mn^{++}	0.91	Al^{+++}	0.57
Fe^{++}	0.83	Se^{+++}	0.83
Co^{++}	0.82	Yt^{+++}	1.06
Ni^{++}	0.78	La^{+++}	1.22
Pb^{++}	1.32	Ga^{+++}	0.62
		In^{+++}	0.92
		Tl^{+++}	1.05

list of Goldschmidt's radii, which have been determined from additive systems. These values are self-consistent to within about 5 per cent.

Other radii have been obtained by other workers using different starting assumptions. For example, Huggins and Mayer[1] obtained the radii, given in Table XXXIII, for the alkali ions and halogen ions by adjusting

TABLE XXXIII.—THE RADII OF THE ALKALI METAL IONS AND THE HALOGEN IONS AS DETERMINED BY HUGGINS AND MAYER
(IN Å)

Li^+	0.475	F^-	1.110
Na^+	0.875	Cl^-	1.475
K^+	1.185	Br^-	1.600
Rb^+	1.320	I^-	1.785
Cs^+	1.455		

the radii that appear in the Born-Mayer equation so that the theory would lead to the observed lattice constant; that is, b in Eq. (7) of Sec. 12 was given a fixed value and the ionic radii were adjusted so that Eq. (1) of Sec. 10 was satisfied for the observed lattice spacing. The values in Table XXXIII differ appreciably from Goldschmidt's values.

[1] M. L. HUGGINS and J. E. MAYER, *Jour. Chem. Phys.*, **1**, 643 (1933).

The cesium chloride type of structure has a larger electrostatic cohesive energy than the sodium chloride type for a given value of r_0, the nearest ion distance. In view of this fact, one might attempt to explain the stability of the sodium chloride type in most alkali halides on the basis of the rigid ion picture, by assuming that one type of ion is so large that like neighbors touch at greater values of r_0 in the cesium chloride structure than unlike ones do in the sodium chloride structure. A condition obviously necessary for this is that one set of ions should be so large that pairs touch in the cesium chloride structure for the observed interionic distance of the sodium chloride type of structure, that is, that the radii R_1 and R_2 should satisfy the conditions

$$R_1 \geq \frac{r_0}{\sqrt{3}}, \qquad R_2 \leq \left(1 - \frac{1}{\sqrt{3}}\right) r_0.$$

The ratios must then obey the inequality

$$\frac{R_2}{R_1} < \sqrt{3} - 1 = 0.73.$$

KF, RbF, RbCl, RbBr, and CsF do not satisfy this condition if Goldschmidt's radii are used; moreover, no reasonable and self-consistent change in radii would make all the ions in face-centered alkali halide crystals satisfy it.

In view of the work of the previous section, the failure of the simple rigid-sphere concept to cope with this problem is not surprising. The stability of an ionic crystal in any given structure is determined by a delicate balance of several types of force.

15. Implications of Deviations from the Cauchy-Poisson Relations.—In all the computational work that is discussed in the preceding sections of this chapter, it is explicitly assumed that the ions are spherically symmetrical and interact with spherically symmetrical forces. If this assumption were correct, the three elastic constants,[1] c_{11}, c_{12}, and c_{44}

[1] We shall use Voigt's notation for the elastic constants [see W. Voigt, *Lehrbuch der Kristallphysik* (Teubner, Leipzig, 1910; reissued 1928)]. If we designate the three tension components of the stress tensor by X_1, X_2, and X_3, the three shear components by X_4, X_5, and X_6, the three tension components of the strain tensor by x_1, x_2, and x_3, and the three shear components by x_4, x_5, and x_6, Hooke's stress-strain relation is given in terms of the thirty-six elastic constants c_{ij} ($i, j = 1, 2, \cdots, 6$) by the equations

$$X_i = \sum_{j=1}^{6} c_{ij} x_j.$$

If $E(x_1, \cdots, x_6)$ is the energy per unit volume of the crystal as a function of the

of a cubic crystal should satisfy the Cauchy-Poisson relation

$$c_{12} = c_{44}.$$

The measurements of Rose,[1] which are plotted in Fig. 2, show that this relation is not valid for sodium chloride at low temperatures, the

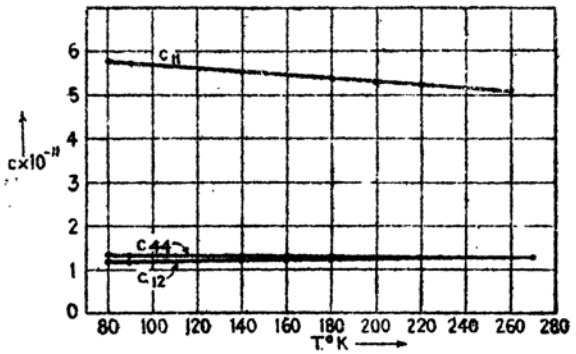

FIG. 2.—The three elastic constants of sodium chloride. c_{44} and c_{12} cross at 240°K. (*After Rose.*)

discrepancy being of the order of 10 per cent. Similar measurements[2]

---

strains, it follows from the relation between force, work, and energy that

$$E(x_1, \cdots, x_6) = \frac{1}{2}\sum_{i,j} c_{ij}x_i x_j.$$

Thus,

$$c_{ij} = \frac{\partial^2 E}{\partial x_i \partial x_j}$$

and

$$c_{ij} = c_{ji};$$

hence, only twenty-one of the thirty-six c_{ij} are independent. For cubic crystals, there are the additional relations

$$c_{11} = c_{22} = c_{33}, \qquad c_{12} = c_{23} = c_{31}, \qquad c_{44} = c_{55} = c_{66}$$

and all other components are zero if the coordinate axes and cube axes are parallel; hence, there are only three independent components in this case. In hexagonal crystals, there are the relations

$$c_{11} = c_{22}, \qquad c_{13} = c_{23}, \qquad c_{44} = c_{55}, \qquad c_{66} = \tfrac{1}{2}(c_{11} - c_{12}),$$

and all other components except c_{12} are zero, if the hexagonal axis is in the z direction. If the atoms of the crystal interact with central forces, there are the additional Cauchy-Poisson relations, which in the cases of cubic and hexagonal crystals are

$$c_{12} = c_{44}, \qquad \text{and} \qquad c_{13} = c_{44}, \qquad c_{11} = \tfrac{3}{2}c_{12},$$

respectively.

[1] F. C. Rose, *Phys. Rev.*, **49**, 50 (1936).

[2] Data furnished by Quimby, Balsmuth, and Rose on magnesium oxide appears in a paper by Barnes, Brattain, and Seitz, *Phys. Rev.* [illegible] (1935). See also the correction, *Phys. Rev.*, **49**, 405 (1936); M. A. Durand, *Phys. Rev.*, **50**, 449 (1936).

on magnesium oxide show a deviation between c_{12} and c_{44} of about 50 per cent. It is natural to ask how this discrepancy affects the Born theory.

The terms in the Born-Mayer equation may be divided into two classes, namely: the electrostatic term f_1, which varies slowly with r, and all others f_2, which vary rapidly in comparison. The first term is rigorously known, for although the charge distribution may not be precisely spherically symmetrical, the electrostatic-correction term arising from this distortion varies at least as rapidly as $1/r^9$ and may be classed in f_2. It follows that the deviations from the Cauchy-Poisson relations occur because some of the terms in f_2 should not be spherically symmetrical.

We have seen in the preceding work that the absolute magnitude of f_1 is about ten times that of f_2 for equilibrium values of r, even though the force terms $\partial f/\partial r$ are equal. The increase in importance of f_2 as we pass from a consideration of energies to one of forces is an immediate consequence of the fact that f_2 varies more rapidly with r than does f_1. This fact is shown in a striking way by the older form of the repulsive term b/r_0^n, where n is of the order of magnitude 10. The ratio of this term to the electrostatic term is

$$\frac{b}{Ar^{n-1}} = \frac{1}{n},$$

whereas the ratio of the derivatives is n times this. Similarly, f_2'' is about ten times larger than f_1'', whence we may conclude that the elastic constants, which are related to second derivatives of the energy, are primarily determined by f_2. Thus, the experimental measurements of elastic constants show that f_2 is in error by about 10 per cent in sodium chloride and about 50 per cent in magnesium oxide. This error affects the cohesive energy by only 1 per cent in the first case and 5 per cent in the second, because f_2 accounts for only about 10 per cent of the cohesive energy. The energy differences of different crystallographic forms, however, are of the same order of magnitude as this error. Hence, it seems safe to say that these energies can be computed accurately only when the nonradial part of f_2 is properly included in the Born-Mayer equation.

16. Surface Energy.—One of the factors that determines whether or not a given crystallographic plane is a cleavage plane is its surface energy, that is, the energy σ per unit area required to separate the crystal along this plane. This energy usually is defined in such a way that the areas of the two halves of the crystal are counted separately. Surface energies of crystals with the sodium chloride lattice have been computed on the

basis of the ionic theory by Madelung,[1] Born and Stern,[2] and Yamada.[3] Madelung took into account only electrostatic forces and found that the energy of a (100) surface is $0.520e^2/a^3$ per unit area, where a is the cube-edge distance of the unit cell. Born and Stern improved this calculation by including repulsive forces and found the following values for the (100) and (110) planes:

$$\left.\begin{aligned}\sigma_{100} &= 0.116\frac{e^2}{a^3},\\ \sigma_{110} &= 0.315\frac{e^2}{a^3}.\end{aligned}\right\} \qquad (1)$$

Since the second is about 2.7 times larger than the first, it follows that rock salt should split more easily along a (100) plane than along a (110) plane. This result is in agreement with experiment, since only (100) planes show cleavage. In both the preceding calculations, it was assumed that the ionic spacing near the surface is the same as in the interior of the lattice. Madelung, however, has shown that, in the case of (100) surfaces, there should be a slight contraction of interionic distance in the direction normal to the (100) plane.

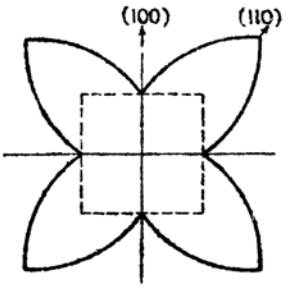

FIG. 3.—Polar diagram of the surface energy of crystals with the sodium chloride lattice. The angular variable is the angle between the normal to the surface plane and the (100) direction. Although this plot is valid only for planes whose normals lie in a principal plane, the minimum in the (100) direction is an absolute minimum. The dotted square represents the crystal form of lowest surface energy, as determined by Wulff's method. (*After Yamada.*)

Yamada extended Born and Stern's computations by calculating σ for all surface planes that have normals lying in the (100) plane. This includes as special cases planes that are equivalent to the (100) and (110) surfaces. His values of σ are shown in the polar diagram of Fig. 3, in which the angular variable is the angle between the (100) direction and the normal to the plane. This figure shows that the value of σ for the (100) plane is a relative minimum, whereas that for the (110) plane is a relative maximum. Yamada also found that the surface energy of the (100) plane is smaller than that of any other plane, so that the (100) value is also an absolute minimum for the three-dimensional σ surface.

Wulff[4] has proved that the crystal form corresponding to the lowest surface energy may be obtained from a polar diagram of this type by taking the envelope of those planes that are orthogonal to lines passing through the origin at the points where these lines intersect the surface.

[1] E. MADELUNG, *Physik. Z.*, **19**, 524 (1918); **20**, 494 (1919).
[2] M. BORN and O. STERN, *Sitzb. preuss. Akad. Wiss.*, 901 (1919).
[3] M. YAMADA, *Physik. Z.*, **24**, 364 (1923); **25**, 52 (1924).
[4] G. WULFF, *Z. Krist.*, **34**, 449 (1901).

It is clear that this form is the dotted square of Fig. 3 in the two-dimensional case corresponding to Yamada's result for the (100) plane and is a cube in the three-dimensional case. In order to obtain the condition for thermodynamical equilibrium above absolute zero, Wulff's construction should be applied to the free-energy polar diagram instead of to the energy diagram.

If ionic crystals were perfect, it should be possible to estimate their breaking strength from the surface energy. Following Polanyi,[1] we shall assume that the interatomic forces between two separating crystal planes are important for a distance Δl of about ten angstrom units and shall determine a rough value of the breaking stress S by use of the energy equation

$$2\sigma \cong \Delta l S \tag{2}$$

where σ is the surface energy per unit area of the crystal. The value of σ for sodium chloride that is given by the first of the equations (1) is about 147 ergs cm^{-2}, which corresponds to a value of S of about 300×10^7 dynes cm^{-1}. The values ordinarily observed are about one hundred times smaller than this. Zwicky[2] carried through a more exact calculation of S on the basis of the ionic theory and found the value

$$S = 2{,}000 \times 10^7 \text{ dyne-cm}^{-1},$$

which is in even worse agreement with the measured values. Since measurements of the surface tension of molten salts lead to values of σ that are in comparative agreement with the theoretical values, we may not conclude that the computational methods for a perfect lattice are badly in error.

At the present time, it is believed that these discrepancies between observed and calculated breaking strengths arise from lattice imperfections. This type of explanation was first proposed by Griffith,[3] who suggested that the weak spots are tiny surface cracks. A more complete theory has been developed by Taylor.[4] He assumes that the weakening centers, which he calls dislocations, are distributed throughout the crystal and play an important role not only in lowering the breaking strength but also in determining the plastic properties of crystals. We shall refer the reader to other sources[5] for a discussion of this work.

[1] M. POLANYI, *Z Physik*, **7**, 323 (1921).

[2] F. ZWICKY, *Physik. Z.*, **24**, 131 (1923).

[3] A. A. GRIFFITH, *Phil. Trans. Roy. Soc.*, **221**, 163 (1920).

[4] G. I. TAYLOR, *Trans. Faraday Soc.*, **24**, 121 (1928); *Proc. Roy. Soc.*, **145**, 362, 388 (1934).

[5] See E. SCHMID and W. BOAS, *Kristallplastizität* (Julius Springer, Berlin, 1935); C. F. ELAM, *The Distortion of Metal Crystals* (Oxford University Press, 1935); also the report of the Bristol Conference on this topic, *Proc. Phys. Soc.*, January, 1940.

CHAPTER III

THE SPECIFIC HEATS OF SIMPLE SOLIDS

17. Introduction.—It was shown in Chap. I that the specific heats of practically all simple solids approach zero monotonically as the temperature approaches absolute zero. Since this fact cannot be explained satisfactorily on the basis of classical mechanics, Einstein's qualitative interpretation[1] on the basis of quantum theory may be regarded as one of the first triumphs of the theory. Einstein postulated, as had been done previously, that a simple crystal could be regarded as an aggregate of atomic oscillators, all of which vibrate with the same natural frequency. In addition, he assumed that the allowed energy states of these oscillators are integer multiples of $h\nu$, where ν is the frequency of oscillation and h is Planck's constant. In classical mechanics, it would have been assumed that the allowed energy states are continuous; and this assumption, according to classical statistical mechanics, would mean that Dulong and Petit's law should be valid at all temperatures. Einstein found upon applying Boltzmann's theorem to his postulated assemblage of quantum oscillators that the observed decrease in specific heat could be explained qualitatively.

A few years later, Debye[2] showed that most of the quantitative discrepancy between Einstein's results and observations could be removed by the introduction of a more accurate analysis of the modes of vibration of a simple solid. In short, he improved upon Einstein's assumption that all atoms oscillate with the same frequency. One outstanding result of Debye's work is the prediction that the specific heat should approach zero as T^3 near the absolute zero of temperature. This law is accurate for a large range of temperature in many cases; in others, it is not accurate.

Born and von Kármán[3] developed a method for computing the modes of vibration of a solid about the same time as did Debye. Although their method actually is more accurate than Debye's, since it is based upon fairly rigorous principles of atomic dynamics, Debye's results agreed so well with available experimental material that his theory was considered sufficiently accurate for practical purposes. In recent years,

[1] A. Einstein, *Ann. Physik*, **22**, 180, 800 (1906); **34**, 170 (1911).
[2] P. Debye, *Ann. Physik*, **39**, 789 (1912).
[3] M. Born and Th. von Kármán, *Physik. Z.*, **13**, 297 (1912); **14**, 15 (1913).

however, a number of striking deviations from Debye's theory have been found in the low-temperature range where it might be expected to be most accurate. For this reason, Blackman[1] reopened the question of determining the distribution of frequencies, and he found that Born and von Kármán's method gives something more nearly like the observed specific-heat curves. Blackman's work is essentially qualitative, however, and it remains to be seen how well the low-temperature anomalies actually can be explained.

The three stages of development of the vibrational theory of specific heats, namely, the theory of Einstein, of Debye, and of Born, von Kármán, and Blackman, may be included in an approximation in which it is assumed that interatomic forces obey Hooke's law. We shall treat specific heats of solids by this method in the present chapter.

It is believed that the details of specific-heat curves that cannot be explained by the Hooke's law approximation have two origins, namely, anharmonic interaction terms and thermal excitation of electrons. These topics will be discussed in later chapters. Anharmonic interaction terms supposedly account for the following facts: (*a*) the anomalous peaks in the specific-heat curves of molecular crystals, such as solid methane, and of ionic crystals, such as ammonium chloride; and (*b*) a part of the high-temperature deviations from Dulong and Petit's law. Electronic interaction is believed to account for: (*a*) the linear temperature dependence of the specific heats of some metals near absolute zero, (*b*) the anomalous peaks in the specific heats of ferromagnetic metals and paramagnetic salts, and (*c*) a part of the high-temperature deviation from Dulong and Petit's law—particularly that of transition metals.

18. The Energy of an Assembly of Oscillators.—It will be shown in Sec. 22 that a crystal which contains N atoms that interact according to Hooke's law is mechanically equivalent to a set of $3N$ independent oscillators. Thus, the mean total energy $\bar{E}$ of such a crystal is equal to the sum of the mean energies of the individual oscillators. We shall derive the expression for the mean energy of an oscillator by assuming discrete energy levels in accordance with quantum theory.

The relative probability[2] of finding in its ith level a system that has levels of energy ϵ_i and degeneracy g_i is

$$r_i = g_i e^{-\frac{\epsilon_i}{kT}}. \tag{1}$$

[1] M. BLACKMAN, *Z. Physik*, **86**, 421 (1933); *Proc. Roy. Soc.*, **148**, 384 (1935), **159**, 416 (1937); *Proc. Cambridge Phil. Soc.*, **33**, 94 (1937).

[2] For a discussion of Boltzmann's theorem, see various books on statistical mechanics such as: E. H. Kennard, *Kinetic Theory of Gases* (McGraw-Hill Book Company, Inc., New York, 1938); L. Brillouin, *Die Quantenstatistik* (Julius Springer, Berlin, 1931).

Using this, we find that the absolute probability p_i is

$$p_i = \frac{g_i e^{-\frac{\epsilon_i}{kT}}}{\sum_j g_j e^{-\frac{\epsilon_j}{kT}}}$$

where j is summed over all levels. The mean energy $\bar{\epsilon}$ of the system obviously is

$$\bar{\epsilon} = \sum_i p_i \epsilon_i = \frac{\sum_i \epsilon_i g_i e^{-\frac{\epsilon_i}{kT}}}{\sum_j g_j e^{-\frac{\epsilon_j}{kT}}}, \tag{3}$$

which is identical with the expression

$$\bar{\epsilon} = -\frac{d}{d(1/kT)} \log \left(\sum_i g_i e^{-\frac{\epsilon_i}{kT}} \right). \tag{4}$$

We shall call the sum

$$\sum g_i e^{-\frac{\epsilon_i}{kT}}$$

the partition function and shall designate it by f.

Let us evaluate Eq. (3) for a harmonic oscillator of which the energy levels, in accordance with quantum theory, are given by

$$\epsilon = nh\nu \tag{5}$$

where n takes all integer values and ν is the natural frequency of the oscillator. The levels are not degenerate in this case, whence g_i is unity. Thus, the partition function for the system is

$$f = \sum_{n=0}^{\infty} e^{-\frac{nh\nu}{kT}} = \sum_{n=0}^{\infty} (e^{-\frac{h\nu}{kT}})^n = \frac{1}{1 - e^{-\frac{h\nu}{kT}}}. \tag{6}$$

According to Eq. (4), the mean energy is

$$\bar{\epsilon} = \frac{d}{d(1/kT)} \log (1 - e^{-\frac{h\nu}{kT}}) = \frac{h\nu e^{-\frac{h\nu}{kT}}}{1 - e^{-\frac{h\nu}{kT}}} = \frac{h\nu}{e^{\frac{h\nu}{kT}} - 1}, \tag{7}$$

whereas the mean total energy of an assembly of $3N$ oscillators of different frequencies ν_i ($i = 1, 2, \cdots, 3N$) is

$$\bar{E} = \sum_{i=1}^{3N} \frac{h\nu_i}{e^{\frac{h\nu_i}{kT}} - 1} \tag{8}$$

This sum may be replaced by an integral if the distribution of frequencies can be represented by an integrable function. In this case, which occurs commonly, as will be seen later,

$$\bar{E} = \int_0^{\nu_m} \frac{h\nu}{e^{\frac{h\nu}{kT}} - 1} q(\nu)d\nu \tag{9}$$

where $q(\nu)d\nu$ is the number of oscillators in the frequency range from ν to $\nu + d\nu$, and ν_m is the maximum frequency of any oscillator. Obviously, we may expect $q(\nu)$ to satisfy the equation

$$\int_0^{\nu_m} q(\nu)d\nu = 3N. \tag{10}$$

The derivative of Eq. (8) or (9) with respect to T is the molar heat C_V in the case in which N is the number of atoms in a mol of substance N_m. We have

$$C_V = \sum_{i=1}^{3N_m} k\left(\frac{h\nu_i}{kT}\right)^2 \frac{e^{\frac{h\nu_i}{kT}}}{(e^{\frac{h\nu_i}{kT}} - 1)^2} \tag{11}$$

in the general case, and we have

$$C_V = \int_0^{\nu_m} k\left(\frac{h\nu}{kT}\right)^2 \frac{e^{\frac{h\nu}{kT}}}{(e^{\frac{h\nu}{kT}} - 1)^2} q(\nu)d\nu, \tag{11a}$$

when Eq. (9) may be used. The summand in (11), namely,

$$k\left(\frac{h\nu}{kT}\right)^2 \frac{e^{\frac{h\nu}{kT}}}{(e^{\frac{h\nu}{kT}} - 1)^2}, \tag{12}$$

increases monotonically with temperature. It is zero at absolute zero; and as T approaches infinity, it approaches the constant value k in the manner indicated in Fig. 1. It should be clear that, with positive coefficients, any sum or integral of such a monotonically increasing function also must increase monotonically. Therefore, we may conclude that C_V is an increasing function of temperature regardless of the frequency

distribution. This conclusion is valid only when the interatomic forces obey Hooke's law and when electrons do not contribute to the specific heat. Since (12) is equal to k when kT is much greater than $h\nu_m$, the sum (11) in this case reduces to

$$C_V = 3N_m k = 3nR \tag{13}$$

where R is the gas constant and n is the number of atoms per molecule of the crystal. Thus, the law of Dulong and Petit is valid at high temperatures for an assembly of oscillators. This limiting case also corresponds to the case in which quantum theory degenerates to classical theory, that is, when h approaches zero.

It is not possible to draw further important conclusions from this theory unless the form of the frequency distribution function $q(\nu)$ is known. In the next section, we shall discuss several different assumptions that have been made about this function.

19. Einstein's and Debye's Frequency Distributions. *a. Einstein's Distribution.*—Einstein[1] postulated, for simplicity, that the atoms of a solid oscillate independently and with the same frequency ν. The summand in Eq. (8) of Sec. 18 is constant under this assumption, whence,

$$\bar{E} = 3N\frac{h\nu}{e^{\frac{h\nu}{kT}} - 1}, \tag{1}$$

and

$$C_V = 3nR\left(\frac{h\nu}{kT}\right)^2 \frac{e^{\frac{h\nu}{kT}}}{(e^{\frac{h\nu}{kT}} - 1)^2}$$

$$= 3nRf_E\left(\frac{h\nu}{kT}\right). \tag{2}$$

f_E will be called the Einstein specific-heat function (*cf.* Fig. 1). The parameter ν in Eq. (2) is usually adjusted so as to obtain the best possible fit of theoretical and observed specific-heat curves. Its magnitude usually is of the order of an infrared frequency. Although the parts of specific-heat curves for which C_v is greater than $3nR/2$ usually can be fitted closely with an Einstein specific-heat function, the low-temperature portions usually cannot. It may be recalled from the discussion of Chap. I that the observed specific heat usually drops to zero as T^3, whereas Eq. (2) varies as

$$3nR\left(\frac{h\nu}{kT}\right)^2 e^{-\frac{h\nu}{kT}}$$

[1] EINSTEIN, *op. cit.*

which approaches zero more rapidly. We shall discuss the comparison of observed and theoretical values in more detail after presenting Debye's theory.

b. Debye's Distribution.—Debye[1] pointed out that the possible oscillation frequencies of a solid body range from very low values, corresponding to sound waves, to very high ones, corresponding to infrared absorption peaks of solids. Since the contribution of a low-frequency oscillator to the low-temperature specific heat is larger than that of a high-frequency one, we might have anticipated that Einstein's assumption, which neglects the long waves, would not give a large enough specific heat at very low temperatures.

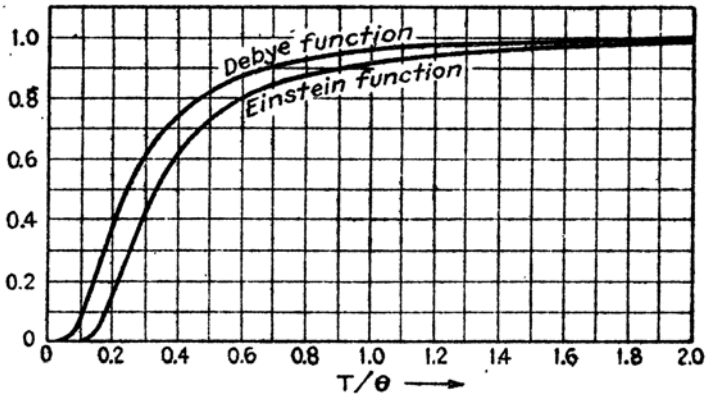

FIG. 1.—A comparison of the Einstein and Debye specific-heat curves. The frequency associated with the Einstein curve is equal to the maximum Debye frequency.

Debye suggested that a more accurate distribution might be found by treating the crystal as though it were a continuous elastic medium which has elastic constants that are independent of frequency. It has become conventional to follow Debye's original simplified treatment of the problem in which the medium is assumed to be isotropic. Instead, one might use the actual elastic constants. The factor that makes the more precise procedure difficult to apply is not so much the problem of finding the normal modes of vibration as the problem of expressing the frequency density in a tractable analytical form. In the isotropic case the frequency is independent of the direction of propagation, so that the distribution function may be expressed simply. We shall proceed with the simplified treatment and discuss the alternative one later.

The following point should be kept in mind. The vibrational frequencies of a continuous medium range from zero to infinity, corresponding to wave lengths that range from infinity to zero, whereas, in an actual solid, wave lengths shorter than interatomic distances are not independ-

[1] DEBYE, *op. cit.*

ent of longer ones. This interdependence is shown for a simple example in Fig. 2. The transverse displacements of the particles in this figure may be described equally well by the function a or b, or by any one of an infinite number of shorter waves that have the same amplitudes as a or b at the atomic positions.

The same conclusion may be drawn from the fact that a system which contains N atoms has only $3N$ degrees of freedom, whereas a strictly continuous medium has an infinite number of degrees of freedom. This interdependence of different modes is not exhibited as long as the medium is treated as though it were continuous. Hence, it is necessary to introduce a condition that limits the number of modes used in determining the specific heat. This may be done conveniently by neglecting all

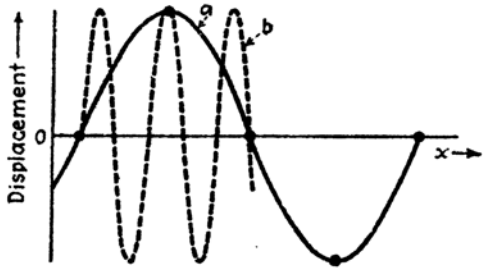

FIG. 2.—An example of the way in which a given periodic displacement of a lattice may be described by different harmonic functions a and b. The wave numbers σ_a and σ_b of the two functions bear the relation $\sigma_b = \sigma_a + 1/l$, where l is the lattice distance.

modes having a frequency greater than the value ν_m defined by the equation

$$\int_0^{\nu_m} f(\nu)d\nu = 3N, \tag{3}$$

in which $f(\nu)$ is the distribution function for a continuous medium. This, however, is not a unique way of selecting the modes of vibration. We could, for example, choose them by placing a restriction upon the range of wave lengths rather than upon the range of frequencies. The question of choosing modes does not arise when they are derived by a detailed atomic treatment like that of Secs. 20, 21, and 22 for this method automatically gives us a unique set of $3N$ modes.

Let us consider a rectangular parallelepiped of an isotropic medium, bounded by the six planes $x = 0$, $x = L_x$, $y = 0$, $y = L_y$, $z = 0$, $z = L_z$. The equations for the propagation of longitudinal and transverse vibrations in this medium are

$$\Delta u_l = \frac{1}{c_l^2}\frac{\partial^2 u_l}{\partial t^2}, \tag{4a}$$

$$\Delta u_t = \frac{1}{c_t^2}\frac{\partial^2 u_t}{\partial t^2}, \tag{4b}$$

respectively, where u_l and u_t are the amplitudes of the vibrations and c_l and c_t are their velocities. These velocities may be related to the two elastic constants for isotropic media by means of the equations[1]

$$c_l = \sqrt{\frac{m}{c_{11}}}, \qquad c_t = \sqrt{\frac{2m}{c_{11} + c_{12}}}$$

in which m is the density.

The boundary conditions employed when one searches for the normal modes of vibration depend upon the specific physical conditions under which the body vibrates. Since the frequency distribution turns out to be insensitive to the choice of boundary conditions, we shall assume for simplicity that the surface planes are rigidly held. In this case, the boundary condition is that u should be zero at these planes. With this condition, the normal modes are

$$u_l = A_l \sin \frac{\pi n_x x}{L_x} \sin \frac{\pi n_y y}{L_y} \sin \frac{\pi n_z z}{L_z} \sin 2\pi \nu_l t, \tag{5a}$$

$$u_t = A_t \sin \frac{\pi n_x x}{L_x} \sin \frac{\pi n_y y}{L_y} \sin \frac{\pi n_z z}{L_z} \sin 2\pi \nu_t t, \tag{5b}$$

where the n_i are arbitrary integers and the A_i are constants. The frequencies ν_l and ν_t are related to the n_i by means of the equations

$$\nu_l^2 = \frac{c_l^2}{4}\left(\frac{n_x^2}{L_x^2} + \frac{n_y^2}{L_y^2} + \frac{n_z^2}{L_z^2}\right), \tag{6a}$$

$$\nu_t^2 = \frac{c_t^2}{4}\left(\frac{n_x^2}{L_x^2} + \frac{n_y^2}{L_y^2} + \frac{n_z^2}{L_z^2}\right). \tag{6b}$$

There are three independent modes of vibration associated with each set of integers n_x, n_y, n_z, namely, two transverse waves of frequency ν_t and one longitudinal wave of frequency ν_l.

The quantities

$$\lambda_x = \frac{2L_x}{n_x}, \qquad \lambda_y = \frac{2L_y}{n_y}, \qquad \lambda_z = \frac{2L_z}{n_z} \tag{7}$$

[1] In an isotropic medium, the elastic constants c_{ij} (*cf.* footnote 1, p. 94) are related by the equations.

$$c_{11} = c_{22} = c_{33},$$
$$c_{12} = c_{23} = c_{31},$$
$$c_{44} = c_{55} = c_{66} = \frac{(c_{11} - c_{12})}{2}$$

and all others are zero. Thus only two constants are independent. The equations for propagation of a wave in an elastic medium degenerate to (4*a*) and (4*b*) in this case. See, for example, W. Voigt, *Lehrbuch der Kristallphysik*, p. 587 (Teubner, Leipzig, 1928); G. Joos, *Theoretical Physics* (G. E. Stechert & Company, New York, 1934).

are the wave lengths of the modes (5*a*) and (5*b*) measured along the x, y, and z axes, respectively. The reciprocals of these quantities give the number of waves per unit length, or *the wave-number components* σ_x, σ_y, and σ_z:

$$\sigma_x = \frac{n_x}{2L_x}, \qquad \sigma_y = \frac{n_y}{2L_y}, \qquad \sigma_z = \frac{n_z}{2L_z}. \tag{8}$$

In terms of these quantities, the frequencies (6*a*) and (6*b*) become

$$\begin{aligned} \nu_l &= c_l\sigma, \\ \nu_t &= c_t\sigma, \end{aligned} \tag{6c}$$

where

$$\sigma = \sqrt{\sigma_x^2 + \sigma_y^2 + \sigma_z^2}. \tag{9}$$

The standing wave (5) may be viewed as the sum of a set of traveling waves, for which the prototype function is the exponential function

$$Ae^{i2\pi(\sigma\cdot\mathbf{r} - \nu t)} \tag{10}$$

where A is a constant, σ is a constant vector having components σ_x, σ_y, and σ_z, ν is the frequency of the wave, and $\mathbf{r}$ is the position vector whose components are the Cartesian coordinates of space x, y, and z. At a given time the phase of (10) is constant at all points that satisfy the equation

$$\sigma \cdot \mathbf{r} = \text{constant},$$

that is, on planes orthogonal to the direction σ. Thus these planes are the wave fronts. The normal distance between planes that differ in phase by 2π clearly is $1/|\sigma|$. Hence, $|\sigma|$ is the reciprocal wave length or the wave number, that is, the number of waves per unit distance normal to the wave front. For this reason, σ is called the wave-number vector. The phase of (10) is constant at points that satisfy the equation

$$\sigma \cdot \mathbf{r} - \nu t = \text{constant},$$

which describes the motion of the wave front planes in the direction of σ with a velocity equal to $\nu/|\sigma|$. The standing wave (5) may be obtained by adding together the sixteen functions that can be derived from (10) by exchanging the signs of ν and of the three components of σ.

The number of modes of vibration having values of σ less than the value σ' may be computed conveniently in the following way:

Let us represent the continuous variables σ_x, σ_y, and σ_z in a three-dimensional coordinate system. The points defined by Eqs. (8) are then distributed throughout the positive octant of this reference system and are arranged at the mesh points of a simple rectangular lattice. The equation $\sigma = \sigma'$ defines a sphere in the same reference system, and

the volume of the part of this sphere lying in the positive octant is $\frac{1}{8}\pi\sigma'^3$. The total number of points (8) that lie in this octant is

$$L_x L_y L_z \tfrac{4}{3}\pi\sigma'^3, \tag{11}$$

since the average volume occupied by any one point is $1/8L_xL_yL_z$. Equation (11) may be written as

$$V\tfrac{4}{3}\pi\sigma'^3 \tag{12}$$

where V is the volume of the crystal.

There are three modes of vibration associated with each point (8); therefore, the total number of modes for which σ is less than σ' is

$$N(\sigma') = V4\pi\sigma'^3. \tag{13}$$

The equation $\nu = \nu'$ defines two spheres in δ space, one corresponding to longitudinal modes of vibration, and the other to transverse modes. The radii of these spheres are given by the equations

$$\sigma_l' = \frac{\nu'}{c_l}, \qquad \sigma_t' = \frac{\nu'}{c_t}, \tag{14}$$

respectively. Using Eq. (12), we find that the number of points in the first sphere is $\frac{4}{3}\pi\sigma_l'^3$, and the number in the second is $\frac{4}{3}\pi\sigma_t'^3$. Hence, the total number of modes having frequency less than ν' is

$$N(\nu') = V\frac{4\pi}{3}\left(\frac{1}{c_l^2} + \frac{2}{c_t^2}\right)\nu'^3, \tag{15}$$

since one longitudinal mode and two transverse modes correspond to each value of δ. We may define a mean velocity c by means of the equation

$$\frac{3}{c^2} = \frac{1}{c_l^2} + \frac{2}{c_t^2}, \tag{16}$$

and we may write Eq. (15) in the form

$$N(\nu') = V\frac{4\pi}{c^3}\nu'^3. \tag{17}$$

The corresponding expression for the case of a nonisotropic medium cannot be derived so simply, for the dependence of ν on the variables σ_x, σ_y, and σ_z usually will not be so simple as in (6*a*) and (6*b*).

Debye determined the highest allowed frequency ν_m, by setting (17) equal to $3N$ where N is the number of atoms in the solid. In addition, he assumed that the number of modes of vibration in the frequency range from ν' to $\nu' + d\nu'$ is given by the differential of (17), which is equivalent to assuming that the distribution function $f(\nu)$ in Eq. (9) of Sec. 18 is

$$f(\nu) = 3V\frac{4\pi}{c^3}\nu^2. \tag{18}$$

Under these conditions, Eq. (9) of Sec. 18 becomes

$$\bar{E} = \frac{V12\pi h}{c^3}\int_0^{\nu_m} \frac{\nu^3 d\nu}{e^{\frac{h\nu}{kT}} - 1}. \tag{19}$$

We may eliminate c by means of Eq. (17), after setting $N(\nu') = 3N$. $\bar{E}$ then is

$$\bar{E} = \frac{9Nh}{\nu_m^3}\int_0^{\nu_m} \frac{\nu^3 d\nu}{e^{\frac{h\nu}{kT}} - 1}. \tag{20}$$

Next, we may set $x = h\nu/kT$ and write (20) in the form

$$\bar{E} = 9Nh\nu_m\left(\frac{kT}{h\nu_m}\right)^4 \int_0^{\frac{h\nu_m}{kT}} \frac{x^3 dx}{e^x - 1}. \tag{21}$$

Thus, the molar heat is equal to

$$C_v = 9nR\left(\frac{kT}{h\nu_m}\right)^3 \int_0^{\frac{h\nu_m}{kT}} \frac{e^x x^4 dx}{(e^x - 1)^2}. \tag{22}$$

Following Debye, we shall define a characteristic temperature Θ_D by the

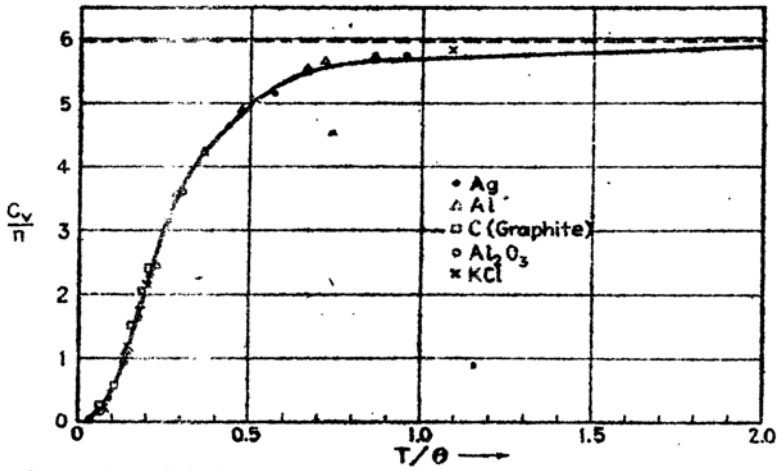

FIG. 3.—Comparison of the Debye specific-heat curve and the observed specific heats of a number of simple substances.

equation $k\Theta_D = h\nu_m$ and shall write (22) in the form

$$C_v = 3nRf_D\left(\frac{\Theta_D}{T}\right) \tag{23}$$

where

$$f_D\left(\frac{\Theta_D}{T}\right) = 3\left(\frac{T}{\Theta_D}\right)^3 \int_0^{\frac{\Theta_D}{T}} \frac{e^x x^4}{(e^x - 1)^2} dx. \tag{24}$$

The function $f_D(x)$ is shown in Fig. 1, along with Einstein's function. It

TABLE XXXIV.—THE DEBYE CHARACTERISTIC TEMPERATURES OF SOLIDS

Substance	Θ_D	Substance	Θ_D
	Metals		
Na	150	Al	390
K	100	Ga	125
Cu	315	In	100
Ag	215	Tl	100
Au	170	La	150
Fe	1,000	Ti	350
Mg	290	Zr	280
Ca	230	Hf	213
Sr	170	Ge	290
Zn	250	Sn	260
Cd	172	Pb	88
Hg	96		
		Sb	140
		Bi	100
Cr	485	Ta	245
Mn	350	W	310
Fe	420	Re	300
Co	385	Os	250
Ni	375	Ir	285
Mo	380	Pt	225
Ru	400		
Rh	370		
Pd	275		
	Ionic Crystals		
KCl	227	CaF_2	474
NaCl	281	FeS_2	645
KBr	177		
AgCl	183		
AgBr	144		

should be noted that f_D approaches zero as $(T/\Theta_D)^3$.

Figure 3 shows the manner[1] in which the atomic-heat curves of a number of simple solids may be fitted with a Debye function. It is obvious

[1] See, for example, the compilation of data in Landolt-Bornstein.

from Fig. 1 that the Einstein function would not apply nearly so well at low temperatures. Table XXXIV contains values of Θ_D that were determined from curves of this type.

The mean velocity c and the characteristic temperature are related by means of the equation

$$\Theta_D = \frac{h\nu_m}{k} = \frac{h}{k}c\left(\frac{3N}{4\pi V}\right)^{\frac{1}{3}}. \tag{25}$$

Born and von Kármán,[1] employing room-temperature values of the elastic constants, found excellent agreement between the observed and calculated temperatures of aluminum, copper, silver, lead, sodium chloride, and potassium chloride for temperatures above 25°K. Similar agreement has been found by Hopf and Lechner[2] for calcium fluoride and iron sulfide and by Schrödinger[3] for iron and carbon. Eucken,[4] however, showed that this close agreement usually disappears if the elastic constants for absolute zero are used instead. For example, Table XXXV contains a comparison of observed and calculated values

TABLE XXXV.—COMPARISON OF CHARACTERISTIC TEMPERATURES DETERMINED FROM THE T^3 LAW WITH THOSE COMPUTED FROM THE LOW-TEMPERATURE ELASTIC CONSTANTS

Substance	Θ_D (T^3 law)	Θ_D (calc.)
Cu	329	353
Ag	212	241
Al	399	502

of Θ_D for copper, silver, and aluminum. It should be noted that the Θ_D calculated from absolute-zero data are larger than the observed values. We may invariably expect this kind of discrepancy if the observed Θ_D agrees with the value computed from room-temperature data, since the elastic constants increase with decreasing temperature.

In addition to the discrepancy pointed out by Eucken, Grüneisen and Goens[5] have found that the Θ_D for zinc and cadmium determined from room-temperature data are larger than the experimental values. The disagreement would be emphasized even more in this case by using elastic constants corrected to absolute zero.

[1] BORN and VON KÁRMÁN, *op. cit.*
[2] L. HOPF and G. LECHNER, *Verh. deut. physik. Ges.*, **16**, 643 (1914).
[3] E. SCHRÖDINGER, *Handbuch der Physik*, Vol. X.
[4] A. EUCKEN, *Verh. deut. physik. Ges.*, **15**, 571 (1913).
[5] E. GRÜNEISEN and E. GOENS, *Z. Physik*, **26**, 250 (1924).

Thus, we may say that at least part of the excellent agreement between the observed specific-heat curves and Debye's curve is fortuitous.

c. Modifications of Debye's Equation Based upon the Continuum Hypothesis.—Several modifications of Debye's equation were suggested in the period following his paper. We shall discuss briefly the work of Born,[1] the theoretical basis of which is examined in Sec. 21.

In treating molecular solids, Born postulated that only those $3N$ degrees of freedom that correspond to *intermolecular* vibrations of a system of N molecules should be treated by the continuum method. The specific heat associated with the remaining $3(n-1)$ sets of N frequencies that correspond to intramolecular vibrations of the n atoms

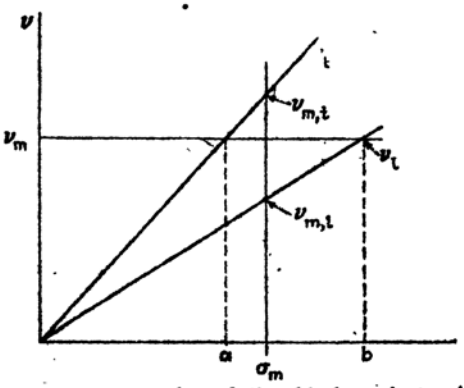

FIG. 4.—Frequency versus wave-number relationship for an isotropic continuous medium. The two lines represent curves for the two types of polarization. In the Debye scheme the $3N$ modes are selected by choosing all waves having frequency less than ν_m. Thus the transverse modes corresponding to the wave-number range a–b are excluded, whereas the longitudinal modes in the same range are retained. In the Born method all modes having wave numbers less than σ_m are retained, all others are discarded.

in a molecule should be approximated by $3(n-1)$ Einstein functions. In addition, he suggested that the continuum frequencies that are used should be selected by taking all modes of vibration associated with a region of wave-number space instead of restricting the range by means of Eq. (17). The difference between Born's and Debye's methods may be understood by considering an isotropic medium. The two lines of Fig. 4 represent the dependence of frequency on wave number for the transverse and longitudinal waves. In Debye's scheme, the modes are selected by taking all frequencies less than ν_m. Thus, waves having only one type of polarization are used in the range of σ extending from a to b. Born suggested that it would be preferable to use all waves for which σ is less than the value σ_m that is defined by the condition that $N(\sigma')$ in Eq. (13) be equal to $3N$; that is,

[1] M. BORN, *Dynamik der Kristallgitter* (Teubner, Leipzig, 1915).

$$\sigma_m = \left(\frac{3n_0}{4\pi}\right)^{\frac{1}{3}} \tag{26}$$

where n_0 is the molecular density. The limiting frequencies of the longitudinal and transverse waves then are

$$\begin{aligned} \nu_{m,l} &= c_l\sigma_m, \\ \nu_{m,t} &= c_t\sigma_m, \end{aligned} \tag{27}$$

and the density of each type of vibration is

$$\left.\begin{aligned} f_l(\nu) &= V\frac{4\pi\nu^2}{c_l^3} \qquad 0 \leq \nu \leq \nu_{m,l}, \\ f_t(\nu) &= 2V\frac{4\pi\nu^2}{c_t^3} \qquad 0 \leq \nu \leq \nu_{m,t}. \end{aligned}\right\} \tag{28}$$

The contribution to the specific heat from these modes is

$$R\left[f_D\left(\frac{\Theta_l}{T}\right) + 2f_D\left(\frac{\Theta_t}{T}\right)\right]$$

where $\Theta_l = h\nu_{m,l}/k$ and $\Theta_t = h\nu_{m,t}/k$. To this, Born would add $3(n-1)$ Einstein terms in order to include the intramolecular vibrations.

Born suggested that the same method of selecting frequencies could also be used in good approximation in anisotropic solids. In this case, the three waves of different polarization have different velocities for each direction of propagation. Thus, the limiting frequencies $\nu_{m,1}$, $\nu_{m,2}$, and $\nu_{m,3}$ are functions of the direction of propagation. If the direction variables are taken to be the polar angles θ and φ, the specific heat associated with the continuum modes of vibration now is

$$C_B = \frac{R}{4\pi}\sum_{j=1}^{3}\int_\Omega f_D\left(\frac{\Theta_j(\theta,\varphi)}{T}\right)d\Omega \tag{29}$$

where

$$\Theta_j(\theta,\varphi) = \frac{h\nu_{m,j}(\theta,\varphi)}{k} = \frac{hc_j(\theta,\varphi)}{k}\left(\frac{3n_0}{4\pi}\right)^{\frac{1}{3}} \tag{30}$$

and $d\Omega$ is the differential of solid angle. If we add to this the Einstein term associated with intramolecular vibrations, we obtain

$$C_V = C_B + \sum_{i=1}^{3(n-1)} Rf_E\left(\frac{h\nu_i}{kT}\right). \tag{31}$$

Försterling[1] computed C_B for the cubic salts sodium chloride, potassium chloride, calcium fluoride, and iron sulfide, using the room-temper-

[1] K. FÖRSTERLING, *Z. Physik*, **3**, 9 (1920); *Ann. Physik*, **61**, 549 (1920); *Z. Physik*, **8**, 251 (1922).

ature elastic constants. As we shall see in Sec. 23, it is doubtful whether Born's method of dividing the specific heat into Debye and Einstein terms can be justified in the case of these ideal ionic crystals. Försterling employed an ingenious expansion method for evaluating the integrals in Eq. (29) and found that the result could be approximated closely by means of three Debye functions. The three Einstein frequencies were taken to be the same in the diatomic case, whereas two sets of three equal values were used in the triatomic cases. These frequencies were chosen in order to give the best fit with the observed results. The three computed characteristic temperatures and the assumed vibrational frequencies are listed in Table XXXVI. Table XXXVII contains a

TABLE XXXVI

Substance	Calculated characteristic temperatures			Assumed frequencies	
	Θ_1	Θ_2	Θ_3	$h\nu_1/k$	$h\nu_2/k$
NaCl	354	216	194	218	
KCl	289	170	150	186	
CaF_2	558	348	274	306	502
FeS_2	631	439	404	406	620

TABLE XXXVII.—COMPARISON OF OBSERVED AND CALCULATED MOLAR HEAT OF POTASSIUM CHLORIDE
(In cal/mol)

T, °K	Observed	Calculated
70	7.54	7.52
87	8.66	8.63
137	10.36	10.38
235	11.46	11.38

comparison of the observed and calculated values of C_V for potassium chloride above 70°K. The accuracy is about the same in the other cases. The contributions from the Einstein terms decrease very rapidly at low temperatures and are negligible below 10°K. Hence, only the C_B term need be considered in this range. This term, however, varies as $T^3/\bar{\Theta}^3$ where $1/\bar{\Theta}^3$ is the mean of $1/\Theta^3(\theta, \varphi)$. Since Keesom and Clark have found that potassium chloride does not satisfy the T^3 law below 10°K, it follows that Born's modification of Debye's theory is subject to some of the same objections as the original theory.

More recently, Lord, Ahlberg, and Andrews[1] have applied a modification of Born's theory to crystalline benzene, which is an ideal molecular

[1] R. C. LORD, J. E. AHLBERG, and D. H. ANDREWS, *Jour. Chem. Phys.*, **5**, 649 (1937)

substance. In addition to using the continuum theory for the $3N$ intermolecular vibrations, they employed the theory for the $3N$ modes of motion in which the molecules undergo coupled torsional oscillations.[1] Einstein functions were employed for the remaining $27N$ intramolecular modes of vibration, the observed vibrational frequencies of the free benzene molecules being used in these terms. Since the intramolecular frequencies are comparatively high, the Einstein terms are negligible at 60°K and only the $6N$ continuum modes need be considered below

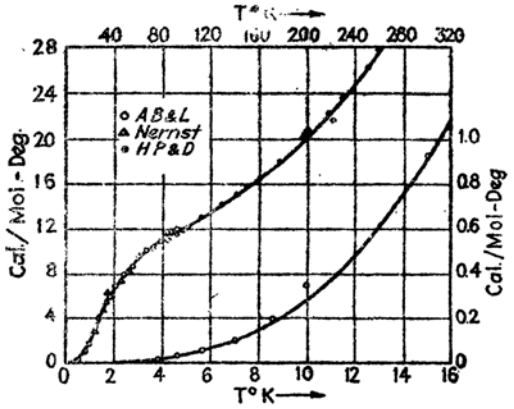

FIG. 5.—The observed and calculated molar heats of benzene. The lower curve represents the low-temperature heat, the upper curve corresponds to the higher temperature values. (*After Lord, Ahlberg and Andrews.*)

this temperature. These workers found that the molecular heat in this region could be fitted closely by the single Debye term

$$6Rf_D\left(\frac{\Theta_D}{T}\right)$$

in which $\Theta_D = 150°$. The agreement is shown in the lower half of Fig. 5. The observed and calculated molar heats for a wider range are given in the upper half of the same figure. Unfortunately, a rather questionable $C_P - C_V$ correction must be added above 150°K, so that the agreement in the higher temperature range is not so significant as that in the lower. It may be observed that the Einstein terms play an important role between 60° and 150°K.

The preceding method cannot be applied to a molecular crystal in a temperature range in which torsional oscillation changes to free rotation. We shall discuss the theory for this case in Chap. XIV.

[1] In a molecular crystal in which the inner molecular forces are much larger than the intermolecular forces, it may be expected that there are a set of $3N$ low-frequency torsional modes of vibration as well as the $3N$ low-frequency translational ones.

A striking example of a substance for which Born's equation (29) is more successful in predicting the specific heat than Debye's is the alkali metal lithium. The elastic constants of this substance have been computed by Fuchs,[1] using a method that will be described in Chap. X. The elastic constants of lithium have not been measured, but those of sodium have been measured by Quimby and Siegel.[2] The agreement between observed and calculated results in this case may be found in Table LXII, Chap. X. It should be observed that the relation

$$\frac{2c_{44}}{c_{11} - c_{12}} = 1$$

for isotropic media is far from being satisfied, showing that these crystals are much less isotropic than the alkali halides. Fuchs used the elastic

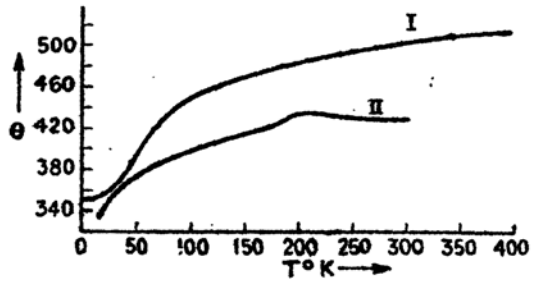

FIG. 6.—The observed and calculated $\Theta_D(T)$ curves of lithium. The theoretical curve I was calculated by Fuchs using Born's modification of Debye's theory and the theoretical elastic constants of lithium. The experimental curve II was determined by Simon and Swain.

constants of lithium to compute the three characteristic temperatures for each direction of propagation [*cf.* Eqs. (29) and (30)] and then computed the specific heat from these. Figure 6 gives a comparison of the observed and computed values of $\Theta(T)$ obtained by equating the observed and calculated specific-heat curves to a single Debye function. It is evident from curve II that Debye's law is inaccurate over a range of temperature comparable with the mean Debye temperature. The theoretical curve duplicates the general trend of the experimental curve and, hence, is in better agreement with it than the straight line corresponding to Debye's theory would be. However, the agreement is still far from exact. It is possible that Blackman's method of computing the specific-heat curve, which is discussed in Sec. 23, would further improve the theoretical values.[3]

[1] K. FUCHS, *Proc. Roy. Soc.*, **153**, 622 (1936); **157**, 444 (1936).

[2] S. L. QUIMBY and S. SIEGEL, *Phys. Rev.*, **54**, 293 (1938).

[3] R. C. Lord has pointed out to the writer that the specific heat of lithium may be fitted closely by a combination of Einstein and Debye functions, which suggests that the lattice may be somewhat molecular. This suggestion awaits X-ray confirmation.

20. Observed Deviations from Debye's Law of Vibrational Specific Heats.—The specific heats of a number of simple solids exhibit deviations from Debye's law that imply that Debye's theory requires modification. We shall exclude molecular crystals from consideration at the present time and shall discuss only the cases that involve metals or simple ionic crystals. The discrepancies may be placed under three headings, as follows:

a. The Linear Term in Metals.—Figure 19, Chap. I, shows that nickel has a linear specific-heat term in addition to the Debye term. A number of other metals show a similar anomaly near absolute zero that is now ascribed to the specific heat of their free electrons. We shall discuss a simple theory of this term in the next chapter.

b. Low-temperature Anomalies of Monotonically Increasing Curves.—Figure 68, Chap. I, shows the variation with temperature of the Debye characteristic temperature of potassium chloride. This curve was obtained by Keesom and Clark by equating the observed molecular heat to Debye's expression

$$C_V = 464.5\left(\frac{T}{\Theta_D}\right)^3$$

and solving for Θ_D. It may be observed that, instead of being a constant, this quantity shows a distinct peak about 4.3°K. As we mentioned in the previous section, Born's modification of Debye's theory cannot explain this because the Einstein terms in Eq. (29) are negligible in this temperature range. Deviations of this type have been explained qualitatively by Blackman. We shall discuss his work in Sec. 23.

c. The Anomalous Peaks of Germanium and Hafnium.—These peaks, which are shown in Fig. 16, Chap. I, cannot be explained by any theory that assumes Hooke's-law forces, for no superposition of Einstein functions leads to a curve with a maximum. Although the peaks in these metals are similar to the peaks occurring in molecular crystals, the explanation in terms of molecular rotation evidently cannot be applied.

21. The Vibrational Modes of One-dimensional Systems. *a. Monatomic Lattice.*—Suppose that we have a one-dimensional lattice of atoms extending along the x axis. We shall assume that the atoms are spaced at distances a from one another and that they interact with Hooke's-law forces. Consider the longitudinal modes of vibration from the following two standpoints: first, the Debye standpoint in which the modes are determined by solving the equation

$$\frac{1}{c_l^2}\frac{d^2\psi(x)}{dt^2} = \frac{d^2\psi(x)}{dx^2} \tag{1}$$

where $\psi(x)$ is the longitudinal displacement and c_l is the constant velocity of propagation of very long waves; and, the atomic standpoint in which the normal modes are determined by solving simultaneously the equations of motion for each atom.

The independent harmonic solutions of (1) for a string of length L fixed at both ends are

$$\psi(x) = A_n \sin \frac{2\pi n_x}{2L} \sin 2\pi\nu t \tag{2}$$

where n is an integer and the frequency ν is related to it by the equation

$$\nu = c_l \frac{n}{2L} = c_l\sigma. \tag{3}$$

It is obvious that $n/2L$ is the wave number σ of the standing wave. Since there are an infinite number of independent modes of vibration for a continuous string, we must limit the frequencies if we want to take account of the fact that there are a finite number of atoms in the string. This may be done by excluding those modes for which n is greater than N, the total number of atoms, that is, by neglecting modes with wave lengths less than $2a$.

Next, let us derive the equations of motion for the atoms of the string. We shall number the atoms consecutively from 1 to N, starting at the origin of coordinates, and we shall let x_n be the variable that measures the displacement of the nth atom from its equilibrium position. In addition, we shall assume that each atom interacts only with its immediate neighbors. Under this condition, the equation of motion of the nth atom is

$$m\frac{d^2x_n}{dt^2} = -\mu[(x_n - x_{n+1}) - (x_{n-1} - x_n)] \tag{4}$$

where μ is Hooke's constant for a pair of atoms and m is the mass of an atom.

The equations for the two end atoms obviously are different from the equations for interior atoms, for the former lack neighbors on one side. This fact could complicate the procedure of finding solutions of (4); however, we shall avoid the difficulty by using a method that was employed first by Born and von Kármán.[1] We shall assume that there are additional atoms on both sides of the string, in order that the equations for the end atoms may be the same as those for interior atoms. This assumption cannot affect the nature of the frequency-distribution function in any important way, as long as the number of atoms is sufficiently large. In addition, we shall assume that the phase of x_1 is the same as

[1] BORN and VON KÁRMÁN, *op. cit.*

the phase of the hypothetical $N + 1$st atom. As an alternative, we might assume that the end atoms are fixed. The "periodic" boundary condition is more advantageous, however, since it allows us to find elemental running-wave solutions without dealing with an infinite string and, at the same time, takes proper account of the number of degrees of freedom.

The function

$$x_n(l, \nu) = Ae^{2\pi i\left(\frac{l}{N}n - \nu t\right)}, \tag{5}$$

where A is a constant and l is an integer, satisfies the periodic boundary condition and reduces all of Eqs. (4) to the same form, namely,

$$-4\pi^2 m\nu^2 = -\mu 2\left(1 - \cos\frac{2\pi l}{N}\right). \tag{6}$$

Hence, the real and imaginary parts of (5) are physically interesting solutions of Eq. (4). Each of these parts defines a real traveling wave of wave length $\lambda = Na/l$, or wave number $\sigma = l/Na$, for which the end atoms move in phase. We may also construct real standing-wave solutions of fixed frequency ν' by properly combining the four functions of type (5) for which $l = \pm l'$ and $\nu = \pm\nu'$. The two independent waves of this type are

$$x_n = A \sin\frac{2\pi l'}{N}n \sin(\nu' t + \delta) \tag{7a}$$

and

$$x_n = A \cos\frac{2\pi l'}{N}n \sin(\nu' t + \delta') \tag{7b}$$

where δ and δ' are arbitrary phases.

The set of independent modes of different wave number are those corresponding to values of l for which the quantities (5) are different functions of n. Thus, the modes belonging to $l = l'$ and $l = l' + N$ are not independent, since $x_n(l', \nu)$ is equal to $x_n(l' + N, \nu)$ for all values of n. It is clear that there are only N independent values of l and that these may be chosen to be those lying in the range from zero to $N - 1$. The sine and cosine functions of $2\pi l'n/N$ appearing in (7a) and (7b) are different only for half this range of l, because of the relation

$$\sin x = -\sin(2\pi - x), \tag{8a}$$
$$\cos x = \cos(2\pi - x). \tag{8b}$$

Thus, if we choose to represent the normal modes by means of the functions (5), the independent range of l extends from zero to $N - 1$, whereas, if we choose the two functions (7a) and (7b) instead, the range extends

from zero to $N/2$. In either case, however, there are just N independent modes, in agreement with the fact that there are only N degrees of freedom in the system.

The relationship between frequency and wave number, given by Eq. (6), may be reduced to

$$\nu = \frac{1}{\pi}\sqrt{\frac{\mu}{m}}\left|\sin\frac{\pi l}{N}\right| = \frac{1}{\pi}\sqrt{\frac{\mu}{m}}\left|\sin \pi\sigma a\right|. \tag{9}$$

This equation becomes identical with Eq. (3) for values of σ small enough to allow us to replace the sine by its argument, for then

$$\nu = a\sqrt{\frac{\mu}{m}}\sigma.$$

Figure 7 illustrates the expression (9) for the independent range of the σ. This curve shows that it is not generally permissible to assume that the

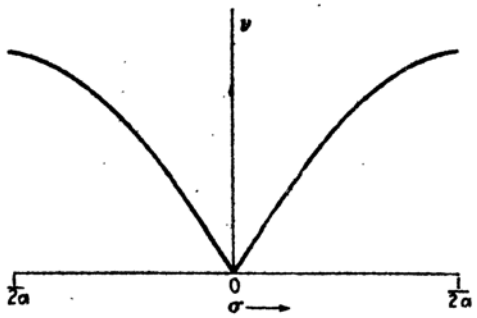

FIG. 7.—The $\nu(\sigma)$ curve for a monatomic linear lattice. The independent range of σ is taken to extend from $-1/2a$ to $1/2a$ corresponding to values of l extending from $-N/2$ to $N/2$.

velocity of propagation of elastic waves is independent of frequency, as one does in deriving Debye's frequency distribution.

Grüneisen and Goens[1] first suggested that one might explain the discrepancy between the characteristic temperatures obtained from the T^3 law and those computed from elastic data on the basis of differences in velocity between long and short waves. The fact that short waves travel more slowly in the linear lattice suggests that in computing Θ_D we should use elastic constants somewhat smaller than those obtained from ordinary measurements. Blackman has investigated more fully the significance of this suggestion (*cf.* Sec. 23).

We shall next find the frequency distribution corresponding to Eq. (9). Equal ranges of σ contain equal numbers of modes, since σ is proportional to l. Moreover,

[1] GRÜNEISEN and GOENS, *op. cit.*

$$d\sigma = \frac{dl}{Na}.$$

Hence, the density $dl/d\nu$ is

$$\frac{dl}{d\nu} = \frac{1}{d\nu/dl} = \frac{Na}{d\nu/d\sigma} = \frac{N}{\sqrt{\mu/m}\cos \pi\sigma a}$$
$$= \frac{N\sqrt{m/\mu}}{\sqrt{1 - (\nu^2\pi^2 m/\mu)}}.$$

Thus, the density is constant for small values of ν, just as in the continuous case, but it becomes infinite when $\sigma = 1/2a$.

b. Diatomic Lattices.—Let us extend the preceding problem by adding particles of mass M at points midway between the particles of mass m so that the distance between neighbors now is $a/2$. We shall label the masses with integers extending from 1 to $2N$ in this case so that the odd integers correspond to the masses m and the even ones to the masses M. The equations of motion are then

$$\left.\begin{aligned} m\frac{d^2x_{2n+1}}{dt^2} &= -\mu[(x_{2n+1} - x_{2n}) - (x_{2n+2} - x_{2n+1})], \\ M\frac{d^2x_{2n}}{dt^2} &= -\mu[(x_{2n} - x_{2n-1}) - (x_{2n+1} - x_{2n})], \end{aligned}\right\} \tag{10}$$

where μ is the force constant between unlike neighbors.

If the Born-von Kármán boundary conditions are used again, a normal coordinate substitution is

$$\left.\begin{aligned} x_{2n+1} &= Ae^{2\pi i\left(l\frac{2n+1}{2N} - \kappa\right)}, \\ x_{2n} &= Be^{2\pi i\left(l\frac{2n}{2N} - \kappa\right)}, \end{aligned}\right\} \tag{11}$$

for integer l. The wave number σ is l/Na in this case. A and B satisfy the homogeneous linear equations

$$\left.\begin{aligned} (4\pi^2 m\nu^2 - 2\mu)A + 2\mu\cos\left(2\pi\frac{l}{2N}\right)B &= 0, \\ 2\mu\cos\left(2\pi\frac{l}{2N}\right)A + (4\pi^2 M\nu^2 - 2\mu)B &= 0, \end{aligned}\right\} \tag{12}$$

which have a solution only for those values of ν that satisfy the compatibility equation

$$\begin{vmatrix} (4\pi^2 m\nu^2 - 2\mu) & 2\mu\cos\left(2\pi\frac{l}{2N}\right) \\ 2\mu\cos\left(2\pi\frac{l}{2N}\right) & (4\pi^2 M\nu^2 - 2\mu) \end{vmatrix} = 0, \tag{13}$$

the roots of which are

$$4\pi^2\nu^2 = \frac{\mu}{mM}\left(M + m \pm \sqrt{M^2 + m^2 + 2Mm\cos 2\pi\frac{l}{N}}\right) \qquad (14)$$

It is easy to show that all independent modes occur in the range of l extending from zero to $2N$ and that there are two modes for each value

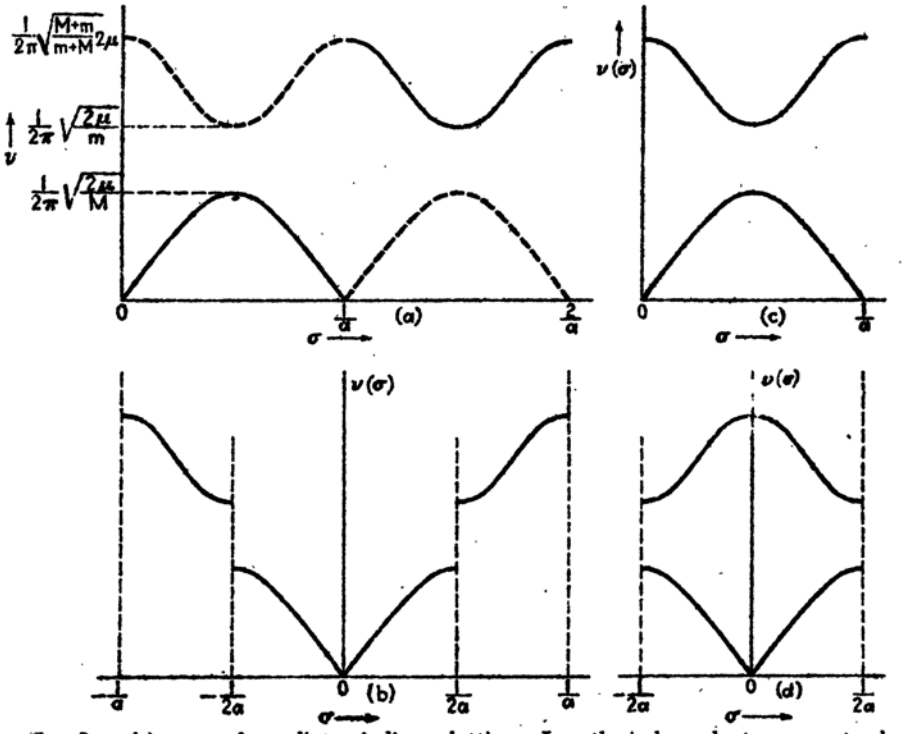

FIG. 8.—$\nu(\sigma)$ curves for a diatomic linear lattice. In *a* the independent range extends from 0 to $2/a$ and the function is single-valued. In *b* the independent range extends from $-1/a$ to $1/a$ so that the curve is symmetrical and single-valued. In *c* and *d* the function is double-valued, being unsymmetrical in the first case and symmetrical in the second. Case *d* is the reduced-zone scheme which will be used frequently in subsequent work.

of l, corresponding to the two roots of (14). Only half of these $4N$ modes are independent, however, since the factor $\cos 2\pi l/2N$ in Eqs. (12) repeats its values in the range of l extending from zero to $2N$.

We may obtain a one-to-one relationship between the modes of vibration and the $2N$ values of σ by arbitrarily restricting ν to be a single-valued function of σ, in the manner illustrated in Fig. 8*a*. The dotted curves show the discarded branches. It is sometimes convenient to choose the domain of σ to extend from $-\frac{1}{a}$ to $+\frac{1}{a}$ instead of from zero to $2/a$. In this case, the single-valued $\nu(\sigma)$ curve is symmetrical, as

in Fig. 8*b*. On the other hand, we may choose the range of σ to extend from zero to $1/a$, and we may use both branches of the frequency curve in the manner shown in Fig. 8*c*. It is also possible to obtain a symmetrical $\nu(\sigma)$ curve in this case by choosing σ so that it ranges from $-1/2a$ to $+1/2a$, as in Fig. 8*d*. We shall usually use this mode of description, which is called the "reduced-zone scheme."

The relationship between A and B may be determined by substituting ν^2 from Eq. (14) in either of the equations (12). The ratio A/B is

$$\frac{A}{B} = \frac{2M \cos 2\pi(l/2N)}{[M + m \pm \sqrt{M^2 + m^2 + 2Mm \cos (2\pi l/N)}]}. \tag{15}$$

It is interesting to note that the two normal modes which correspond to the points of discontinuity at $\sigma = \pm 1/2a$ are those in which one of the two types of mass is stationary. This fact may be shown by setting $l = N/2$ in Eq. (12). In addition, it should be noted that ν approaches zero linearly near the origin, just as in the monatomic case, a fact showing that acoustical waves travel with constant velocity.

The number of modes of vibration per unit range of σ is constant for each branch of the $\nu(\sigma)$ curve and is equal to Na. Hence, the distribution as a function of frequency is

$$f(\nu) = Na\frac{d\sigma}{d\nu},$$

which may be evaluated from Eq. (14). This function, which is shown in Fig. 9, has nonvanishing values over two ranges of frequency corresponding to the two branches of the $\nu(\sigma)$ curve.

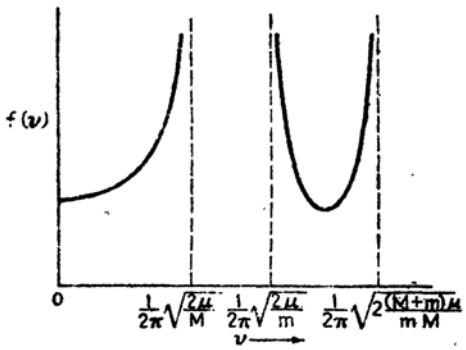

FIG. 9.—Diagrammatic plot of the frequency distribution of modes of vibration for the diatomic linear lattice. The density approaches infinity at the three points indicated by vertical dashed lines.

It may be seen that we obtain a new equation of type (12) for each particle added to the unit cell of the lattice and that a new root is simultaneously added to the determinantal equation that determines the frequency. Hence, if we want to retain ν as a single-valued function of σ, we must extend the domain of σ by $1/a$ for each particle added.

Another interesting one-dimensional case is that in which the two particles in the unit cell of a one-dimensional lattice have equal masses but interact more strongly with one another than with their neighbors. This case is analogous to that of a molecular crystal in which intramolecular forces are stronger than intermolecular ones. If α is the

Hooke's constant for interaction between the pairs in the unit cell, β the constant for interaction of one of these atoms with the neighbor in a different cell, and m the atomic mass, the equation analogous to (14) is $4\pi^2\nu^2 = \frac{2(\alpha+\beta) \pm \sqrt{4(\alpha+\beta)^2 - 16\alpha\beta \sin^2(2\pi l/2N)}}{2m}$. In the reduced-zone scheme, l is an integer extending from $-\frac{N}{2}$ to $\frac{N}{2}$, and the wave number is $\sigma = l/Na$, as before. When α becomes very large in comparison with β, the upper branch of this function reduces to the constant value

$$\nu = \frac{1}{\pi}\sqrt{\frac{\alpha}{2m}},$$

which is the natural frequency of two atoms having mass m and force constant β. The lower curve reduces to

$$\nu = \frac{1}{\pi}\sqrt{\frac{\beta}{2m}}\left|\sin\frac{\pi l}{N}\right|$$

under the same conditions. This expression is almost the same as (9); however, $2m$ appears instead of m. It is clear that the pair of atoms in the unit cell move as one in the modes associated with the low-frequency branch of the $\nu(\sigma)$ curve and oscillate as though in free space in the modes of the high-frequency branch. This condition obtains, of course, only when α is much larger than β. If α and β are comparable, the pair of atoms in the unit cell will behave less as an independent unit.

These simple one-dimensional examples illustrate most of the important features of the general three-dimensional case, which we shall discuss next. In particular, it is possible to see the origin of Born's postulates which were presented in part *c* of Sec. 19.

In connection with Born's first postulate, it should be pointed out that only the modes associated with the lowest branch of the $\nu(\sigma)$ curve approach acoustical vibrations at long wave lengths. Hence, this is the only branch we may reasonably expect to approximate by means of the $\nu(\sigma)$ curve for a continuous medium. There are no other branches for a monatomic lattice (case *a*), but there are others in polyatomic cases. The number of degrees of freedom in the lowest branch or in any single branch is equal to the number of unit cells in the lattice in the linear case and is equal to three times this number in the three-dimensional case. From this fact, Born concluded that it is not permissible to obtain the distribution of more than $3N$ modes by treating the solid as though it were continuous. He effectively assumed that the other branches of the $\nu(\sigma)$ curve are constants when he assumed that their contribution to the specific heat could be expressed in terms of Einstein

functions. This assumption is reasonable for molecular crystals, in which the intramolecular forces are much stronger than intermolecular ones (*cf.* the last one-dimensional example); but it cannot apply to crystals such as potassium chloride, in which there is high coordination between the atoms in the lattice.

Born's second assumption follows from the fact that in three dimensions each of the three branches of $\nu(\sigma)$, corresponding to the three directions of polarization, extends over the same range of wave numbers, just as the different branches do in the one-dimensional case.

22. General Three-dimensional Case.*—Let us now consider a general three-dimensional lattice[1] of atoms that interact with Hooke's-law forces. We shall discuss a crystal having translational vectors τ_1, τ_2, and τ_3 that are not necessarily orthogonal to each other. The unit cell may be taken as a rhombohedral parallelepiped the edges of which are determined by the translational vectors. The crystal specimen may have an arbitrary shape; we shall conveniently assume, however, that it is a rhombohedral parallelepiped of which the faces are parallel to the faces of the unit cell and the edges L_1, L_2, and L_3 are large integer multiples of the lengths of the edges of a unit cell; that is,

$$L_1 = N_1|\tau_1|, \qquad L_2 = N_2|\tau_2|, \qquad L_3 = N_3|\tau_3|. \tag{1}$$

The crystal obviously contains $N_1N_2N_3$ unit cells. A given cell may be specified relative to one corner of the crystal by the vector

$$\mathbf{T}(p_1,p_2,p_3) = p_1\tau_1 + p_2\tau_2 + p_3\tau_3, \tag{2}$$

which extends from the corner of the crystal to the corner of the cell in question. We shall call the cell that is specified by the integers p_1', p_2', p_3' the p'th cell in the lattice. In addition, we shall assume that there are n atoms in the unit cell and that the positions of these relative to the corner nearest the origin of coordinates are specified by the n position vectors $\varrho_\alpha(\alpha = 1, 2, \cdots, n)$. The position vector $\mathbf{r}_\alpha(p_1,p_2,p_3)$, of the αth atom in the pth unit cell relative to the origin is then given by the sum

$$\mathbf{r}_\alpha(p_1,p_2,p_3) = \mathbf{T}(p_1,p_2,p_3) + \varrho_\alpha.$$

* This section may be omitted in a first reading. It is demonstrated that the general conclusions that were drawn from the one-dimensional case concerning the $\nu(\sigma)$ curves also apply to the three-dimensional one with the difference that the wave number is now a three-dimensional vector whose independent values range throughout a polyhedron.

[1] For additional details concerning the problem of determining the vibrational modes of lattices, see R. B. Barnes, R. R. Brattain, and F. Seitz, *Phys. Rev.*, **48**, 582 (1935); R. R. Lyddane and K. F. Herzfeld, *Phys. Rev.*, **54**, 846 (1938).

We shall designate the coordinates of the αth atom in the pth unit cell relative to its equilibrium position by $x_\alpha{}^i(p_1,p_2,p_3)$ where i takes values 1, 2, and 3 corresponding to each of three orthogonal Cartesian coordinates. The equations of motion for this atom have the form

$$m_\alpha \frac{d^2 x_\alpha{}^i(p_1,p_2,p_3)}{dt^2} = f_\alpha{}^i(p) \tag{3}$$

where m_α is the mass of the αth atom and $f_\alpha{}^i(p)$ is a homogeneous linear function of the relative displacements of the α-pth atom and all other atoms:

$$f_\alpha{}^i = \sum_{j,\beta,p'} \mu_{\alpha\beta}{}^{ij}(p,p')[x_\beta{}^j(p') - x_\alpha{}^j(p)]. \tag{4}$$

We shall consider the interaction with all neighbors rather than, as in the one-dimensional case, merely the interaction with nearest neighbors. The coefficients $\mu_{\alpha\beta}{}^{ij}(p,p')$, which obviously are the interatomic Hooke constants, depend only on the relative values of p' and p for given α, β, i, j, that is, only on the relative equilibrium positions of two atoms. Thus, the equations of motion for the atoms in the pth cell differ from those for equivalent atoms in the qth cell only by the translation $\mathbf{T}(p_1,p_2,p_3) - \mathbf{T}(q_1,q_2,q_3)$. We shall make no assumptions about the number of physically important terms that appear in the right-hand side of (4), although we may expect that nearest neighbors will have the largest coefficients.

Since we shall generally be interested in forces that may be expressed as the gradient of a potential, we may require that

$$f_\alpha{}^i(p) = -\frac{\partial V_i}{\partial x_\alpha(p)}. \tag{5}$$

The necessary and sufficient conditions for these equations to be satisfied are

$$\frac{\partial f_\alpha{}^i(p)}{\partial x_\beta{}^j(p')} = \frac{\partial f_\beta{}^j(p')}{\partial x_\alpha{}^i(p)},$$

or, as one may readily see,

$$\mu_{\alpha\beta}{}^{ij}(p,p') = \mu_{\beta\alpha}{}^{ji}(p',p), \tag{6a}$$

$$\sum_{\beta,p'} \mu_{\alpha\beta}{}^{ij}(p,p') = \sum_{\gamma,p''} \mu_{\alpha\gamma}{}^{ji}(p,p''). \tag{6b}$$

In order to introduce the Born-von Kármán boundary conditions, we shall assume that surface atoms of the crystal have the same equations of motion as the interior atoms. We are then able to show that the

$3nN_1N_2N_3$ equations (3) may be reduced to $3n$ independent equations. The appropriate substitution that satisfies the Born-von Kármán boundary conditions is, in fact,

$$x_\alpha{}^i(p_1,p_2,p_3) = \frac{\xi_\alpha{}^i(\delta)}{\sqrt{m_\alpha}} e^{2\pi i(\sigma\cdot\mathbf{r}_\alpha(p_1,p_2,p_3)-\nu t)} \tag{7}$$

where the $\xi_\alpha{}^i$ are constants and δ is a vector that satisfies the conditions

$$\left.\begin{aligned} \delta\cdot\mathbf{T}(N_1,0,0) &\equiv N_1\delta\cdot\tau_1 = l_1,\\ \delta\cdot\mathbf{T}(0,N_2,0) &\equiv N_2\delta\cdot\tau_2 = l_2,\\ \delta\cdot\mathbf{T}(0,0,N_3) &\equiv N_3\delta\cdot\tau_3 = l_3, \end{aligned}\right\} \tag{8}$$

where l_1, l_2, and l_3 are integers. The solution of Eqs. (8), which obviously express the Born-von Kármán boundary conditions, is

$$\delta = \frac{l_1}{N_1}\frac{\tau_2\times\tau_3}{|\tau_1\tau_2\tau_3|} + \frac{l_2}{N_2}\frac{\tau_3\times\tau_1}{|\tau_1\tau_2\tau_3|} + \frac{l_3}{N_3}\frac{\tau_1\times\tau_2}{|\tau_1\tau_2\tau_3|} \tag{8a}$$

where the cross designates the conventional vector product and $|\tau_1\tau_2\tau_3|$ is the determinant

$$\begin{vmatrix} t_{11} & t_{21} & t_{31}\\ t_{12} & t_{22} & t_{32}\\ t_{13} & t_{23} & t_{33} \end{vmatrix}.$$

Equation (8a), like Eq. (2), defines the mesh points of a lattice having primitive translation vectors $\mathbf{s}_1$, $\mathbf{s}_2$, and $\mathbf{s}_3$, where

$$\mathbf{s}_1 = \frac{\tau_2\times\tau_3}{N_1|\tau_1\tau_2\tau_3|}, \qquad \mathbf{s}_2 = \frac{\tau_3\times\tau_1}{N_2|\tau_1\tau_2\tau_3|}, \qquad \mathbf{s}_3 = \frac{\tau_1\times\tau_2}{N_3|\tau_1\tau_2\tau_3|}. \tag{9}$$

This lattice has been called the reciprocal lattice by Gibbs, who first used it.

Let us consider for a moment the simple lattice consisting of one type of atom, say the αth in the unit cell. When the $\xi_\alpha{}^i$ ($i = 1, 2, 3$) are fixed, expression (6) describes a running wave for which the wave-number vector is δ. There are $N_1N_2N_3$ ways of choosing the l_i in (8) to give independent waves, and these may be selected to lie in the range

$$-\frac{N_1}{2} \le l_1 \le \frac{N_1}{2}, \qquad -\frac{N_2}{2} \le l_2 \le \frac{N_2}{2}, \qquad -\frac{N_3}{2} \le l_3 \le \frac{N_3}{2}.$$

This domain obviously corresponds to values of the wave-number vector δ that lie within the rhombohedral parallelepiped having corners at the eight points

$$\delta = \pm\frac{N_1}{2}\mathbf{s}_1 \pm \frac{N_2}{2}\mathbf{s}_2 \pm \frac{N_3}{2}\mathbf{s}_3. \tag{10}$$

The waves (7) are independent for a larger domain of $\mathfrak{d}$ if we consider all atoms at once; in fact, it is easy to see that the domain may be chosen to be n times larger. However, the difference between modes for which $\mathfrak{d}$ lies inside (10) and those for which it does not is simply a matter of the relative phases of the motions of different atoms in the unit cell. Hence, we may condense the larger domain of $\mathfrak{d}$ to the domain determined by (10) by absorbing these phases in the ξ^i and introducing n sets of ξ^i instead of one. This procedure is analogous to the one used in passing from the description of Fig. 8*b* to the reduced-zone scheme of Fig. 8*d* in the one-dimensional case.

We shall frequently use the relation

$$\sum_p e^{2\pi i(\sigma-\sigma')\cdot r_\alpha(p)} = N_1N_2N_3\delta_{\sigma,\sigma'} \tag{11}$$

where the summation extends over all $N_1N_2N_3$ values of p, and $\delta_{\sigma,\sigma'}$, the Kronecker delta function, is zero when $\mathfrak{d} \neq \mathfrak{d}'$ and is equal to 1 when $\mathfrak{d} = \mathfrak{d}'$.

Substituting (7) in Eq. (3) and multiplying the result by $e^{-2\pi i[\sigma\cdot r_\alpha(p')-\nu t]}$, we obtain the following $3n$ equations for the $\xi_\alpha{}^i$ after summing over p':

$$-4\pi^2\nu^2\xi_\alpha{}^i = g_\alpha{}^i(\xi_1^1, \xi_1^2, \cdots, \xi_n^3, \mathfrak{d}). \tag{12}$$

$g_\alpha{}^i$ clearly is a homogeneous linear function of the $\xi_\alpha{}^i$ and of the differences $\mathbf{r}_\alpha(p) - \mathbf{r}_\beta(q)$. We have, in fact,

$$\begin{aligned} g_\alpha{}^i &= \sum_{j,\beta,p'} \frac{\mu_{\alpha\beta}{}^{ij}}{\sqrt{m_\alpha}}\left(\frac{\xi_\beta{}^j}{\sqrt{m_\beta}}e^{2\pi i\sigma\cdot[r_\beta(p')-r_\alpha(p)]} - \frac{\xi_\alpha{}^j}{\sqrt{m_\alpha}}\right) \\ &\equiv \sum_{j,\beta}\lambda_{\alpha\beta}{}^{ij}(\mathfrak{d})\xi_\beta{}^j \end{aligned} \tag{13}$$

where

$$\lambda_{\alpha\beta}{}^{ij}(\mathfrak{d}) = \sum_{p'}\frac{\mu_{\alpha\beta}{}^{ij}(p,p')}{\sqrt{m_\alpha m_\beta}}e^{2\pi i\sigma\cdot[r_\beta(p')-r_\alpha(p)]} \tag{14a}$$

for $\alpha \neq \beta$ and $p \neq p'$, and

$$\lambda_{\alpha\alpha}{}^{ij}(\mathfrak{d}) = -\frac{1}{\gamma n_\alpha}\sum_{\beta\neq\alpha,p'\neq p}\mu_{\alpha\beta}{}^{ij}(p,p'). \tag{14b}$$

The $\lambda_{\alpha\beta}$ are real only in the special case in which each atom is a center of symmetry, that is, in the case in which there is an atom of type β at the position $-[r_\beta(p') - r_\alpha(p)]$, relative to atom α in the p'th cell, for each atom at the position $+[r_\beta(p') - r_\alpha(p)]$. This condition is

satisfied in some of the simplest lattices, but the $\lambda_{\alpha\beta}$ are usually complex. We see, however, that

$$\lambda_{\alpha\beta}{}^{ij}(\mathfrak{d}) = \lambda_{\beta\alpha}{}^{ji*}(\mathfrak{d}), \tag{15}$$

because of Eqs. (6*a*) and (6*b*). Hence, the matrix of the $\lambda_{\alpha\beta}$ is Hermitian.

The homogeneous linear equations (12), namely,

$$-4\pi\nu^2\xi_\alpha{}^i(\mathfrak{d}) = \sum_{j,\beta}{}' \lambda_{\alpha\beta}{}^{ij}(\mathfrak{d})\xi_\beta{}^j(\mathfrak{d}), \tag{16}$$

will have solutions only for those values of ν which satisfy the usual determinantal condition. This secular equation has $3n$ real roots ν_t^2 ($t = 1, \cdots, 3n$), since $\lambda_{\alpha\beta}$ is Hermitian. The $3n$ independent sets of $\xi_\alpha{}^i$, which we may distinguish by a superscript t, satisfy the orthogonality relation

$$\sum_{i,\alpha}{}' \xi_{\alpha,t}{}^i(\mathfrak{d})\xi_{\alpha,t'}{}^{i*}(\mathfrak{d}) = 0 \tag{17}$$

when $t \neq t'$. These coefficients may be normalized in such a way that this sum is unity when $t = t'$; that is,

$$\sum_{i'}{}' \xi_{\alpha,t}{}^i(\mathfrak{d})\xi_{\alpha,t'}{}^{i*}(\mathfrak{d}) = \delta_{t,t'}. \tag{17a}$$

Thus, we see that in all there are $3nN_1N_2N_3$ independent functions of type (7) when we include the $N_1N_2N_3$ independent values of $\mathfrak{d}$.

The $\xi_\alpha{}^i$ are usually complex when the $\lambda_{\alpha\beta}$ are complex. We may obtain physically interesting real functions, however, by taking the real and imaginary parts of the quantities $\xi_{\alpha,t}{}^i(\mathfrak{d})e^{2\pi i[\sigma\cdot\mathbf{r}_\alpha(p)-\nu(\sigma)t]}$. This procedure does not double the modes of vibration, since spatial parts of the $3n$ complex functions associated with $-\mathfrak{d}$ are the complex conjugates of those associated with $+\mathfrak{d}$, as may be seen by taking the complex conjugate of Eq. (16). The real functions have the form

$$x_\alpha{}^i(p) = a_\alpha{}^i \sin 2\pi[\mathfrak{d}\cdot\mathbf{r}_\alpha(p) - \nu t] + b_\alpha{}^i \cos 2\pi[\mathfrak{d}\cdot\mathbf{r}_\alpha(p) - \nu t]$$

where $a_\alpha{}^i$ and $b_\alpha{}^i$ are now real. Thus the motion of the αth atom is described by the vector

$$\mathbf{A}_\alpha \sin 2\pi[\mathfrak{d}\cdot\mathbf{r}_\alpha(p) - \nu t] + \mathbf{B}_\alpha \cos 2\pi[\mathfrak{d}\cdot\mathbf{r}_\alpha(p) - \nu t], \tag{18}$$

where $\mathbf{A}_\alpha$ and $\mathbf{B}_\alpha$ are the vectors

$$\mathbf{A}_\alpha = \begin{pmatrix} a_\alpha{}^1 \\ a_\alpha{}^2 \\ a_\alpha{}^3 \end{pmatrix}, \qquad \mathbf{B}_\alpha = \begin{pmatrix} b_\alpha{}^1 \\ b_\alpha{}^2 \\ b_\alpha{}^3 \end{pmatrix}. \tag{19}$$

If **A** and **B** are not equal and if neither is zero, the quantity (18) describes harmonic motion in an elliptical path, the plane of the ellipse being the plane containing the two vectors **A** and **B**. The direction of the major and minor axes may be found in the following way: Let us choose Cartesian coordinates x and y in the plane determined by **A** and **B**. Then, the motion of the particle is given by

$$x = A_x \sin\alpha + B_x \cos\alpha = \sqrt{A_x^2 + B_x^2}\,\sin\left(\alpha + \tan^{-1}\frac{B_x}{A_x}\right),$$
$$y = A_y \sin\alpha + B_y \cos\alpha = \sqrt{A_y^2 + B_y^2}\,\cos\left(\alpha - \tan^{-1}\frac{A_y}{B_y}\right),$$

where $\alpha = 2\pi[\boldsymbol{\sigma}\cdot\mathbf{r}_a(p) - \nu t]$ and A_x, A_y and B_x, B_y are the components of **A** and **B**. The condition that the ellipse should be in normal form is that

$$\frac{B_x}{A_x} = -\frac{A_y}{B_y}. \tag{20}$$

If θ is the angle between **B** and **A** and if ϕ is the angle that **A** makes with the x axis, we have

$$A_x = |\mathbf{A}|\cos\phi,$$
$$A_y = |\mathbf{A}|\sin\phi,$$
$$B_x = |\mathbf{B}|\cos(\theta+\phi),$$
$$B_y = |\mathbf{B}|\sin(\theta+\phi),$$

and (20) becomes

$$|\mathbf{B}|^2 \sin 2(\theta+\phi) = -|\mathbf{A}|^2 \sin 2\phi,$$

which determines ϕ.

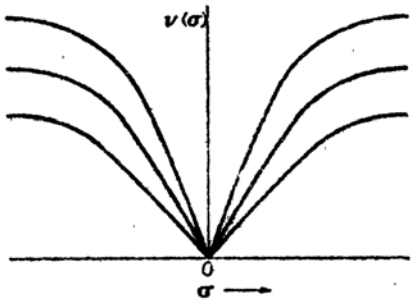

FIG. 10.—Schematic diagram of the $\nu(\sigma)$ curves of a monatomic three-dimensional lattice. Actually the values of σ range over a three-dimensional zone rather than a one-dimensional one so that this corresponds to the $\nu(\sigma)$ relation for a line in σ space that passes through the origin. It should be noted that the $\nu(\sigma)$ relation is linear near the origin.

Thus, we see that the real normal modes usually describe elliptically polarized elastic waves. They are plane-polarized in the special case in which **A** or **B** is zero, that is, when $\xi_a{}^i$ is real.

The $\lambda_{\alpha\beta}$ that appear in (16) are continuous functions of $\boldsymbol{\sigma}$. Hence, it may be expected that the frequencies $\nu(\boldsymbol{\sigma})$ are continuous functions of $\boldsymbol{\sigma}$ in the domain (10). In general, there are $3n$ continuous branches of the frequency curve. These branches may coincide for certain values of $\boldsymbol{\sigma}$, particularly if the lattice has a high degree of symmetry. In analogy with the one-dimensional case, there are three branches of the $\nu(\boldsymbol{\sigma})$ curve that approach zero linearly as $\boldsymbol{\sigma}$ approaches zero. The long wave-

length modes of this type correspond to ordinary acoustical waves which travel with constant velocity. Thus, we may expect that the three branches of $\nu(\sigma)$ will behave schematically, as illustrated in Fig. 10, if there is one atom per unit cell. Similarly, the six branches will behave as illustrated in Fig. 11, if there are two atoms in the unit cell, etc.

Before leaving this treatment of the three-dimensional problem, we shall derive expressions for the total energy and for the Lagrangian and Hamiltonian functions. In terms of the $\dot{x}_\alpha{}^i(p)$, the kinetic energy is

$$T = \sum_{i,\alpha,p} \frac{m_\alpha}{2} [\dot{x}_\alpha{}^i(p)]^2. \tag{21}$$

The potential energy, which is defined by Eqs. (5), is given by the expression

$$-V = \tfrac{1}{2} \sum_{i,\alpha,p} \sum_{j,\beta,p'} \mu_{\alpha\beta}{}^{ij}(p,p')[x_\beta{}^j(p') - x_\alpha{}^i(p)]x_\alpha{}^i(p)$$

$$\equiv \tfrac{1}{2} \sum_{i,\alpha,p} f_\alpha{}^i(p) x_\alpha{}^i(p). \tag{22}$$

Let us now express $x_\alpha{}^i(p)$ as a series of the form

$$x_\alpha{}^i(p) = \sum_{t,\sigma} a_t(\sigma) \frac{\xi_{\alpha,t}{}^i(\sigma)}{\sqrt{m_\alpha N}} e^{2\pi i \sigma \cdot r_\alpha(p)} \tag{23}$$

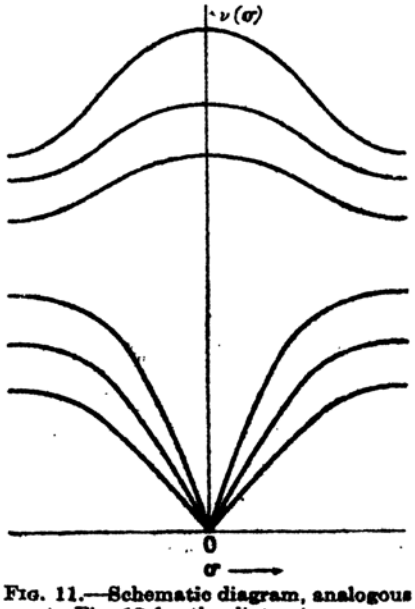

FIG. 11.—Schematic diagram, analogous to Fig. 10 for the diatomic case.

where N is the total number of unit cells and $a_t(\sigma)$ is the time-dependent amplitude of the tth mode having wave number σ. We shall assume that the ξ are normalized in the sense of Eq. (17a). The amplitude $a_t(\sigma)$ may then be expressed in terms of the $x_\alpha{}^i(p)$ in the following way. If Eq. (23) is multiplied by

$$\sqrt{\frac{m_\alpha}{N}} \xi_{\beta,t'}{}^{i*}(\sigma') e^{-2\pi i \sigma' \cdot r_\alpha(p)} \tag{24}$$

and the result is summed over i, α, and p, the right-hand side reduces to $3na_{t'}(\sigma')$ because of the orthogonality relations. Hence,

$$a_{t'}(\sigma') = \frac{1}{3n} \sum_{i,\alpha,p} \sqrt{\frac{m_\alpha}{N}} x_\alpha{}^i(p) \xi_{\alpha,t'}{}^{i*}(\sigma') e^{-2\pi i \sigma' \cdot r_\alpha(p)}. \tag{25}$$

Since the secular equations for σ and $-\sigma$ are identical, we may choose the index t in such a way that

$$\xi_{\alpha,t}{}^i(-\sigma) = \xi_{\alpha,t}{}^{i*}(\sigma) \tag{26}$$

and write Eq. (23) in the form

$$x_\alpha^i(p) = \frac{1}{\sqrt{m_\alpha N}} \sum_{\sigma,t}'' [a_t(\mathfrak{s})\xi_{\alpha,t}^i(\mathfrak{s})e^{2\pi i\sigma\cdot r_\alpha(p)} + a_t(-\mathfrak{s})\xi_{\alpha,t}^{i*}(\mathfrak{s})e^{-2\pi i\mathfrak{s}\cdot r_\alpha(p)}] \tag{27}$$

where Σ'' implies that the summation of $\mathfrak{s}$ is carried out in such a way that only one of the two points $\pm\mathfrak{s}$ is counted. We may now specify that

$$a_t^*(\mathfrak{s}) = a_t(-\mathfrak{s}) \tag{28}$$

in order that (27) may be real. If (27) is substituted into Eqs. (21) and (22), it is found that these equations reduce to

$$\left.\begin{aligned} T &= \sum_{\sigma,t}'' \dot{a}_t(\mathfrak{s})\dot{a}_t^*(\mathfrak{s}), \\ V &= \sum_{\sigma,t}'' 4\pi^2\nu_t^2(\mathfrak{s})a_t(\mathfrak{s})a_t^*(\mathfrak{s}), \end{aligned}\right\} \tag{29}$$

because of the orthogonality conditions. Now it is convenient to replace the complex variable $a_t(\mathfrak{s})$ by the two real variables

$$[a_t(\mathfrak{s}) + a_t^*(\mathfrak{s})]\sqrt{2}, \tag{30a}$$

$$[a_t(\mathfrak{s}) - a_t^*(\mathfrak{s})]\frac{\sqrt{2}}{i}. \tag{30b}$$

We shall do this in such a way that the $3n$ variables (30a) are associated with the point $\mathfrak{s}$ and the $3n$ variables (30b) are associated with the point $-\mathfrak{s}$; that is, we shall introduce real variables $\alpha_t(\mathfrak{s})$ defined by the relations

$$\left.\begin{aligned} \alpha_t(\mathfrak{s}) &= \frac{[a_t(\mathfrak{s}) + a_t^*(\mathfrak{s})]}{\sqrt{2}}, \\ \alpha_t(-\mathfrak{s}) &= \frac{[a_t(\mathfrak{s}) - a_t^*(\mathfrak{s})]}{i\sqrt{2}}. \end{aligned}\right\} \tag{31}$$

Since the inverses of the equations are

$$\left.\begin{aligned} a_t(\mathfrak{s}) &= \frac{\alpha_t(\mathfrak{s}) + i\alpha_t(-\mathfrak{s})}{\sqrt{2}}, \\ a_t^*(\mathfrak{s}) &= \frac{\alpha_t(\mathfrak{s}) - i\alpha_t(-\mathfrak{s})}{\sqrt{2}}, \end{aligned}\right\} \tag{32}$$

Eqs. (29) may be transformed to

$$\left.\begin{aligned} T &= \frac{1}{2}\sum_{t,\sigma} \dot{\alpha}_t^2(\mathfrak{s}), \\ V &= \frac{1}{2}\sum_{t,\sigma} 4\pi^2\nu_t^2(\mathfrak{s})\alpha_t^2(\mathfrak{s}), \end{aligned}\right\} \tag{33}$$

where δ is now summed over the entire zone of wave numbers. Thus, the Lagrangian function of the system is

$$L = T - V = \frac{1}{2}\sum_{t,\sigma}[\dot{\alpha}_t^2(\delta) - 4\pi^2\nu_t^2(\delta)\alpha_t^2(\delta)], \tag{34}$$

whereas the Lagrangian equations are

$$\ddot{\alpha}_t(\delta) = -4\pi^2\nu_t^2(\delta)\alpha_t(\delta), \tag{35}$$

which are the same as those for a linear harmonic oscillator, as we should expect.

The momentum variables $p_t(\delta)$ are defined in terms of the Lagrangian function by the equations

$$p_t(\delta) = \frac{\partial L}{\partial \dot{\alpha}_t(\delta)} \equiv \dot{\alpha}_t(\delta). \tag{36}$$

Thus, the Hamiltonian function is

$$\left.\begin{aligned} H &= \sum_{t,\sigma} p_t(\delta)\dot{\alpha}_t(\delta) - L \\ &= \frac{1}{2}\sum_{t,\sigma}[p_t^2(\delta) + 4\pi^2\nu_t^2(\delta)\alpha_t^2(\delta)]. \end{aligned}\right\} \tag{37}$$

23. Blackman's Computations.—Blackman[1] has determined by direct computation the frequency-distribution function of the normal modes of several simple lattices and has used the results to determine specific heats. The lattices he treated do not correspond to actual cases; however, the consequences of these computations make it seem reasonable that the discrepancies between experiment and the Debye theory which we listed under *b* in Sec. 20 may be explained, at least in part, by an extension of his work.

We shall discuss the results of two of Blackman's computations, namely, those for the linear lattice with two different atoms and those for a three-dimensional simple cubic lattice.

a. The Linear Lattice.—In Sec. 21, we derived the expression for the energy of the vibrational modes of a one-dimensional lattice containing two atoms. The result is

$$4\pi^2\nu^2(\delta) = \frac{\mu}{mM}(M + m \pm \sqrt{M^2 + m^2 + 2mM\cos 2\pi\delta a}) \tag{1}$$

where δ is the wave number, which may be taken to extend from $-1/2a$ to $1/2a$. The distribution of modes as a function of frequency is shown in

[1] BLACKMAN, *op. cit.*

Fig. 9. We may compare the specific heat derived by the use of this distribution function $f(\nu)$ with that derived by the use of the distribution for a continuous string. It is easy to show from the results of the first part of a in Sec. 21 that the distribution function in the second case is

$$f_c(\nu) = \frac{2Na}{c_l}.$$

If we restrict the domain of ν so that the number of modes is $2N$, then

$$f_c(\nu) = \frac{2N}{\nu_m} \qquad 0 \leq \nu \leq \nu_m \tag{2}$$

from which we obtain a one-dimensional Debye specific heat

$$C_v\left(\frac{\Theta}{T}\right) = 2Nk\left(\frac{T}{\Theta}\right)\int_0^{\frac{\Theta}{T}} \frac{x^2e^x}{(e^x-1)^2}dx \tag{3}$$

where $\Theta = h\nu_m/k$.

Blackman evaluated the specific heat, using the actual distribution function for the one-dimensional lattice, and then equated this to (3) at each temperature and obtained a value of Θ. The maximum frequency of the lattice was chosen in each case so that $h\nu_m/k = 200°$. The dependence of Θ on temperature for several values of the ratio m/M is shown in Fig. 12. It may be observed that the curves approach a nearly constant value at high temperatures but deviate considerably in the range below 40°K. The details near $T = 0$ are not given exactly in this figure; more precise computations show that the $\Theta(T)$ curve for $m/M = 8$ has a minimum near 10°K.

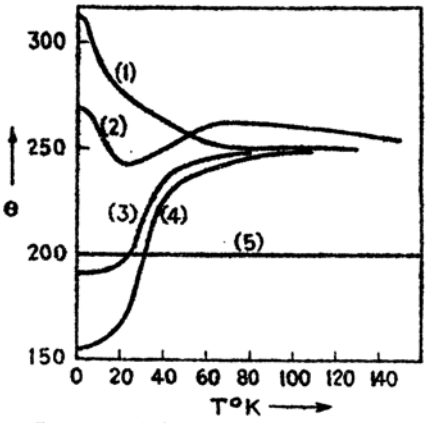

FIG. 12.—$\Theta(T)$ curves for a one-dimensional lattice. (1) $m/M = 1$, (2) $m/M = 3$, (3) $m/M = 8$, (4) $m/M = 13$, (5) continuum. (*After Blackman.*)

These curves leave little room for doubt that the deviations from Debye's distribution function can be important.

b. Simple Cubic Lattice.—Blackman carried through a similar computation for a monatomic simple cubic lattice (*cf.* Fig. 13). He assumed that each atom interacts only with its six nearest neighbors and twelve next nearest neighbors. For simplicity, he fixed the ratio γ/α of the

force constants at 0.05, where α is the Hooke's constant for nearest neighbors and γ is that for next nearest neighbors.

There is one particle per unit cell in this case; hence, the secular equation is of third degree, the three roots corresponding to the three directions of polarization of vibrational waves. It turns out that the three branches of the $\nu(\delta)$ curve meet at $\delta = 0$ and at the eight corners of the cube defined by

$$\sigma_x = \pm 1/2a, \sigma_y = \pm 1/2a, \sigma_z = \pm 1/2a,$$

where a is the distance between nearest neighbors. For this reason, the range of frequency happens to be the same for each branch of $\nu(\delta)$.

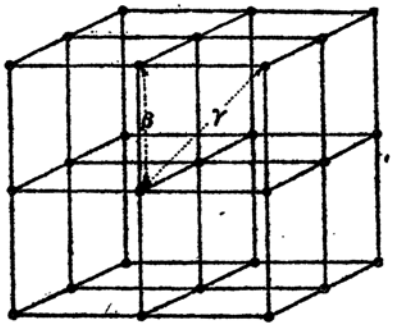

FIG. 13.—The simple cubic lattice. The Hooke's constant for nearest neighbors is β. That for next nearest is γ.

Figure 14 shows the relative number of modes as a function of ν for each of the branches of $\nu(\delta)$. The unit of frequency has been chosen arbitrarily to make $\nu_m = 1.55$. The fourth continuous curve is the relative distribution of all modes. Actually,

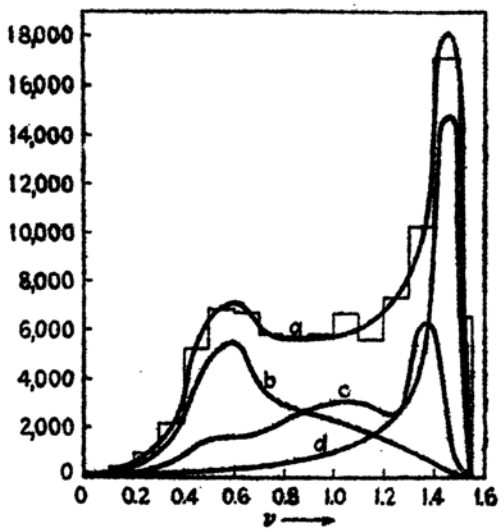

FIG. 14.—Relative scale plot of the frequency distribution of modes of vibration for the simple cubic lattice. Curves b, c, and d correspond to separate directions of polarization; curve a is the sum. (*After Blackman.*)

these curves were determined by approximate means; the stepwise curve illustrates the total distribution function as Blackman computed it.

One striking difference between the three-dimensional and one-dimensional cases is that the peak in $f(\nu)$ does not occur at ν_m in the

former. Although the gradient of $\nu(\delta)$ does vanish at the corners of the cube in δ space within which all independent modes lie, this fact does not lead to a peak in $f(\nu)$, for the volume of δ space in which the gradient of $\nu(\delta)$ vanishes is an infinitesimal of higher order. Blackman believes that the three-dimensional and one-dimensional cases usually differ in this respect.

Figure 15 is a plot of the Θ versus T curve that Blackman obtained by comparing the specific heat of the simple cubic lattice with that for the Debye continuum. The absolute units of ν_m were fixed arbitrarily in order that the high temperature Θ should be 144°. It may be seen that Θ varies considerably in the range below 40° and that it does not approach the high-temperature value at 0°K. Thus, if this were an actual crystal, Debye's law would appear to be valid experimentally above 40°K, but the value of Θ_D that would be obtained would differ

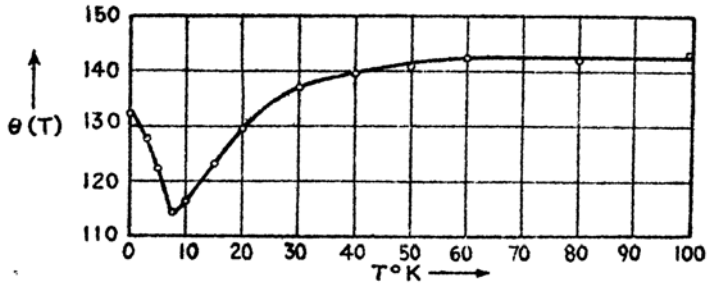

FIG. 15.—The $\Theta(T)$ curve for the simple cubic lattice. (*After Blackman.*)

from the value obtained from observations near absolute zero. In addition, the T^3 law would not be valid below 10°K, as one might expect from Debye's theory.

The substances that conform most closely to Blackman's model are the alkali halides, such as potassium chloride, in which the masses of positive and negative ions are nearly equal. The low-temperature behavior of Θ for KCl has been measured by Keesom and Clark and is described in part *b*, Sec. 20. Blackman's results do not agree with their experimental results very closely, since Θ passes through a minimum at low temperatures in his model, whereas a maximum actually is observed. It is possible that more extensive assumptions about the interaction forces between neighboring atoms would give better agreement with experiment.

24. The $C_P - C_V$ Correction.—The molar heat ordinarily measured is C_P, the heat at constant pressure. However, the theories we have been discussing in this chapter are based on the assumption that the interatomic distance is kept constant as the temperature changes; hence,

they refer to C_V, the molar heat at constant volume. C_P and C_V are connected by the thermodynamical equations[1]

$$C_P - C_V = -T\frac{(\partial V/\partial T)_P^2}{(\partial V/\partial P)_T} \tag{1}$$

where V is the molar volume and P is the pressure. If we set

$$\alpha_V = \frac{1}{V}\left(\frac{\partial V}{\partial T}\right)_P, \qquad \beta = -\frac{1}{V}\left(\frac{\partial V}{\partial P}\right)_T, \tag{2}$$

where α_V is the coefficient of volume expansion and β is the compressibility, Eq. (1) becomes

$$C_P - C_V = TV\frac{\alpha_V^2}{\beta}. \tag{3}$$

The coefficient of volume expansion is practically equal to three times the coefficient of linear expansion; hence,

$$C_P - C_V = TV\frac{9\alpha_l^2}{\beta}. \tag{4}$$

Although α_l is comparatively easy to measure at any temperature, β is usually measured only in the vicinity of room temperature. For this reason, it is necessary to obtain values of β at other temperatures by an extrapolation method of some kind.

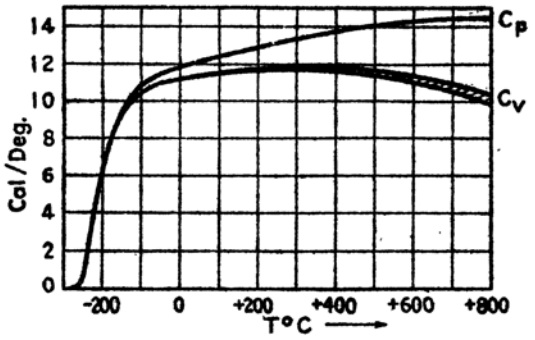

FIG. 16.—The C_P and C_V curves for sodium chloride. The ordinate is in cal/mol-deg.

Figures 16 and 17 show C_P and C_V for sodium chloride and for lead as determined by Eucken and Dannöhl.[2] They used their own values of α_l and values of β that were extrapolated linearly from room-temperature values of Slater and Bridgman. One of the most interesting features of these results is the fact that C_V seems to drop below the Dulong and

[1] See, for example, G. BIRTWISTLE, *The Principles of Thermodynamics*, pp. 71 *ff.* (Cambridge University Press, 1925).

[2] A. EUCKEN and W. DANNÖHL, *Z. Elektrochem.*, **40**, 814 (1934).

Petit value at high temperatures in the case of sodium chloride. The same workers have found a similar drop in silver. Since the rise at high temperatures in the case of lead may be explained in terms of an electronic specific heat (*cf.* Chap. IV), they conclude that the drop observed in the other cases is a common property of the contribution to C_V from lattice vibrations.

Unfortunately, this conclusion rests upon the assumption that the linearly extrapolated values of β are correct. Eucken and Dannöhl

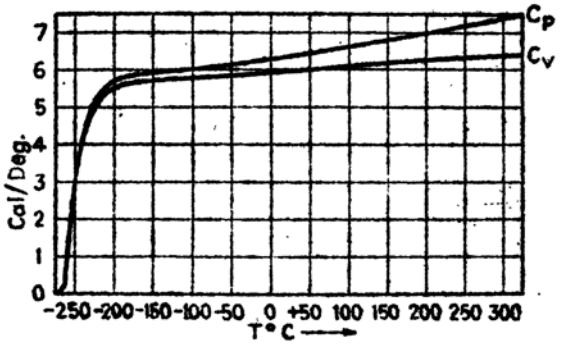

FIG. 17.—The C_P and C_V curves for lead. The ordinate is in cal/mol-deg.

believe that this assumption is justified by the fact that the quantity γ defined by the equation

$$\gamma = \frac{3\alpha_l V}{C_V \beta}, \tag{5}$$

in which C_V is their value of the specific heat at constant volume, is practically independent of temperature, the reason for this being that Grüneisen,[1] in developing an equation of state for metals on the assumption of central-force interaction of atoms, found that γ should be temperature-independent. It does not seem entirely safe to accept this result of Grüneisen's theory, however, since the interatomic forces are far from central. If Grüneisen's relation were correct, the determination of C_V would be greatly simplified, for then β could be replaced by γ and we should have

$$\frac{C_P}{C_V} = 1 + 3\gamma\alpha_l T. \tag{6}$$

Thus, γ could be determined from room-temperature measurements, and values of α_l alone would need to be measured at other temperatures.

It should be added that the deviations from Dulong and Petit's law implied by Fig. 17 are not unreasonable, for Born and Brody[2] have shown that potential interaction terms that are cubic in atomic displacements have an effect of this kind at high temperatures.

[1] E. GRÜNEISEN (see *Handbuch der Physik*, Vol. X, for a survey of this work).

[2] M. BORN and E. BRODY, *Z. Physik*, 6, 132 (1921).

CHAPTER IV

THE FREE-ELECTRON THEORY OF METALS AND SEMI-CONDUCTORS

25. Introduction.—Drude[1] first suggested that the electrical and thermal properties of metals might be correlated by assuming that metals contain free electrons in thermal equilibrium with the atoms of the solid. This hypothesis has passed through several stages of development and remains the cornerstone of the theory of metals. Drude employed the hypothesis to derive approximate expressions for electrical and thermal conductivity. In this work, he introduced the concept of a mean free path for collision of the free electrons, which has also been retained in a modified form in subsequent developments.

Lorentz[2] carried Drude's postulates to their logical conclusion in a more accurate and extensive treatment of the problem. He assumed that the electron velocities in a metal that is in field free space at constant temperature obey the Maxwell-Boltzmann distribution laws, and he determined by an ingenious method the appropriate modification of this distribution when electric fields and temperature gradients are present. Using these results, he was able to make more precise computations of the conductivities than Drude had made. In addition, he was able to treat various thermoelectric effects. As sometimes happens in such cases, Drude's results were in somewhat better agreement with experiment than Lorentz's results. These differences are of minor importance, however, when compared with two major objections to the theory, namely: (1) The manner in which Maxwell-Boltzmann statistics were employed implies that the electrons contribute a larger share of the specific heat of metals than is possible if the Einstein-Debye theory is applicable to atomic vibrations in metals. (2) It was necessary to assume that the electronic mean free path becomes infinite at the absolute zero of temperature in order to explain the vanishing of resistance at absolute zero. The theory presented no plausible reason for this fact.

The theory remained in this unsatisfactory state until after the discovery of the Pauli principle and the development of Fermi-Dirac

[1] P. DRUDE, *Ann. Physik*, **1**, 566 (1900). See H. A. LORENTZ, *The Theory of Electrons* (Teubner, Leipzig, 1909 and G. P. Stechert & Co., New York, 1923) for a discussion of this early work.

[2] H. A. LORENTZ, *Amsterdam Proc.*, 1904–1905; see also Lorentz, *op. cit.*

statistics. Then Sommerfeld[1] modified Lorentz's treatment by employing quantum statistics instead of classical statistics. This procedure, as we shall see, removed practically all the difficulties except those relating to the behavior of the mean free path at low temperatures. Houston[2] and Bloch,[3] however, were able to justify the occurrence of large mean free paths on the basis of a quantum mechanical investigation of the interaction between electrons and lattice ions.

On the whole, then, Drude's original idea has withstood the test of time. The free-electron theory merits a thorough discussion, partly for this reason and partly because it furnishes us with a clear semiquantitative picture of some of the most useful properties of metals. We may point out, however, that the free-electron picture, as we shall present it in this chapter, does not include an interpretation of the cohesive properties of metals and does not explain why some substances are metals and others are not. These topics can be understood clearly only when solids are treated on the basis of quantum mechanics.

A. METALS

26. Distribution of Electron Velocities.—Following Drude, we shall employ the following simple model of a metal. We shall assume that

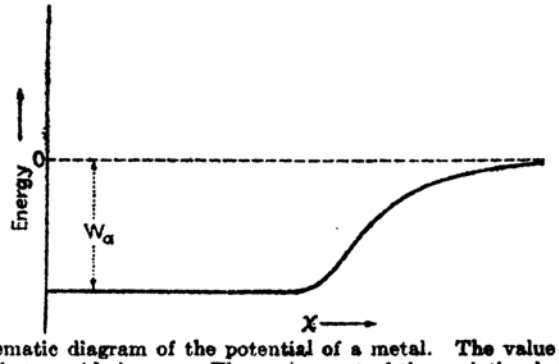

FIG. 1.—Schematic diagram of the potential of a metal. The value of the potential inside is $-W_a$, that outside is zero. The major part of the variation between $-W_a$ and zero takes place near the surface.

the electronic potential energy is constant in the interior of the metal and is equal to $-W_a$, relative to an arbitrary zero of potential at infinite distance. The precise way in which the potential energy varies in the vicinity of the surface need not concern us at present; a schematic diagram of the variation is shown in Fig. 1. The total energy of a moving

[1] A. SOMMERFELD, *Z. Physik*, **47**, 1 (1928). See also the review articles: A. Sommerfeld and H. Bethe, *Handbuch der Physik*, Vol. XXIV/2 (1934). A. Sommerfeld and N. H. Frank, *Rev. Modern Phys.*, **3**, 1 (1931).

[2] W. V. HOUSTON, *Z. Physik*, **48**, 449 (1928).

[3] F. BLOCH, *Z. Physik*, **52**, 555 (1928).

electron in the interior of the metal then is

$$E = \frac{1}{2m^*}\mathbf{p}^2 - W_a \tag{1}$$

where $\mathbf{p}$ is the momentum vector. If the electronic potential actually were constant, m^* would be the electronic mass. It happens, however, that it often is possible to approximate the electronic energy by an equation of type (1), even for fairly complex internal potential fields, by allowing m^* to take values different from the electronic mass. The reason for this fact will be made clear in later sections of the book that deal with methods of computing the energies of electrons in solids.

We shall now develop the expressions for the velocity distribution of electrons in a metal in those cases in which classical and Fermi-Dirac statistics[1] are valid. Actually, it is permissible to use only the second type of statistics, but it is interesting to compare the differences between the two forms.

Before proceeding, we shall determine the number of states having energy E, since this quantity is involved in the expressions for the distribution functions. In a rigorous treatment, the degeneracy should be determined by solving the Schrödinger equation; however, the following simple method leads to correct results. We may associate with each electron a six-dimensional phase space of which the six coordinates are the three positional coordinates x, y, z and the three components of momenta p_x, p_y, p_z of an electron. If we arbitrarily divide this phase space into cells of volume h^3, we may obtain the proper density of states by associating two states with each cell. These two states correspond to electrons that move in the same orbit with opposite orientation of electron spin. This procedure may be justified roughly by use of the

[1] For a discussion of the differences between Fermi-Dirac and classical statistics, see, for example, G. Joos, *Theoretical Physics*, Chap. 37 (G. E. Stechert & Company, New York, 1934); also Sommerfeld and Bethe, *op. cit.*, and L. Brillouin, *Die Quantenstatistik* (Julius Springer, Berlin, 1930).

In classical statistics, the number of particles in an assembly that are in the level of degeneracy g_i and energy ϵ_i is

$$n_i = g_i A e^{-\frac{\epsilon_i}{kT}}$$

where A is a constant that is to be adjusted so that the sum of n_i over all levels is equal to the total number of particles.

In Fermi-Dirac statistics, which is valid for electrons,

$$n_i = \frac{g_i}{e^{\frac{\epsilon_i - \epsilon'}{kT}} + 1}$$

in which ϵ' is the adjustable parameter analogous to A in the classical case. It may be shown that ϵ' is the free energy per particle.

phase-integral condition of classical quantum mechanics, which states that the volume of phase space associated with each level is h for each positional coordinate. Thus, the volume is h^3 for a particle in ordinary three-dimensional space. A more rigorous justification will be given in a later chapter.

In the present case, in which the electrons are allowed to wander freely throughout the volume V of the metal, we may assume that each cell in phase space extends throughout the volume V associated with the positional coordinates. Then the different cells may be completely specified by giving the domain of momentum space they cover. The average number of cells in the parallelepiped that has edges extending from p_x to $p_x + \Delta p_x$, p_y to $p_y + \Delta p_y$, and p_z to $p_z + \Delta p_z$, etc., is

$$\Delta G = V\frac{\Delta p_x \Delta p_y \Delta p_z}{h^3}, \tag{2}$$

since $V\Delta p_x \Delta p_y \Delta p_z$ is the volume of phase space occupied by the parallelepiped. In a crystal of ordinary size, for which V is greater than 10^{-12} cc, the values of Δp_x, Δp_y, Δp_z that are associated with one cell are infinitesimally small for all practical purposes. Hence, we may replace the discrete distribution of cells by a continuous one and say that the number of cells dG, in the volume $V dp_x dp_y dp_z$ of phase space, is

$$dG = V\frac{dp_x dp_y dp_z}{h^3}. \tag{3}$$

The number dG_s of states of both kinds of spin, associated with the same volume, is twice dG; that is,

$$dG_s = 2V\frac{dp_x dp_y dp_z}{h^3}. \tag{4}$$

We may now derive the expression for the number of levels lying in the range from E to $E + dE$. According to Eq. (1), the relation

$$E = \text{constant}$$

defines a sphere in momentum space of radius $\sqrt{(E + W_a)2m^*}$. The volume dP between concentric spheres the radii of which differ by dE is clearly

$$\begin{aligned} dP &= 4\pi(E + W_a)2m^* d\sqrt{(E + W_a)2m^*} \\ &= 2\pi\sqrt{(E + W_a)}(2m^*)^{\frac{3}{2}}dE. \end{aligned} \tag{5}$$

This may be simplified by setting

$$\epsilon = E + W_a, \tag{6}$$

for then

$$dP = 2\pi(2m^*)^{\frac{3}{2}}\sqrt{\epsilon}d\epsilon. \tag{7}$$

The variable ϵ evidently measures the electronic energy relative to $-W_a$. According to Eq. (4) the number of states associated with this volume of momentum space is

$$g(\epsilon)d\epsilon = \frac{4\pi V(2m^*)^{\frac{3}{2}}}{h^3}\sqrt{\epsilon}d\epsilon = C\sqrt{\epsilon}d\epsilon \tag{8}$$

where

$$C = \frac{4\pi V(2m^*)^{\frac{3}{2}}}{h^3} \tag{9}$$

and $g(\epsilon)$ is the density of states. It should be emphasized that Eq. (8) is valid only when (1) is true.

We are now prepared to discuss the distribution function for classical and Fermi-Dirac statistics.

a. Classical Distribution.—The classical, or Maxwell-Boltzmann, distribution function, is[1]

$$n_i = g_i A e^{-\frac{\epsilon_i}{kT}}, \tag{10}$$

where n_i is the number of particles in the level of energy ϵ_i which is g_i-fold degenerate, and A is the normalizing parameter. Thus, according to Eq. (8), the number dn of electrons in the energy range $d\epsilon$ is

$$dn = CAe^{-\frac{\epsilon}{kT}}\sqrt{\epsilon}d\epsilon. \tag{11}$$

If the total number of particles is N, we find, upon integrating (11) over all values of ϵ from zero to infinity, that A is related to N by the equation

$$A = \frac{2}{\sqrt{\pi}}\frac{N}{(kT)^{\frac{3}{2}}}\frac{1}{C} = \frac{h^3 n_0}{2(2\pi m^* kT)^{\frac{3}{2}}} \tag{12}$$

where $n_0 = N/V$ is the number of electrons per unit volume. Hence,

$$dn = \frac{2}{\sqrt{\pi}}\frac{N}{(kT)^{\frac{3}{2}}}e^{-\frac{\epsilon}{kT}}\sqrt{\epsilon}d\epsilon. \tag{13}$$

This result, which is independent of h and of V, shows that the size of the cell in phase space is not important so long as the cell is small enough to permit the use of a continuous distribution of levels. Incidentally, Eq. (13) is identical with the corresponding equation for the Maxwell distribution of kinetic energies of gas molecules.

[1] See footnote 1, p. 141.

The mean energy per electron $\bar{\epsilon}$ is equal to $1/N$ times the integral of ϵdn over all values of ϵ; that is,

$$\bar{\epsilon} = \frac{2}{\sqrt{\pi}} \frac{1}{(kT)^{\frac{3}{2}}} \int_0^\infty e^{-\frac{\epsilon}{kT}} \epsilon^{\frac{1}{2}} d\epsilon$$
$$= \tfrac{3}{2} kT. \tag{14}$$

Thus, according to classical statistics, the total electronic energy per mol of a monatomic metal is

$$E = \tfrac{3}{2} zRT$$

where z is the number of free electrons per atom. This result is contradicted by experiment, if we assume that z is equal to the number of valence electrons per atom, for it predicts an electronic heat per gram atom of $3zR/2$. We have seen in Chap. III, however, that practically all the specific heats of most metals can be ascribed to lattice vibrations. This contradiction is sufficient to rule out the use of classical statistics for describing the distribution of the free electrons in metals.

b. Fermi-Dirac Distribution.—The quantum statistical distribution function that should replace (10) is[1]

$$n_i = \frac{g_i}{e^{-\alpha + \frac{\epsilon_i}{kT}} + 1} = \frac{g_i}{e^{\frac{(\epsilon_i - \epsilon')}{kT}} + 1} \tag{15}$$

where, for convenience, α has been replaced by ϵ'/kT. This equation becomes

$$dn = C \frac{\sqrt{\epsilon} d\epsilon}{e^{\frac{(\epsilon - \epsilon')}{kT}} + 1} \tag{16}$$

when the continuous distribution method is used.

The parameter ϵ', which must be fixed in such a way that the integral of (16) is equal to the total number of electrons, cannot be determined so easily as the corresponding parameter in the classical case. We shall evaluate ϵ' for several limiting cases, using different methods in each one.

1. *Absolute zero.*—The form of

$$f\left(\frac{\epsilon - \epsilon'}{kT}\right) = \frac{1}{e^{\frac{\epsilon - \epsilon'}{kT}} + 1} \tag{17}$$

as a function of ϵ is illustrated in Fig. 2 for various relations between T and ϵ'. In the limiting case of absolute zero, (17) is unity when ϵ is less

[1] See footnote 1, p. 141.

than ϵ_0', the value of ϵ' at 0°K, and is zero for greater values of ϵ. Hence, Eq. (16) may be written

$$dn = \begin{cases} C\sqrt{\epsilon}d\epsilon & 0 \leq \epsilon \leq \epsilon_0' \\ 0 & \epsilon_0' < \epsilon \end{cases} \tag{18}$$

That is, at absolute zero all the energy levels below ϵ_0' are completely filled with electrons, whereas all above ϵ_0' are completely empty (*cf.* Fig. 3). Also, all cells in momentum space for which p is less than

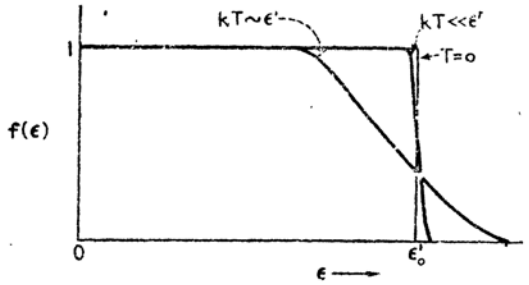

FIG. 2.—The Fermi-Dirac distribution function for several relationships between T and ϵ'.

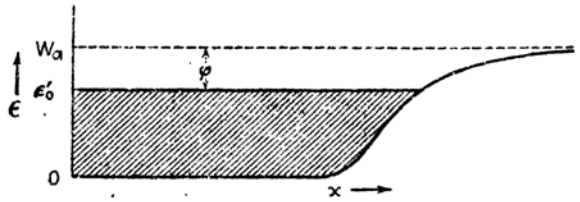

FIG. 3.—The filling of the lowest energy levels of the metal at absolute zero. All levels below $\epsilon = \epsilon_0'$ are completely occupied; all above are empty. The work function φ is the difference between the top of the filled region and the potential at $x = \infty$.

$\sqrt{2m\epsilon_0'}$ are entirely filled, whereas those that have greater momenta are completely empty.

The integral of (18) is

$$N = C\int_0^{\epsilon_0'} \sqrt{\epsilon}d\epsilon = \frac{2}{3}C\epsilon_0'^{\frac{3}{2}}. \tag{19}$$

Hence,

$$\epsilon_0' = \frac{\hbar^2}{2m^*}\left(\frac{3n_0}{8\pi}\right)^{\frac{2}{3}} \tag{20}$$

where $n_0 = N/V$, as previously. It is interesting to note that ϵ_0' depends only upon the density of particles.

Using Eq. (20) we find that the mean energy is

$$\bar{\epsilon}_0 = \frac{1}{N}\int_0^{\epsilon_0'} C\epsilon\sqrt{\epsilon}d\epsilon = \frac{3}{5}\epsilon_0' = \frac{3}{10}\frac{\hbar^2}{m^*}\left(\frac{3n_0}{8\pi}\right)^{\frac{2}{3}}. \tag{21}$$

The value of ϵ_0' for a number of metals is given in Table XXXVIII, in which we have assumed that m^* is the actual electronic mass. Since these numbers are of the magnitude of several electron volts, we see that the results of quantum statistics are appreciably different from those of classical statistics, in which $\bar{\epsilon}$ is zero at absolute zero. Thus, in classical statistics, an energy W_a would be required to remove an electron from a metal, whereas, in quantum statistics, the minimum energy required is $\phi = W_A - \epsilon_0'$ (*cf.* Fig. 3). The quantity ϕ, which is called the work function, may be evaluated experimentally by determining the light quantum of lowest energy that will eject electrons from a metal. Photoelectrically determined[1] values of ϕ are listed in Table XXXVIII. ϕ is greatly dependent on the condition of the surface through which electrons are ejected, and the values in Table XXXVIII are averages for the cleanest surfaces that have been obtained.

TABLE XXXVIII

Metal	Valence	ϵ_0', ev	ϕ, ev	$W_a = \epsilon_0' + \phi$
Li	1	4.72	2.2	6.9
Na	1	3.12	1.9	5.0
K	1	2.14	1.8	3.9
Cu	1	7.04	4.1	11.1
Ag	1	5.51	4.7	10.2
Au	1	5.54	4.8	10.3
Be	2	14.3		
Ca	2	4.26	3.2	7.5
Al	1	5.63	3.0	8.6
Al	3	11.7	3.0	14.7

We shall see below that at low temperatures $\bar{\epsilon}$ is equal to ϵ_0' plus a term that varies with temperature as $(kT)^2/\epsilon_0'$. Since this term ordinarily is very small compared with unity, the complete filling of the lowest energy levels that are shown in Fig. 3 is not appreciably altered at ordinary temperatures.

2. *The case in which kT is small in comparison with ϵ_0'.*—This case is a very useful one since ϵ_0' is much larger than kT below the melting point of most metals.

The equation for determining ϵ' in this case, as well as the general one, is

$$N = C\int_0^\infty \frac{\sqrt{\epsilon}\,d\epsilon}{e^{\frac{\epsilon-\epsilon'}{kT}}+1}. \tag{22}$$

[1] See, for example, the compilations by J. A. Becker, *Rev. Modern Phys.*, **7**, 95 (1935), and A. L. Hughes and L. A. DuBridge, *Photoelectric Phenomena* (McGraw-Hill Book Company, Inc., New York, 1932).

Sommerfeld and Bethe[1] have developed a convenient method for evaluating integrals of this type. Let us write Eq. (22) in the form

$$N = C\int_0^\infty f\sqrt{\epsilon}\,d\epsilon, \tag{23}$$

using the notation of Eq. (17). Integrating this expression by parts, we obtain

$$N = -C\int_0^\infty \frac{df}{d\epsilon}\frac{2}{3}\epsilon^{\frac{3}{2}}d\epsilon. \tag{24}$$

It will be found that a number of important integrals may be expressed in an analogous form, namely,

$$a = C\int_0^\infty \frac{df}{d\epsilon}\alpha(\epsilon)d\epsilon \tag{25}$$

where $\alpha(\epsilon)$ is a continuous function of ϵ.

Following Sommerfeld and Bethe, let us transform the integration variable in (25) from ϵ to η defined by

$$\eta = \frac{\epsilon - \epsilon'}{kT}.$$

This is equivalent to choosing ϵ' to be the origin of the energy scale and expressing the energy in units of kT. In addition, let us write

$$\alpha(\epsilon) = \beta(\eta).$$

Then, Eq. (25) becomes

$$a = C\int_{-\frac{\epsilon'}{kT}}^\infty \beta(\eta)\frac{df(\eta)}{d\eta}d\eta. \tag{26}$$

For small values of kT, ϵ'/kT is so large that the lower limit may be replaced by $-\infty$, whence

$$a \cong C\int_{-\infty}^\infty \beta(\eta)\frac{df(\eta)}{d\eta}d\eta. \tag{27}$$

Let us assume that β may be expanded into a Taylor series. Then,

$$\beta(\eta) = \beta(0) + \left[\frac{\partial\beta}{\partial\eta}\right]_0\eta + \frac{1}{2}\left[\frac{\partial^2\beta}{\partial\eta^2}\right]_0\eta^2 + \cdots,$$

[1] SOMMERFELD and BETHE, *op. cit.*

and Eq. (27) is

$$a = C\int_{-\infty}^{\infty} \frac{\partial f}{\partial \eta}\left(\beta(0) + \left[\frac{\partial \beta}{\partial \eta}\right]_0 \eta + \frac{1}{2}\left[\frac{\partial^2 \beta}{\partial \eta^2}\right]_0 \eta^2 + \cdots\right) d\eta.$$

The first integral is

$$\beta(0)C\int_{-\infty}^{\infty} \frac{\partial f}{\partial \eta} d\eta = \beta(0)C[f(\eta)]_{-\infty}^{\infty} = -\beta(0)C.$$

The second integral vanishes because $\partial f/\partial \eta$ is an even function of η.

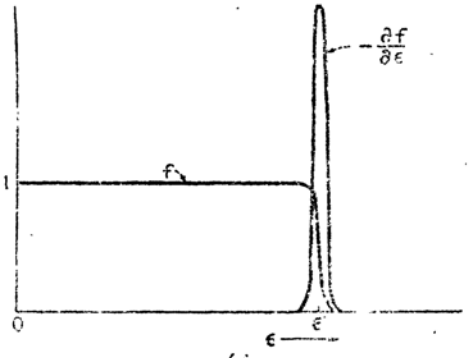

FIG. 4.—The functions f and $-\frac{\partial f}{\partial \epsilon}$ as functions of ϵ (schematic).

The third term, which is the only other one that need be considered in the low-temperature approximation, is

$$-\frac{1}{2}\left[\frac{\partial^2 \beta}{\partial \eta^2}\right]_0 C\int_{-\infty}^{\infty} \frac{\eta^2 e^{-\eta}}{(e^{-\eta}+1)^2} d\eta \tag{28}$$

This may be evaluated by expanding the integrand in powers of $e^{-\eta}$. The integral then is

$$2\int_0^{\infty} \eta^2(e^{-\eta} - 2e^{-2\eta} + 3e^{-3\eta} - \cdots) d\eta = 4\left(1 - \frac{1}{2^2} + \frac{1}{3^2} - \frac{1}{4^2} + \cdots\right) \cdot$$
$$= \frac{\pi^2}{3}.$$

Hence,

$$a \cong -C\left(\beta(0) + \frac{\pi^2}{6}\left[\frac{\partial^2 \beta}{\partial \eta^2}\right]_0\right).$$

Transforming back to the variable ϵ, we have

$$a \cong -C\left(\alpha(\epsilon') + \frac{\pi^2}{6}(kT)^2\left[\frac{\partial^2 \alpha}{\partial \epsilon^2}\right]_{\epsilon=\epsilon'}\right). \tag{29}$$

It is easy to see that the infinite series of which the terms in (29) are the first few members converges very rapidly when kT is much less than ϵ'. The form of the function $df/d\epsilon$ for this case is illustrated in Fig. 4. It is obvious that this function vanishes everywhere except in a region of width kT at $\epsilon = \epsilon'$, where it has a steep maximum. Since

$$\int_{-\infty}^{\infty} \frac{df}{d\epsilon} d\epsilon = -1,$$

the main term in the expansion of a is simply $-C\alpha(\epsilon')$, the value of the integrand at ϵ'. The additional terms are corrections for the finite width of the function $\partial f/\partial\epsilon$, and are small as long as the width is small.

Let us return to Eq. (24). Using Eq. (29), we find that

$$N = C\left(\frac{2}{3}\epsilon'^{\frac{3}{2}} + \frac{\pi^2}{12}\frac{(kT)^2}{\epsilon'^{\frac{1}{2}}}\right), \tag{30}$$

which reduces to Eq. (19) when T becomes zero. We may solve this for ϵ' when T is not zero by replacing ϵ' in the denominator of the second term by ϵ_0'. It is found in this way that

$$\begin{aligned}\epsilon' &\cong \epsilon_0'\left[1 - \frac{\pi^2}{8}\left(\frac{kT}{\epsilon_0'}\right)^2\right]^{\frac{2}{3}} \\ &\cong \epsilon_0'\left[1 - \frac{\pi^2}{12}\left(\frac{kT}{\epsilon_0'}\right)^2\right].\end{aligned} \tag{31}$$

The mean energy per electron is

$$\bar{\epsilon} = \frac{C}{N}\int_0^{\infty} f\epsilon^{\frac{3}{2}} d\epsilon = -\frac{C}{N}\int_0^{\infty} \frac{df}{d\epsilon}\frac{2}{5}\epsilon^{\frac{5}{2}} d\epsilon. \tag{32}$$

When Eqs. (19) and (29) are used, this reduces to

$$\begin{aligned}\bar{\epsilon} &= \frac{C}{N}\left(\frac{2}{5}\epsilon'^{\frac{5}{2}} + \frac{\pi^2}{4}\epsilon'^{\frac{1}{2}}(kT)^2\right) \\ &= \frac{3}{2}\frac{1}{\epsilon_0'^{\frac{3}{2}}}\left(\frac{2}{5}\epsilon'^{\frac{5}{2}} + \frac{\pi^2}{4}\epsilon'^{\frac{1}{2}}(kT)^2\right).\end{aligned} \tag{33}$$

Upon simplifying this equation by means of Eq. (31), we find

$$\begin{aligned}\bar{\epsilon} &= \frac{3}{5}\epsilon_0'\left[1 + \frac{5}{12}\pi^2\left(\frac{kT}{\epsilon_0'}\right)^2\right] \\ &= \bar{\epsilon}_0\left[1 + \frac{5}{12}\pi^2\left(\frac{kT}{\epsilon_0'}\right)^2\right]\end{aligned} \tag{34}$$

where $\bar{\epsilon}_0$ is the value of $\bar{\epsilon}$ for absolute zero [*cf.* Eq. (21)].

The derivative of Eq. (34) with respect to T is the electronic heat γ_V; that is,

$$\gamma_V = k\frac{\pi^2}{2}\frac{kT}{\epsilon_0}.$$

A linear term, which ordinarily is much smaller than the classical value of $3k/2$, has been observed in a number of metals and will be discussed in Secs. 27 and 28. It is easy to see from the method we have used to arrive at Eq. (34) that only the part of the electronic distribution near $\epsilon = \epsilon'$ contributes to the temperature-dependent part of $\bar{\epsilon}$. The electrons of lower energy are hemmed in by filled cells to such an extent that ordinarily they are not excited. Thus, only those electrons that are in the energy range of width kT near the top of the occupied levels are free in the classical sense. In fact, one may obtain a specific heat of the same order of magnitude as (35) by assuming that a fraction kT/ϵ' of the electrons are free and have the classical electronic heat $3k/2$.

Equation (29) may be used to derive expressions for the mean value of various quantities in cases more general than that in which $g(\epsilon)$ has the form (8). Suppose that $g(\epsilon)$ is an arbitrary function of ϵ. Then, the integrals for N and $\bar{\epsilon}$ are

$$N = \int_0^\infty g(\epsilon)f\,d\epsilon = -\int_0^\infty \left[\int_0^\epsilon g(x)\,dx\right]\frac{df}{d\epsilon}d\epsilon \tag{36}$$

$$N\bar{\epsilon} = \int_0^\infty \epsilon g(\epsilon)f\,d\epsilon = -\int_0^\infty \left[\int_0^\epsilon xg(x)\,dx\right]\frac{df}{d\epsilon}d\epsilon. \tag{37}$$

Using Eq. (29), we find that these expressions may be reduced to

$$N = \int_0^{\epsilon'} g(\epsilon)\,d\epsilon + \frac{\pi^2}{6}(kT)^2 g'(\epsilon'), \tag{36a}$$

$$N\bar{\epsilon} = \int_0^{\epsilon'} \epsilon g(\epsilon)\,d\epsilon + \frac{\pi^2(kT)^2}{6}[\epsilon' g'(\epsilon') + g(\epsilon')]. \tag{37a}$$

The total electronic heat may be derived from (37a) by differentiating this equation with respect to T and is

$$N\gamma_V = \left\{\epsilon' g(\epsilon') + \frac{\pi^2 k^2 T^2}{6}[2g'(\epsilon') + \epsilon' g''(\epsilon')]\right\}\frac{d\epsilon'}{dT} + \frac{\pi^2 k^2 T}{3}[\epsilon' g'(\epsilon') + g(\epsilon')]. \tag{38}$$

An expression for the derivative of ϵ' with respect to temperature may be obtained by taking the temperature derivative of Eq. (36a). The result is

$$\frac{d\epsilon'}{dT} \cong -\frac{\pi^2}{3}k^2 T\frac{g'(\epsilon')}{g(\epsilon')}.$$

Substituting this in Eq. (38), we find

$$\gamma_V \cong \frac{\pi^2 k^2 T}{3} \frac{g(\epsilon')}{N}. \tag{39}$$

Equation (35) is the special case of this equation in which g has the value $C\sqrt{\epsilon}$.

3. *The case in which ϵ' is large and negative.*—We shall see that this case is of interest at very high temperatures. When ϵ' is negative, the quantity

$$e^{\frac{\epsilon'}{kT}} \tag{40}$$

is less than unity. Hence, we may expand the expressions for N and $\bar{\epsilon}$ in terms of it. The results, to terms in the first power of (40), are

$$N \cong Ce^{\frac{\epsilon'}{kT}} \int_0^\infty e^{-\frac{\epsilon}{kT}} \sqrt{\epsilon}\, d\epsilon, \tag{41a}$$

$$\bar{\epsilon} \cong \frac{Ce^{\frac{\epsilon'}{kT}}}{N} \int_0^\infty e^{-\frac{\epsilon}{kT}} \epsilon^{\frac{3}{2}} d\epsilon. \tag{41b}$$

These equations become identical with those for the case of classical statistics, which was discussed in part *a*, when we set

$$A = e^{\frac{\epsilon'}{kT}}.$$

This result shows that classical statistics are valid under the conditions in which (40) is small in comparison with unity. Substituting A from Eq. (12), we find that

$$\frac{\epsilon'}{kT} = \log \frac{h^3 n_0}{2(2\pi mkT)^{\frac{3}{2}}}.$$

The condition that must be satisfied if this is to be negative and large is

$$kT \gg \left(\frac{16}{9\pi}\right)^{\frac{2}{3}} \frac{h^2}{2m} \left(\frac{3n_0}{8\pi}\right)^{\frac{2}{3}} = \left(\frac{16}{9\pi}\right)^{\frac{2}{3}} \epsilon_0'.$$

For most metals, this condition is satisfied only at temperatures far above the melting point.

4. *Intermediate case.*—The evaluation of integrals in the intermediate case, in which kT is comparable with ϵ_0', involves a relatively large amount of computation and will not be discussed here. Some of the more important results have been derived by Mott[1] and by Stoner.[2]

[1] N. F. MOTT, *Proc. Roy. Soc.*, **152**, 42 (1935).
[2] E. C. STONER, *Phil. Mag.*, **21**, 145 (1936).

For example, the behavior of the electronic heat, as a function of kT/ϵ_0', is shown in Fig. 5.

27. The Specific Heats of Nontransition Metals.—The expression that was derived in the preceding section for the heat of free electrons in a metal, namely,

$$\gamma_V = \frac{k}{2}\pi^2\frac{kT}{\epsilon_0'} \tag{1}$$

per electron, is small compared with the contribution from lattice vibrations at ordinary temperatures. At low temperatures, however, it should become comparable with the Debye value of

$$k\frac{12\pi^4}{5}\left(\frac{T}{\Theta_D}\right)^3 \tag{2}$$

per atom, since the quantity (2) decreases more rapidly with decreasing temperature than (1). The ratio of (1) to (2), namely,

$$\frac{5}{24\pi^2}\frac{kT}{\epsilon_0'}\left(\frac{\Theta_D}{T}\right)^3,$$

approaches unity in the neighborhood of 1°K when Θ_D is of the order of 100° and ϵ_0' is of the order of 1 electron volt. Thus, according to this result, an appreciable part of the specific heat of simple metals in the neighborhood of 1°K should have electronic origin.

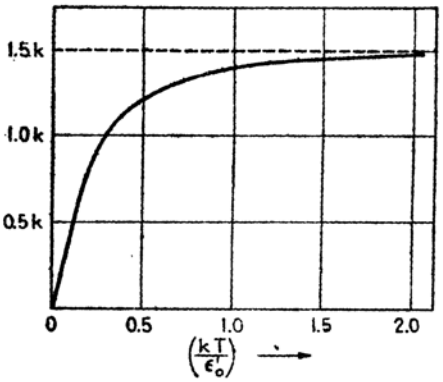

FIG. 5.—The electronic heat as a function of temperature during the transition from the degenerate to the nondegenerate state. (*After Mott.*)

Keesom[1] and Kok have observed an electronic specific heat of this type in the simple metals silver, zinc, copper, and aluminum. The most accurate measurements have been made on copper and aluminum, for which the molar heats are

$$\text{Cu: } C_V = 0.888 \cdot 10^{-4}RT + 3Rf_D\left(\frac{335}{T}\right), \tag{3}$$

$$\text{Al: } C_V = 1.742 \cdot 10^{-4}RT + 3Rf_D\left(\frac{419}{T}\right), \tag{4}$$

where $f_D(\Theta/T)$ is the Debye function. Equation (4) is valid only above 1.13°K, for the metal changes to the superconducting phase at this temperature. It is possible to determine an "experimental" ϵ_0' from α,

[1] Ag, Zn: W. H. KEESOM and J. A. KOK, *Physica*, **1**, 770 (1934). Cu: J. A. KOK and W. H. KEESOM, *Physica*, **3**, 1035 (1936). Al: J. A. KOK and W. H. KEESOM, *Physica*, **4**, 835 (1937).

the coefficient of T in the linear term in (3) and (4). In fact, we have from Eq. (1)

$$\epsilon_0'(\text{exp.}) = \frac{\pi^2 kR}{2\alpha} z \tag{5}$$

where z is the number of valence electrons per atom. Values of this quantity are given in Table XXXIX. In the case of copper, it is assumed that there is 1 free electron per atom, whereas with aluminum values are given for the cases in which it is assumed that there are 1 and 3 electrons per atom. Evidence that will be discussed in Chap. XIII indicates that there actually are 3 free electrons per atom. Dividing the computed ϵ_0' of Table XXXIX by the experimental ones, we obtain the ratio of the effective electronic mass to the real mass. These ratios also appear in Table XXXIX.

Equations so precise as (3) and (4) are not given for silver and zinc. The experimenters estimate, however, that m^*/m is of the order of unity in both cases.

TABLE XXXIX

	ϵ_0' (exp.), ev	z	m^*/m
Cu	4.78	1	1.47
Al	2.44	1	2.30
	7.30	3	1.61

28. The Electronic Specific Heats of Transition Metals at Low Temperatures.—Experimental investigations that we shall discuss presently show that transition metals often have a much larger electronic specific heat than do simple metals. In order to treat this topic at the present point, it is necessary to accept some simple consequences of the band theory of solids; these will be justified in later chapters.

According to the band theory,[1] the ten states of an atomic d shell contribute a quasi-continuous band of $10N$ electronic levels to the metal, where N is the total number of atoms. Of these levels, $5N$ correspond to one orientation of spin, and the other $5N$ correspond to the opposite orientation. Figure 6 shows the position and width of the d electron band relative to the levels of the ordinary valence electrons. At present, we shall not discuss the quantum mechanical principles that determine the width ΔE of the d band, and the relative position of the valence and d levels. The density of valence-electron levels is so much less than the density of d levels that the point below which there are $11N$ levels of both valence and d type is above the top of the d band. In other words, the

[1] This d-electron-band model was first proposed by N. F. Mott, *Proc. Phys. Soc.*, **47**, 571 (1935), and has been extended by J. C. Slater, *Phys. Rev.*, **49**, 537 (1936).

ordinary valence levels are filled above the top of the d band if N electrons are added to the valence band (*cf.* Fig. 6). Thus, the d band is completely filled in metals such as copper, silver, and gold, which have $11N$ electrons outside the rare gas shells. Conversely, the d band is not completely filled in the transition metals that immediately precede copper, silver, or gold in the periodic chart.

We shall let $g_v(\epsilon)$ and $g_d(\epsilon)$, respectively, designate the density of levels in the valence and d bands, as functions of energy. The origin of ϵ will be taken to be at the bottom of the valence band. For present

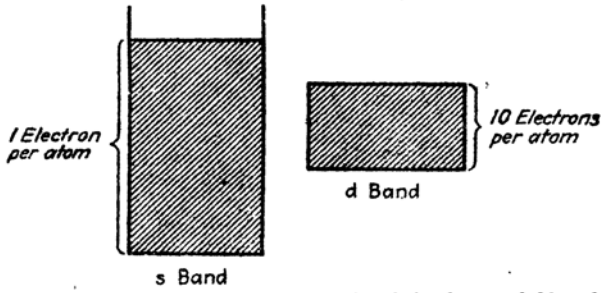

FIG. 6.—The relative widths and the density of levels in the s and d bands. The shaded area of the s band represents the width of energy levels occupied by one electron per atom. There is room for ten electrons per atom in the d band. If there are eleven valence electrons per atom, as in copper, silver and gold, both bands are filled as shown in this figure. For the case in which there are ten electrons per atom see Fig. 7.

purposes, we shall assume that $g_v(\epsilon)$ has the form of $g(\epsilon)$ in Eq. (8) of Sec. 26, namely,

$$g_v(\epsilon) = C\sqrt{\epsilon} \tag{1}$$

where the mass appearing in C does not differ from the true electronic mass by a factor larger than 2 or 3. As we shall see, there is good evidence that near the bottom of the d band $g_d(\epsilon)$ has the same form as the expression (1) for free electrons, namely,

$$g_d(\epsilon) = C_b\sqrt{\epsilon - E_d} \qquad \epsilon \geq E_d. \tag{2}$$

where

$$C_b = \frac{4\pi V(2m_b)^{\frac{3}{2}}}{h^3} \tag{3}$$

in which m_b is the effective mass at the bottom of the band. We shall assume that near the top of the band

$$g_d(\epsilon) = C_t\sqrt{(E_d + \Delta E - \epsilon)} \qquad \epsilon \lesssim E_d + \Delta E \tag{4}$$

where

$$C_t = \frac{4\pi V(2m_t)^{\frac{3}{2}}}{h^3} \tag{5}$$

in which m_t is the effective mass near the top of the band.

It may be expected that m_b and m_t will turn out to be much larger than the electronic mass, since the density of levels in the d band is much larger than the density of levels in the valence-electron band.

According to Eq. (39), which is valid when kT is much less than ϵ_0', the electronic heat is

$$\gamma_V = \frac{\pi^2 k^2 T}{3} \frac{g(\epsilon_0')}{N} \tag{6}$$

per electron, where $g(\epsilon_0')$ is the total density of levels at the top of the filled region. If there are more than $10N$ electrons,

$$g(\epsilon_0') = g_v(\epsilon_0'), \tag{7}$$

since then the d band is completely filled and the valence band is filled beyond the top of the d band. Hence, we may expect that in this case the electronic heat is of the same magnitude as that of other simple metals. This was shown to be true in the previous section, for it was found there that the electronic heats of copper, silver, and zinc are the same order of magnitude as that of aluminum. On the other hand, $g(\epsilon_0')$ is

$$g(\epsilon_0') = g_v(\epsilon_0') + g_d(\epsilon_0') \tag{8}$$

when there are fewer than $11N$ electrons, for then the d band is only partly filled. We may expect (8) to be much greater than (7), since we have seen that the density of d levels is much greater than the density of valence levels. Consequently we may expect a larger electronic specific heat for transition metals.

Before presenting the experimental facts for nickel, we shall find it convenient to discuss the implications of ferromagnetism from the energy-band picture. Let us divide the d band into two bands, each having density $g_d(\epsilon)/2$, corresponding to two opposite orientations of electron spin (*cf.* Fig. 7). In a paramagnetic transition metal, these two bands are filled to exactly the same level when no external magnetic field is present (*cf.* Fig. 7*a*). Thus, the intrinsic magnetic moment is zero in this case. We may interpret many of the properties of ferromagnetic substances by assuming that at the absolute zero of temperature one of the two d bands is completely filled, whereas the other is filled to the same height as the valence band (*cf.* Fig. 7*b*). The metal then possesses an intrinsic magnetic moment per unit volume, since there is an excess of electrons with spin oriented in the direction associated with the filled band. If β is the magnetic moment per electron, the magnetic moment M per unit volume is seen to be

$$\mathsf{M} = \beta \Delta n_d \tag{9}$$

where Δn_d is the difference between the numbers of electrons per unit volume in the d band having each orientation. This occurrence of the freezing of electrons in one-half of the d band is accompanied by a decrease in the total energy of the metal. We shall discuss this in a later chapter. Since all ferromagnetic metals become paramagnetic at sufficiently high temperatures, after a continuous decrease of the intrinsic magnetization, we may conclude that the frozen electronic structure gradually melts as the temperature is raised.

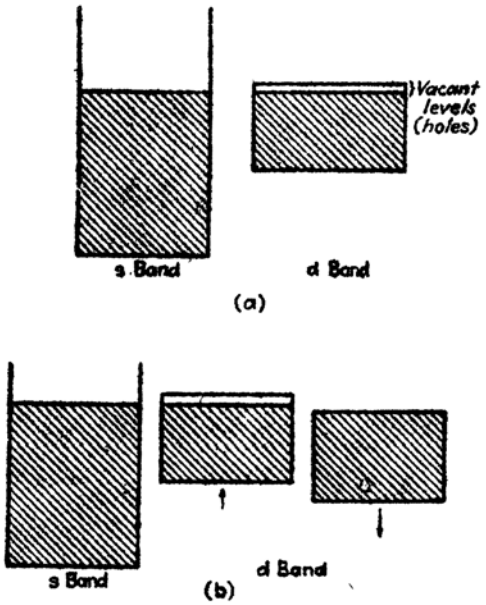

FIG. 7.—Filling of levels in transition metals, which have ten or fewer valence electrons per atom. Case *a*, paramagnetic metal in which levels of both spins in the *d* band are filled to the same height. Case *b*, ferromagnetic metal in which the levels in the *d* band of one spin are preferentially filled. The levels in the other half of the *d* band and in the *s* band are filled to the same height.

The number Δn_d may be determined for ferromagnetic metals from Eq. (9) by using experimental values of M and assuming that β is the Bohr magneton

$$\beta = \frac{eh}{4\pi mc}. \tag{10}$$

This number, which is given for nickel, cobalt, and iron in Table VII, Chap. I, is also equal to the number of electrons per unit volume missing from the unfilled half of the d band. For this reason, it is often called the number of holes in the d band. Thus, there is 0.6 hole per atom in the d band of nickel, etc. Since the nickel d band is nearly filled, we shall

assume that the density $g_{d/2}$ of vacant levels in the partly filled half band is half of (4), or that

$$g_{d/2}(\epsilon) = \frac{C_t\sqrt{E_d + \Delta E - \epsilon}}{2}.$$

The parameter ϵ' may be determined from this by means of the following equation which states that the number of holes is $0.6N$, where N is the total number of atoms:

$$0.6 = \frac{C_t}{2N}\int_{\epsilon'}^{E_d+\Delta E}\sqrt{E_d + \Delta E - \epsilon}\,d\epsilon. \tag{11}$$

Integrating this, we find that

$$\delta\epsilon' \equiv E_d + \Delta E - \epsilon_d' = \left(\frac{1.8N}{C_t}\right)^{\frac{2}{3}} \tag{12}$$

where $\delta\epsilon'$ is the width of the unfilled region. Hence, the density of levels at the point ϵ' in the partly filled half of the d band is

$$g_{d/2}(\epsilon') = \frac{C_t^{\frac{1}{2}}}{2}(1.8N)^{\frac{1}{2}}. \tag{13}$$

According to Eq. (6), the contribution from the d electrons to the electronic heat is

$$\gamma_V = \frac{\pi^2k^2T}{3}\frac{C_t^{\frac{1}{2}}}{2N^{\frac{1}{2}}}(1.8)^{\frac{1}{2}} \tag{14}$$

per electron. If we use Eq. (12), we may replace C_t by $\delta\epsilon'$. The result is

$$\gamma_V = 0.3\frac{\pi^2k^2T}{\delta\epsilon'}, \tag{15}$$

or the molar heat is

$$C_V = 0.3R\pi^2\frac{kT}{\delta\epsilon'}. \tag{16}$$

Keesom and Clark[1] have found that the molar heat of nickel at very low temperatures can be expressed by the equation

$$C_V = 8.72 \cdot 10^{-4}RT + 3Rf_D\left(\frac{413}{T}\right). \tag{17}$$

The electronic term is about ten times larger than that for simple metals

[1] W. H. KEESOM and C. W. CLARK, *Physica*, 2, 513 (1935).

such as copper or aluminum. Equating (16) to the first term in (17), we find

$$\frac{\delta\epsilon'}{k} \cong 3400°\text{K},$$

which is about 0.29 ev. This should be corrected slightly to include the small contribution of the valence electrons to the specific heat. The result, however, would not be appreciably different.

The ratio of m_t to the true electronic mass may be determined easily from the preceding equations. The result is

$$\frac{m_t}{m} \sim 28,$$

which shows that the d electrons behave as though they were relatively heavy particles.

Keesom and Kurrelmeyer[1] have measured the specific heat of α iron at low temperatures and find

$$C_v = 0.60 \cdot 10^{-3}RT + 2.36 \cdot 10^{-6}RT^3.$$

In this case, the number of holes in the d band is of the order of 2.2 per atom, if we assume that we may interpret the data of Table VII of Chap. I (page 23) in the same way as for nickel. The corresponding values of $\delta\epsilon'$ and m_t/m are listed in Table XL. The fact that m_t/m is smaller than for nickel is partly connected with the fact that Eq. (4) is not valid for the holes in the d band of iron, as will be seen in Chap. XIII.

It has become conventional to associate the specific heat of the d electrons in transition metals with the holes in the d bands. This procedure is convenient because the use of quantities such as Δn in (9) and $\delta\epsilon'$ in (12) and (16), which are, respectively, the number of holes in the unfilled region and the width of this region, reduces the expression for the specific heat to its simplest form. We shall see later that this convention has many other advantages.

Large electronic specific heats have been observed in palladium and platinum.[2] However, it is not possible to determine $\delta\epsilon'$ for these metals from the observed data, for they are not ferromagnetic. Since these metals occupy positions in the periodic table similar to nickel, we shall assume that they also have 0.6 hole per atom. These holes are distributed equally among levels of both kinds of spin so that the equation relating C_t and $\delta\epsilon'$ is now different from (12). Instead of (12), we find

$$\delta\epsilon' = \left(\frac{0.9N}{C_t}\right)^{\frac{1}{2}};$$

[1] W. H. KEESOM and B. KURRELMEYER, *Physica*, **6**, 364 (1939).

[2] Pd: G. L. PICKARD, *Nature*, **138**, 123 (1936).
Pt: J. A. KOK and W. H. KEESOM, *Physica*, **3**, 1035 (1936).

however, the equation connection C_V and $\delta\epsilon'$ turns out to be the same as (16). Using the observed electronic heats per mol of metal, namely,

$$\text{Pd: } C_V = 1.6 \cdot 10^{-3} RT,$$
$$\text{Pt: } C_V = 0.804 \cdot 10^{-3} RT,$$

we obtain the values of $\delta\epsilon'$ and m_i/m given in Table XL.

TABLE XL.—THE EFFECTIVE MASSES OF THE HOLES IN THE *d* BANDS OF SEVERAL TRANSITION METALS

(Derived from the observed electronic heats on the assumption that the holes are perfectly free)

Metal	$\delta\epsilon'$, ev	m_i/m
Ni	0.29	28
α Fe	1.58	12
Pd	0.16	43
Pt	0.32	22

29. The Pauli Theory of the Paramagnetism of Simple Metals.—It may be seen from Fig. 20, Chap. I, that the metals that follow the rare gases in the periodic chart are weakly paramagnetic. Since it may be shown that the inner closed-shell electrons of these substances give a diamagnetic contribution to the total susceptibility, we may conclude that the paramagnetic susceptibility is associated with the valence electrons. Pauli[1] proposed the following simple semiquantitative interpretation of this paramagnetic term.

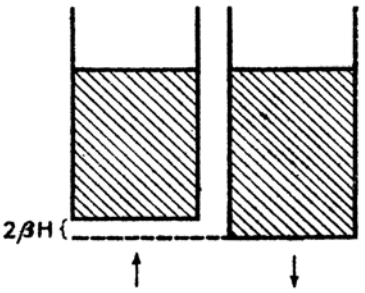

FIG. 8.—The relative displacement of the levels of different spin. The levels with magnetic moment parallel to the field are lowered by βH, those of opposite spin are raised by the same amount. The energy difference 2βH is exaggerated.

Let us divide the quasi-continuous band of levels shown in Fig. 3 into two bands, one for electrons of a given spin and one for electrons of the opposite spin, just as we did for the *d* band in the preceding section. The density of levels in each of these two bands obviously is just half the density in the band of Fig. 3. In the absence of a magnetic field, each band is filled to exactly the same value of ϵ. If the metal is placed in a homogeneous magnetic field of intensity H, the band of levels associated with electrons having spin parallel to the field is lowered by an amount βH, and the other band is raised by the same amount. Here, β is the magnetic moment of the electron which, according to the theory of

[1] W. PAULI, *Z. Physik*, **41**, 81 (1927).

electron spin, is given by Eq. (10) of the preceding section. Figure 8 shows schematically the behavior of the levels. Evidently, the equilibrium distribution of electrons is that in which both types of level are filled to exactly the same point on an energy scale, for otherwise we could gain energy by removing electrons from the highest filled band to the lowest. Hence, more electrons have their magnetic moment parallel to the field than antiparallel to it. The number Δn of electrons that leave the antiparallel band and enter the other is equal to the number of electrons in the energy range of width $\beta\mathsf{H}$ at the top of one of the bands, that is,

$$\Delta n = \beta \mathsf{H} g_s(\epsilon') \tag{1}$$

where $g_s(\epsilon')$ is the density of levels of given spin at the top of the filled region. The relation (1) is valid only for fields that satisfy the condition

$$\beta \mathsf{H} \ll \epsilon'.$$

All ordinary fields are included in this condition, since $\beta\mathsf{H}$ is of the order of 10^{-3} ev for the strongest attainable fields and ϵ' is of the order of 1 volt. The difference between the number of electrons in the two bands, namely $2\Delta n$, is

$$2\Delta n = 2\beta \mathsf{H} g_s(\epsilon'), \tag{2}$$

whence the magnetic moment per unit volume M is

$$\mathsf{M} = \frac{2\beta^2 \mathsf{H} g_s(\epsilon')}{V}$$

and the susceptibility is

$$\chi = \frac{2\beta^2 g_s(\epsilon')}{V}. \tag{3}$$

We shall assume that the electrons are perfectly free and shall substitute $g_s(\epsilon')$ from Eq. (8) of Sec. 28; that is,

$$g_s(\epsilon') = \frac{g(\epsilon')}{2} = \frac{C}{2}\sqrt{\epsilon'} \cong \frac{3}{4} N \frac{1}{\epsilon_0'}.$$

Then,

$$\chi = \frac{3}{2} n_0 \beta^2 \frac{1}{\epsilon_0'}$$

The relationship between χ, ϵ_0', the number of valence electrons per atom, Z, and the atomic volume a, is

$$\chi = \frac{81.0Z}{a\epsilon_0'} 10^{-6} \tag{4}$$

where χ is now expressed in cgs units, a is expressed in cubic angstroms, and ϵ_0' is expressed in electron volts.

Table XLI contains values of χ, computed from Eq. (4), for the alkali metals and several alkaline earth metals. It should be observed that the computed susceptibilities are often more positive than the observed ones, a fact indicating that a diamagnetic correction is frequently needed. One of the principal sources of this correction is the closed-shell ion cores

TABLE XLI.—A COMPARISON OF THE OBSERVED SUSCEPTIBILITIES OF SEVERAL METALS WITH THOSE COMPUTED ON THE BASIS OF THE FREE-ELECTRON THEORY
(In cgs units)

Metal	$\chi \cdot 10^6$	
	Observed	Calculated
Li	2.0	0.80
Na	0.63	0.65
K	0.58	0.53
Be	−1.85	1.38
Mg	0.87	0.98
Ca	1.70	0.89

of the atoms. In the cases of copper, silver, and gold, the diamagnetic contribution from the newly filled d shells is large enough to cancel the paramagnetic term and to make the metal diamagnetic

The transition metals such as platinum and palladium are strongly paramagnetic. Since the free-electron theory may be used to explain the electronic heat (*cf.* Sec. 28), we naturally should attempt to apply to these metals the analogue of Eq. (4), namely,

$$\chi = \frac{81.0Z}{a\delta\epsilon_0'}10^{-6}$$

where Z is the number of holes per atom. It was first shown by Mott and Jones,[1] and may be verified by simple calculations, that the values of $\delta\epsilon_0'$ that are required to explain the observed values of χ are four or five times smaller than those derived from the experimental values of the electronic heat. This discrepancy shows that the free-electron model is much too simple for the d-shell electrons.

30. Thermionic and Schottky Emission.—Up to this point, we have had no cause to consider the way in which the electronic potential of a metal varies near the surface. The form of this curve is important, however, when we consider any process in which electrons pass through

[1] N. F. MOTT and H. JONES, *Theory of the Properties of Metals and Alloys*, pp. 194 *ff.* (Oxford University Press, 1936).

the surface. We shall discuss two phenomena of this type in the present section, namely, thermionic emission and Schottky emission. Thermionic emission[1] is the phenomenon in which an electronic current evaporates from a heated metal in the absence of an external electric field, whereas Schottky emission refers to the evaporation that occurs when the metal is at a negative potential. Both these emission phenomena are strongly temperature-dependent. An additional temperature-independent emission occurs when the potential of the metal becomes suffi-

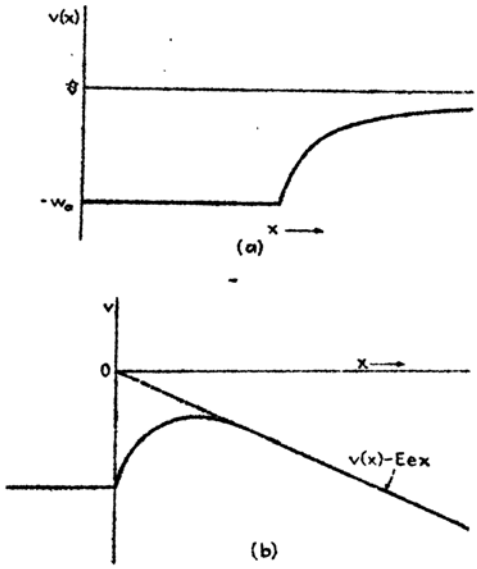

FIG. 9.—*a*, the image-force potential barrier [*cf.* Eq. (3)]; *b*, the effect of a field on the image-force barrier [*cf.* Eq. (5)].

ciently negative. This *field emission*, which can be explained in terms of the quantum mechanical process of penetration through a barrier, will not concern us in the present chapter.

Suppose that we have an electron at a distance x from the surface of an uncharged metal, where x is large compared with an interatomic distance and small compared with the dimensions of the surface. Then the only force that acts upon the electron is the classical attractive image force F_i, which is given by the equation

$$F_i = -\frac{e^2}{4x^2} \tag{1}$$

[1] See the following review articles: S. Dushman, *Rev. Modern Phys.*, 2, 381 (1930); Becker, *op. cit.* See also J. H. de Boer, *Electron Emission and Adsorption Phenomena* (The Macmillan Company, New York, 1935) and A. L. Reimann, *Thermionic Phenomena* (Chapman and Hall, London, 1934).

in which e is the electronic charge. The potential energy associated with this force is clearly

$$V_i = -\frac{e^2}{4x}. \tag{2}$$

The form of the potential may be expected to deviate from this when the surface is not homogeneous, when x is comparable with the interatomic distance, or when x is comparable with the linear dimensions of the surface.

We shall assume for simplicity that the actual potential has the form

$$V(x) = \begin{cases} -\dfrac{e^2}{4x + (e^2/W_a)} & x \geq 0, \\ -W_a & x \leq 0, \end{cases} \tag{3}$$

where the region in which x is negative corresponds to the interior of the metal. It is clear that this function, which is illustrated in Fig. 9, changes continuously from $-W_a$ to zero as we pass from the interior of the metal to infinity and that it approaches (2) for large distances. In an actual case, we might expect dV/dx to be continuous; however, the function (3) does not satisfy this condition.

If the metal is charged negatively, so that the repulsive field is E, the potential

$$V_e = -\mathsf{E}ex \tag{4}$$

must be added to (3). The total field V_t then is

$$V_t = \begin{cases} -\dfrac{e^2}{4x + (e^2/W_a)} - \mathsf{E}ex & x \geq 0 \\ -W_a & x \leq 0. \end{cases} \tag{5}$$

The additional field term in (5) has the effect of lowering the height of the potential barrier at the surface of the metal. The maximum value of V_t is

$$V_m \cong -\sqrt{\mathsf{E}e^3}, \tag{6}$$

as may be proved by solving the equation $dV/dx = 0$.

Hence, when there is a field present, the effective work function φ_{E} is

$$\varphi_{\mathsf{E}} = \varphi - \sqrt{\mathsf{E}e^3}, \tag{7}$$

where φ is the work function defined in Sec. 26.

Let us now compute the number of electrons that evaporate from a unit area of the metal in unit time. We shall assume that the energy ϵ of the electrons inside the metal may be written as a function of the

components of momentum, and we shall choose the x axis to be normal to the surface. Then, as long as the potential is a function of x alone, the y and z components of the momentum of any electron that passes through the surface are preserved. Hence, if a given electron moving toward the surface is to surmount the surface barrier, its total energy ϵ must be greater than the barrier height $W_a - \sqrt{E}e^{\frac{3}{2}} = \epsilon' + \varphi_E$, by an amount $(p_y^2 + p_z^2)/2m$. Thus, we must have

$$\epsilon \geq \epsilon' + \varphi_E + \frac{1}{2m}(p_y^2 + p_z^2) = \epsilon''. \tag{8}$$

The total number of electrons with momenta in the range p_x to $p_x + dp_x$, etc., striking a unit area of the surface in unit time, is

$$n(p_x,p_y,p_z)v_x dp_x dp_y dp_z = n(p_x,p_y,p_z)\frac{\partial \epsilon}{\partial p_x} dp_x dp_y dp_z, \tag{9}$$

since

$$v_x = \frac{\partial \epsilon}{\partial p_x}. \tag{9a}$$

According to Eqs. (4) and (15), Sec. 26, $n(p_x,p_y,p_z)$, which is the number of electrons per unit volume of phase space, is

$$n(p_x,p_y,p_z) = \frac{2}{h^3}\frac{1}{e^{\epsilon(p_x,p_y,p_z)-\epsilon'} + 1}. \tag{10}$$

The total number of electrons ν that strike a unit area in unit time is equal to the integral of (9) over all values of p_x, p_y, and p_z that satisfy the relation (8). The integration over dp_x may be replaced by an integration over the variable ϵ, since

$$\frac{\partial \epsilon}{\partial p_x} dp_x = d\epsilon.$$

The resulting integral is

$$\begin{aligned} \nu &= \frac{2}{h^3}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{\epsilon''}^{\infty} \frac{d\epsilon dp_y dp_z}{e^{\frac{\epsilon-\epsilon'}{kT}} + 1} \\ &= \frac{2kT}{h^3}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \log\left[1 + e^{-\frac{\varphi_E + \frac{(p_y^2+p_z^2)}{2m}}{kT}}\right] dp_y dp_z. \end{aligned} \tag{11}$$

The exponential function appearing in the integrand ordinarily is very small, since φ_E is much greater than kT. Hence, we may expand the logarithm and keep only the first terms. The result is

$$\nu \cong \left(\frac{2kT}{h^3}\right) e^{-\frac{\varphi E}{kT}} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} e^{-\frac{p_y^2+p_z^2}{2mkT}} dp_y dp_z$$
$$= \frac{8\pi m(kT)^2}{h^3} e^{-\frac{\varphi E}{kT}} \int_0^{\infty} e^{-x^2} x dx = \frac{4\pi m(kT)^2}{h^3} e^{-\frac{\varphi E}{kT}}. \tag{12}$$

Equation (12) is the Richardson-Dushman[1] equation which was first derived on the basis of thermodynamics.

If we let r designate the probability that the electrons which have sufficient energy to get over the barrier are reflected back, we find that the current per unit area is

$$I = AT^2(1 - r)e^{-\frac{\varphi}{kT}} e^{-\sqrt{\frac{e^3}{kT}}} \tag{13}$$

where

$$A = \frac{4\pi mek^2}{h^3} = 120 \frac{\text{amp}}{\text{cm}^2\text{deg}^2}. \tag{14}$$

Had we used the classical distribution function

$$n(p_x,p_y,p_z) = \frac{n_0}{(2\pi mkT)^{\frac{3}{2}}} e^{-\frac{\epsilon}{kT}}, \tag{15}$$

where n_0 is the number of electrons per unit volume, we should have found that

$$I = en_0\sqrt{\frac{kT}{2\pi m}}(1 - r)e^{-\frac{W_a}{kT}}. \tag{16}$$

This becomes identical with Eq. (13) at high temperatures if we set

$$W_a = \epsilon' + \varphi$$

and if we recall that

$$e^{-\frac{\epsilon'}{kT}} = \frac{2(2\pi mkT)^{\frac{3}{2}}}{h^3 n_0}.$$

On the other hand, the emission corresponding to Eq. (16) is much larger than that corresponding to (13) if W_a in (16) is regarded as the measured work function. As a matter of fact, Eq. (16) gives the same result as Eq. (13) in this case only if a fraction

$$\frac{3\sqrt{\pi}}{4}\left(\frac{kT}{\epsilon_0}\right)^{\frac{3}{2}}$$

of all electrons is assumed to be free.

The reflection coefficient r is much less than unity if the barrier is described fairly well by the image-force potential. This fact was shown

[1] See footnote 1, p. 162.

first by Nordheim[1] on the basis of wave mechanics. With this simplification, Eq. (13) may be written

$$I = AT^2 e^{-\frac{\varphi}{kT}} e^{-\sqrt{E}\frac{a\mathfrak{F}}{kT}}. \tag{17}$$

The relative dependence of I upon field strength has been found to obey this equation very closely. Table XLII contains values of A and of φ that have been determined by experimental measurements[2] on thermionic emission. It should be observed that A actually does not have the theoretical value (14) in any case. The observed values do not depart from the calculated ones by a very large factor in most cases, but there are large deviations in a few. The possible interpretations of these deviations may be understood by examining the assumptions upon

TABLE XLII.—THERMIONIC DATA FOR SEVERAL METALS

Metal	A, amp/cm^2-deg^2	φ, ev
Ca	$\sim$ 60	3.2
Cs	$\sim$160	1.8
Mo	$\sim$ 60	4.3
Ni	$\sim$ 27	$\sim$5.0
Pt	$\sim$ 10^4	5.0
Ta	$\sim$ 50	4.1
Th	$\sim$ 60	3.4
W	$\sim$ 60	4.5

which the derivation of Eq. (13) was based. We shall present these assumptions categorically.

a. Apparently, we were assuming that the electrons are strictly free inside the metal when we set the mass m that appears in Eq. (8) equal to the electronic mass. Actually, this assumption[3] has not been made, for the momenta and mass in Eq. (8) may be regarded as the values *outside* the metal. It is necessary to assume, however, that the components of momenta in the plane of the surface are preserved as the electrons pass through the surface, if the manner in which Eq. (8) was used in deciding the limits of integration of (11) is to be correct. This assumption is justifiable only if the potential gradient parallel to the surface is zero.

b. We were implicitly assuming that the electrons do not interact appreciably when we used the Fermi distribution function for each

[1] L. W. NORDHEIM, *Proc. Roy. Soc.*, **121**, 626 (1928).

[2] These values of A and φ have been taken from the compilations referred to in footnote 1, p. 162.

[3] *Cf.* Sommerfeld and Bethe, *op. cit.*, p. 436.

electron without including interaction terms. The density of electrons emitted through the surface is so small that it is easy to justify[1] this assumption.

c. The fact that the integration over p_x could be replaced by an integration over ϵ is valid under broader conditions than those contained in our model. In wave mechanics, the integration variable that replaces the x component of momentum is h times the x component of electronic wave number k_x. It turns out that the group velocity in the x direction v_x, is related to the energy by the equation[2]

$$v_x = \frac{1}{h}\frac{\partial \epsilon}{\partial k_x},$$

which is equivalent to the relation

$$v_x = \frac{\partial \epsilon}{\partial p_x}$$

that we used previously.

d. We also assumed that the surface is plane. There seems to be little doubt that even the smoothest surfaces are rough in a submicroscopic sense. Metal surfaces are apparently made[3] of many different crystallographic planes which are inclined relative to one another. The form of the potential function before each plane depends both upon the crystallographic orientation of the plane and upon the nature of any contamination that may be present on it. Although the total potential difference between any point inside the metal and a point at infinity is W_a when there is no external field, the barrier before some of the surfaces may rise to values higher than W_a and drop to W_a at larger distances. Since the highest point of the barrier determines the work function, we should expect the entire surface to behave as though composed of many surfaces which emit more or less in accordance with Eq. (13) but which have different work functions. The variations in surface potential from point to point also imply that there is a tangential force. The existence of this force makes the process of separating the integral (11) into independent integrals over p_x and over p_y and p_z an approximation.

e. The assumption that the reflection coefficient r is zero seems to be justifiable for any reasonably clean surface. It is not justifiable,[4] how-

[1] This does not mean that the electrons will not congregate outside the metal and give rise to space-charge effects.

[2] See Chap. VIII.

[3] Direct evidence for this has been given by R. P. Johnson and W. Shockley, *Phys. Rev.*, **49**, 436 (1936).

[4] Evidence for other barriers has been presented by W. A. Nottingham, *Phys. Rev.*, **49**, 78 (1936).

ever, if the potential function has a large peak at a distance comparable with the electronic wave length before it approaches the image-force value. There is no reason for expecting such peaks for clean metal surfaces; they may occur, however, when the surface has been oxidized or when layers of other foreign atoms have been absorbed.

f. Finally, it should be pointed out that in deriving Eq. (13) we have assumed that the work function is independent of temperature. Since metals are heated to very high temperatures during thermionic experiments, this assumption is not justifiable.[1] This topic will be discussed further in Chap. XI.

By way of summary, it may be said that the deviations from the Richardson-Dushman equation probably arise from a combination of effects that are connected with the composite nature of metal surfaces and the temperature dependence of the work function.

31. Boltzmann's Equation of State; Lorentz's Solution*.—Let us consider a system of particles that is in dynamic equilibrium under external forces. For example, the system may consist of the electrons in a metal that is acted upon by stationary external electric and magnetic fields. When the steady-state current is flowing, this system is in a state of dynamic equilibrium of the type we wish to consider.

Let

$$f_n(x,y,z,v_x,v_y,v_z)dxdydzdv_xdv_ydv_z \tag{1}$$

be the number of particles having position coordinates in the range from x to $x + dx$, etc., and velocity coordinates in the range from v_x to $v_x + dv_x$, etc. We may obtain a condition on f_n, whenever the system is in a steady state, by asking that it should be independent of time. Now, f may vary with time in two independent ways: (1) It may vary because particles are moving from one region of space to another and are accelerated by the external field during this motion. This variation, which takes place continuously, is called the "drift variation" and may be evaluated in the following way. The number of particles that, at time $t + dt$, have drifted to the cell of phase space corresponding to the coordinates x, y, z, v_x, v_y, v_z must be equal to the number that were in the cell located at $x - v_xdt$, $y - v_ydt$, $z - v_zdt$, $v_x - \alpha_xdt$, $v_y - \alpha_ydt$, $v_z - \alpha_zdt$ at time t, where α_x, α_y, α_z are the components of acceleration. This relationship holds only for a time interval dt so short that collisions have not had a large effect on the distribution. Thus, the change due to drift in the number of particles having coordinates x, y, and z and velocity v_x, v_y, v_z in a time dt is

[1] See J. A. BECKER and W. H. BRATTAIN, *Phys. Rev.*, **45**, 694 (1934).

$$(\Delta f)_d = f_n(x - v_x dt, y - v_y dt, z - v_z dt, v_x - \alpha_x dt, v_y - \alpha_y dt, v_z - \alpha_z dt, t) - f_n(x,y,z,v_x,v_y,v_z,t)$$

$$= -\left(\frac{\partial f_n}{\partial x}v_x + \frac{\partial f_n}{\partial y}v_y + \frac{\partial f_n}{\partial z}v_z + \frac{\partial f_n}{\partial v_x}\alpha_x + \frac{\partial f_n}{\partial v_y}\alpha_y + \frac{\partial f_n}{\partial v_z}\alpha_z\right)dt.$$

Consequently the rate of change of f caused by drift is

$$\left(\frac{df_n}{dt}\right)_d = -\frac{\partial f_n}{\partial x}v_x - \frac{\partial f_n}{\partial y}v_y - \frac{\partial f_n}{\partial z}v_z - \frac{\partial f_n}{\partial v_x}\alpha_x - \frac{\partial f_n}{\partial v_y}\alpha_y - \frac{\partial f_n}{\partial v_z}\alpha_z. \tag{2}$$

(2) f may vary because of the relatively discontinuous changes in velocity that accompany collisions. If

$$\Theta(v_x,v_y,v_z;v'_x,v'_y,v'_z)dv'_x dv'_y dv'_z \tag{3}$$

is the probability per unit time that a particle will change its velocity from v_x, v_y, v_z to a value having components in the ranges extending from v'_x to $v'_x + dv'_x$, etc., the total number the velocity of which alters from v_x, v_y, v_z to some other value is

$$a = f_n(x,y,z,v_x,v_y,v_z)\int \Theta(v_x,v_y,v_z;v'_x,v'_y,v'_z)dv'_x dv'_y dv'_z. \tag{4}$$

Similarly, the number the velocity of which changes to v_x, v_y, v_z from another value is

$$b = \int f_n(v''_x,v''_y,v''_z)\Theta(v''_x,v''_y,v''_z;v_x,v_y,v_z)dv''_x dv''_y dv''_z. \tag{5}$$

Thus, the rate of change of f_n caused by collisions is

$$\left(\frac{df_n}{dt}\right)_c = b - a. \tag{6}$$

The total rate of change of f_n is the sum of (2) and (6). The condition for equilibrium is that this sum should vanish or that

$$\frac{\partial f_n}{\partial x}v_x + \frac{\partial f_n}{\partial y}v_y + \frac{\partial f_n}{\partial z}v_z + \frac{\partial f_n}{\partial v_x}\alpha_x + \frac{\partial f_n}{\partial v_y}\alpha_y + \frac{\partial f_n}{\partial v_z}\alpha_z = b - a, \tag{7}$$

which is Boltzmann's equation[1] of state.

In a homogeneous specimen of metal that is at constant temperature in a field-free space, the components of the gradient of f_n, $\partial f_n/\partial x$, $\partial f_n/\partial y$, $\partial f_n/\partial z$, and the components of acceleration α_x, α_y, α_z vanish. Equation (7) then reduces to

$$a = b$$

which states that the numbers of particles that leave and enter a given volume of momentum space as a result of collisions are equal. On the

[1] L. BOLTZMANN, *Vorlesung über Gastheorie* (J. A. Barth, Leipzig, 1923).

other hand, the left-hand side of Eq. (7) does not vanish if there is a temperature gradient in the metal or if there is an external field. Hence, a is not equal to b in this case.

Suppose that we have homogeneous electrical fields in the x and y directions and a homogeneous magnetic field in the z direction. According to classical mechanics, the acceleration α is given by the equation

$$m\alpha = -e\mathsf{E} - \frac{e}{c}\mathbf{v} \times \mathsf{H} \tag{8}$$

where E and H are the electric and magnetic fields, respectively, and e is the absolute value of the electronic charge. For the field assumed above, we have

$$\left.\begin{aligned} m\alpha_x &= -\left(e\mathsf{E}_x + \frac{e}{c}v_y\frac{\mathsf{H}_z}{v_g}\right), \\ m\alpha_y &= -\left(e\mathsf{E}_y - \frac{e}{c}v_x\mathsf{H}_z\right), \\ m\alpha_z &= 0. \end{aligned}\right\} \tag{9}$$

Hence, Eq. (7) becomes

$$\frac{\partial f_n}{\partial x}v_x + \frac{\partial f_n}{\partial y}v_y + \frac{\partial f_n}{\partial z}v_z - \frac{\partial f_n}{\partial v_x}\left(\frac{e\mathsf{E}_x}{m} + \frac{e}{c}\frac{v_y}{m}\mathsf{H}_z\right) - \frac{\partial f_n}{\partial v_y}\left(\frac{e}{m}\mathsf{E}_y - \frac{e}{c}\frac{v_x}{m}\mathsf{H}_z\right) = b - a. \tag{10}$$

We shall discuss this equation in the case first treated by Lorentz,[1] namely, when the medium is homogeneous and isotropic, no magnetic field is present, and the electrical field is in the x direction. Equation (10) then reduces to

$$\frac{\partial f_n}{\partial x}v_x - \frac{\partial f_n}{\partial v_x}\frac{e\mathsf{E}_x}{m} = b - a \tag{11}$$

since f_n does not depend upon y and z.

Lorentz simplified the collision terms by making the following three assumptions.

1. The electrons undergo only elastic collisions. This assumption seemed to be reasonable at the time it was made since it had been postulated that the electrons were deflected principally by direct collisions with the ions. It may be shown that electrons, being relatively light, would lose little energy in such processes. This interpretation of the electronic collisions is not accepted as completely rigorous at the present time. Nevertheless we shall employ Lorentz's assumption since the results that may be derived from it have semiquantitative value.

[1] H. A. LORENTZ, *The Theory of Electrons* (Teubner, Leipzig, 1909, and G. E. Stechert & Company, New York, 1923).

2. The electronic scattering is isotropic; that is, Θ in (3) is independent of the relative directions of the velocity before and after collisions.

3. The distribution function f_n, for the case when a field is present, is related to the function f_n^0, for the case when no field is present, by the equation

$$f_n = f_n^0 + v_x\chi(v) \tag{12}$$

where χ is a small undetermined function that depends upon the velocity only through the speed $v = \sqrt{v_x^2 + v_y^2 + v_z^2}$. Equation (12) may be regarded as expressing the form of f_n when it is expanded as a series in powers of E_x and only first-power terms are retained.

According to assumption 1, the probability Θ is zero unless

$$v = v'.$$

It is important to note that Θ should be infinite when this condition is satisfied if the total probability of a collision, namely,

$$\int \Theta(v_x,v_y,v_z;v_x',v_y',v_z')dv_x'dv_y'dv_z',$$

is different from zero. In order to avoid the mathematical difficulties that accompany the use of a discontinuous function, we shall introduce a new collision function

$$\eta(v;\theta,\varphi;\theta',\varphi') \sin \theta' d\theta' d\varphi' \tag{13}$$

which gives the probability that a particle that is traveling with speed v in the direction described by the polar angles θ, φ (*cf.* Fig. 10) is deflected into the solid angle $\sin \theta' d\theta' d\varphi'$ in the direction θ', φ' without a change in speed. Θ and η are obviously connected by the equation

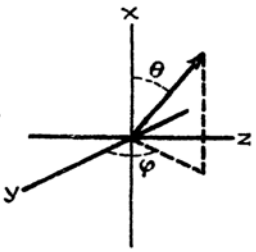

FIG. 10.—The polar angles φ and θ.

$$\eta(v;\theta,\varphi;\theta'\varphi') = \int_0^\infty \Theta(v_x,v_y,v_z;v_x',v_y',v_z')v'^2dv'. \tag{14}$$

According to assumption 2, η is a constant; hence, we see that

$$a = f_n\eta 4\pi, \tag{15}$$

or according to Eq. (12),

$$a = [f_n^0(v) + v_x\chi(v)]4\pi\eta. \tag{16}$$

Substituting (12) in the expression (5), we find

$$b = f_n^0 4\pi\eta + v\chi(v)\eta \int \frac{v_x'}{v'} \sin \theta' d\theta' d\varphi'.$$

The second integral vanishes since $v_x'/v' = \cos\theta'$. Hence,

$$b = f_n^0(v)4\pi\eta, \tag{17}$$

and

$$b - a = -v_x\chi(v)4\pi\eta. \tag{18}$$

We might have foreseen that $b - a$ would not contain f_n^0 from the fact that a is equal to b when E_x is zero.

Since the factor $4\pi\eta$ is the total number of collisions an electron makes in a second, $4\pi\eta/v$ is the total number that it makes in traveling 1 cm, and the reciprocal of this quantity is the mean free path l. Thus,

$$l = \frac{v}{4\pi\eta}. \tag{19}$$

We may now simplify the drift terms by using Eq. (12). Substituting (12) in the left-hand side of Eq. (11), we find

$$\left(\frac{df_n^0}{dt}\right)_{\text{drift}} = v_x\frac{\partial f_n^0}{\partial x} + v_x^2\frac{\partial\chi}{\partial x} - \frac{e}{m}\mathsf{E}_x\frac{\partial f_n^0}{\partial v_x} - \frac{e}{m}\mathsf{E}_x v_x\frac{\partial\chi}{\partial v_x} - \frac{e}{m}\mathsf{E}_x\chi. \tag{20}$$

We may drop the terms in χ since, by assumption, they are considerably smaller than those in f_n^0. Hence, Eq. (11) becomes

$$v_x\frac{\partial f_n^0}{\partial x} - \frac{e}{m}\mathsf{E}_x\frac{\partial f_n^0}{\partial v_x} = -\frac{v_x v}{l}\chi(v). \tag{21}$$

We may replace the derivative with respect to v_x by one with respect to ϵ if we use the relation

$$\epsilon = \frac{m}{2}v^2 = \frac{m}{2}(v_x^2 + v_y^2 + v_z^2).$$

Equation (21) then becomes

$$\frac{\partial f_n^0}{\partial x} - e\mathsf{E}_x\frac{\partial f_n^0}{\partial\epsilon} = -\frac{v}{l}\chi(v). \tag{22}$$

Solving this equation for χ we obtain

$$\chi = -\frac{l}{v}\left(\frac{\partial f_n^0}{\partial x} - e\mathsf{E}_x\frac{\partial f_n^0}{\partial\epsilon}\right), \tag{23}$$

which relates the unknown function χ to the properties of the known distribution function f_n^0.

The process of reasoning that leads to Eq. (23) may be summarized as follows: In using (12), we assume that the distribution function is modified by the addition of a small term if an electric field is present. The form of this term implies that there are more electrons moving

in the direction of the force than in the opposite direction, since this term is negative for negative values of v_x if it is positive for positive ones. Since the collision probability is unaffected by the field, the contribution of f_0 to the collision terms is zero in the presence of a field, just as in the absence of one, and the collision terms depend only upon χ. The drift terms are calculated by considering the effect of the field on the unperturbed distribution function. It is assumed that the small additional term in χ does not appreciably affect the drift terms.

The equation for χ reduces to the simple form (23) only as a result of the simplifying assumptions that were made concerning the collision probability Θ. It would not be possible to remove χ from under the integral sign in the expression for b if the particle speeds were not conserved during collisions. Hence, we should arrive at an integral equation in place of (23) in a more general case. This occurs, for example in the quantum mechanical treatment[1] of the equation of state for electrons in metals, since the electrons are found to be inelastically scattered by the lattice when the collision process is examined in the light of quantum mechanical laws.

Gans[2] has generalized the Lorentz equation to include the case in which there are both electric and magnetic fields present. We shall assume that the electric field is in the x and the y directions and that the magnetic field is in the z direction, so that the equation of state is (10). In place of (12), Gans assumes that

$$f_n = f_n^0 + v_x\chi_1 + v_y\chi_2 \tag{24}$$

where χ_1 and χ_2 depend upon the velocity components only through the speed v. In place of Eq. (18) he then obtains

$$b - a = -\frac{v}{l}(v_x\chi_1 + v_y\chi_2). \tag{25}$$

We may substitute Eq. (24) in the drift term of Eq. (9). The expression may be simplified by setting

$$\frac{\partial}{\partial v_x} = mv_x\frac{\partial}{\partial \epsilon} \qquad \text{and} \qquad \frac{\partial}{\partial v_y} = mv_y\frac{\partial}{\partial \epsilon}.$$

Many terms then cancel, and the result is

$$-\left(\frac{\partial f_n}{\partial t}\right)_d = v_x\frac{\partial f_n^0}{\partial x} + v_y\frac{\partial f_n^0}{\partial y} - \frac{\partial f_n^0}{\partial \epsilon}(v_x e\mathsf{E}_x + v_y e\mathsf{E}_y) - \frac{e}{c}\frac{v_y}{m}\mathsf{H}_z\chi_1 + \frac{e}{c}\frac{v_x}{m}\mathsf{H}_z\chi_2. \tag{26}$$

[1] See Chap. XV.

[2] R. GANS, *Ann. Physik*, **20**, 293 (1906).

Terms that involve products of E_x or E_y and χ_1, χ_2 or involve derivatives of the χ have been regarded as negligible in comparison with the terms in E_x and E_y that are retained.

Equating coefficients of v_x and of v_y in Eqs. (25) and (26), we obtain

$$\left.\begin{aligned} \frac{\partial f_n^0}{\partial x} - c\mathsf{E}_x\frac{\partial f_n^0}{\partial \epsilon} &= -\frac{v}{l}\chi_1 - \frac{e}{mc}\mathsf{H}_z\chi_2, \\ \frac{\partial f_n^0}{\partial y} - e\mathsf{E}_y\frac{\partial f_n^0}{\partial \epsilon} &= \frac{e}{mc}\mathsf{H}_z\chi_1 - \frac{v}{l}\chi_2. \end{aligned}\right\} \tag{27}$$

The result of solving these equations for χ_1 and χ_2 is

$$\left.\begin{aligned} \chi_1 &= -\frac{l}{v}\frac{f_1 - sf_2}{s^2+1}, \\ \chi_2 &= -\frac{l}{v}\frac{f_2 + sf_1}{s^2+1}, \end{aligned}\right\} \tag{28}$$

where

$$s = \frac{el}{mvc}\mathsf{H}_z = \frac{l}{v}k, \qquad k = \frac{e\mathsf{H}_z}{mc}, \tag{29}$$

and

$$\left.\begin{aligned} f_1 &= \frac{\partial f_n^0}{\partial x} - e\mathsf{E}_x\frac{\partial f_n^0}{\partial \epsilon}, \\ f_2 &= \frac{\partial f_n^0}{\partial y} - e\mathsf{E}_y\frac{\partial f_n^0}{\partial \epsilon}. \end{aligned}\right\} \tag{30}$$

32. Electrical and Thermal Conductivity*.—The electrical and thermal currents[1] $\mathbf{i}$ and $\mathbf{c}$ that are associated with an electron having velocity $\mathbf{v}$ are, respectively,

$$\mathbf{i} = -e\mathbf{v} \qquad \text{and} \qquad \mathbf{c} = \mathbf{v}\frac{mv^2}{2}. \tag{1}$$

The total electrical and thermal currents I_x and C_x that pass through a unit area at x, y, z, normal to the x direction of a metal, may be expressed by the integrals

$$I_x = -\int ev_x f_n d\sigma, \tag{2a}$$

$$C_x = \int v_x\frac{1}{2}mv^2 f_n d\sigma, \tag{2b}$$

where f_n is the distribution function of electron velocities, $d\sigma$ is $dv_x dv_y dv_z$, and the integral extends over all values of v_x, v_y, and v_z. I_x and C_x depend upon x, y, and z whenever f_n is dependent upon these variables.

[1] First treated by Sommerfeld, *op. cit.*

The integrals in (2*a*) and (2*b*) vanish when f_n is dependent on the components of velocity only through the scalar velocity, for then the current associated with electrons traveling in one direction is exactly compensated by the current associated with electrons traveling in the opposite direction. Thus, the contribution to the currents from f_n^0 in Eqs. (12) and (24) in Sec. 31 is zero.

Let us consider the form of Eqs. (2*a*) and (2*b*) in the case in which there are an electric field and a thermal gradient in the x direction. The distribution function then is given by Eq. (12) of the preceding section where χ satisfies Eq. (23) of the same section. Consequently Eqs. (2*a*) and (2*b*) become

$$I_x = e\int \frac{v_x^2 l}{v}\left(\frac{\partial f_n^0}{\partial x} - e\mathsf{E}_x\frac{\partial f_n^0}{\partial \epsilon}\right)d\sigma, \tag{3a}$$

$$C_x = -\frac{m}{2}\int v_x^2 l v\left(\frac{\partial f_n^0}{\partial x} - e\mathsf{E}_x\frac{\partial f_n^0}{\partial \epsilon}\right)d\sigma. \tag{3b}$$

The quantities in parentheses in each of these integrals involve v_x, v_y, and v_z only through the scalar velocity v. Hence, we may replace v_x^2 by $v^2/3$, $d\sigma$ by $4\pi v^2 dv$, and the triple integration by a single integral of v extending from zero to infinity. Thus,

$$I_x = \frac{4\pi e}{3}\int_0^\infty v^3 l\left(\frac{\partial f_n^0}{\partial x} - e\mathsf{E}_x\frac{\partial f_n^0}{\partial \epsilon}\right)dv, \tag{4a}$$

$$C_x = -\frac{4\pi m}{6}\int_0^\infty v^5 l\left(\frac{\partial f_n^0}{\partial x} - e\mathsf{E}_x\frac{\partial f_n^0}{\partial \epsilon}\right)dv. \tag{4b}$$

a. Electrical Conductivity.—The simplest case to consider is that in which the temperature of the metal is constant. Then, f_n^0 is independent of x, and the current is

$$I_x = -\frac{4\pi e^2}{3}\mathsf{E}_x\int_0^\infty v^3 l\frac{\partial f_n^0}{\partial \epsilon}dv. \tag{5}$$

We shall use the Fermi-Dirac distribution function for metals. Thus,

$$f_n^0(v_x,v_y,v_z) = \frac{2m^3}{h^3}f(\epsilon) \tag{6}$$

where

$$f(\epsilon) = \frac{1}{e^{\frac{(\epsilon-\epsilon')}{kT}} + 1}. \tag{7}$$

At ordinary temperatures,

$$\epsilon' \cong \epsilon_0' = \frac{h^2}{2m}\left(\frac{3n_0}{8\pi}\right)^{\frac{2}{3}} \tag{8}$$

where n_0 is the number of electrons per unit volume (*cf.* Sec. 26).

It is convenient to change the variable of integration from v to ϵ. Equation (5) then becomes

$$I_x = -\mathsf{E}_x \frac{e^2}{3} \frac{C}{V} \sqrt{\frac{2}{m}} \int \epsilon l \frac{\partial f(\epsilon)}{\partial \epsilon} d\epsilon \tag{9}$$

where [*cf.* Eq. (9), Sec. 26]

$$\frac{C}{V} = \frac{4\pi(2m)^{\frac{3}{2}}}{h^3} = \frac{3}{2} \frac{n_0}{\epsilon_0'^{\frac{3}{2}}}. \tag{10}$$

We may now employ the approximation expressed by Eq. (29), Sec. 26, and write

$$\int_0^\infty \epsilon l \frac{\partial f(\epsilon)}{\partial \epsilon} d\epsilon = -\epsilon_0' l(\epsilon_0') \tag{11}$$

where $l(\epsilon_0')$ is the value of l for $\epsilon = \epsilon_0'$. This approximation is very accurate for ordinary valence electrons. Then, I_x becomes

$$\begin{aligned} I_x &= \mathsf{E}_x \frac{e^2}{3} \frac{C}{V} \sqrt{\frac{2}{m}} \epsilon_0' l(\epsilon_0') \\ &= \mathsf{E}_x \frac{e^2}{2} \frac{n_0}{\epsilon_0'^{\frac{1}{2}}} \sqrt{\frac{2}{m}} l(\epsilon_0') \\ &= \mathsf{E}_x \frac{e^2 n_0 l(\epsilon_0')}{m v(\epsilon_0')}. \end{aligned} \tag{12}$$

Hence, the electrical conductivity is

$$\sigma = \frac{I_x}{\mathsf{E}_x} = \frac{e^2 n_0 l(\epsilon_0')}{m v(\epsilon_0')}. \tag{13}$$

According to this equation, the conductivity depends upon temperature through the factor l alone, since all other quantities in Eq. (13) should be practically constant.

b. Thermal Conductivity.—The electronic thermal conductivity may be obtained by solving Eqs. (4*a*) and (4*b*) for C_x in the case in which there is a uniform temperature gradient dT/dx in the x direction and in which there is no electrical current flowing. The thermal conductivity κ is defined by the ratio

$$\kappa = -C_x \frac{dT}{dx}. \tag{14}$$

$\partial f_n^0/\partial x$ is not zero when there is a temperature gradient. In fact, if we recall that f_n^0 is a function of $\alpha = (\epsilon - \epsilon')/kT$ [*cf.* Eqs. (6) and (7)], we see that

$$\frac{\partial f_n^0}{\partial x} = \frac{\partial f_n^0}{\partial T}\frac{dT}{dx} = kT\frac{\partial f_n^0}{\partial \epsilon}\frac{d\alpha}{dT}\frac{dT}{dx} = -\frac{\partial f_n^0}{\partial \epsilon}\left[\frac{\epsilon}{T} + T\frac{d}{dT}\left(\frac{\epsilon'}{T}\right)\right]\frac{dT}{dx}. \tag{15}$$

Equations (4*a*) and (4*b*) now become

$$I_x = -\frac{e}{3}\sqrt{\frac{2}{m}}\left\{K_1\left[eE_x + T\frac{d}{dT}\left(\frac{\epsilon'}{T}\right)\frac{dT}{dx}\right] + \frac{K_2}{T}\frac{dT}{dx}\right\}, \tag{16a}$$

$$C_x = \frac{1}{3}\sqrt{\frac{2}{m}}\left\{K_2\left[eE_x + T\frac{d}{dT}\left(\frac{\epsilon'}{T}\right)\frac{dT}{dx}\right] + \frac{K_3}{T}\frac{dT}{dx}\right\}, \tag{16b}$$

where

$$K_i = \frac{C}{V}\int_0^\infty \epsilon^i l\frac{df}{d\epsilon}d\epsilon. \tag{17}$$

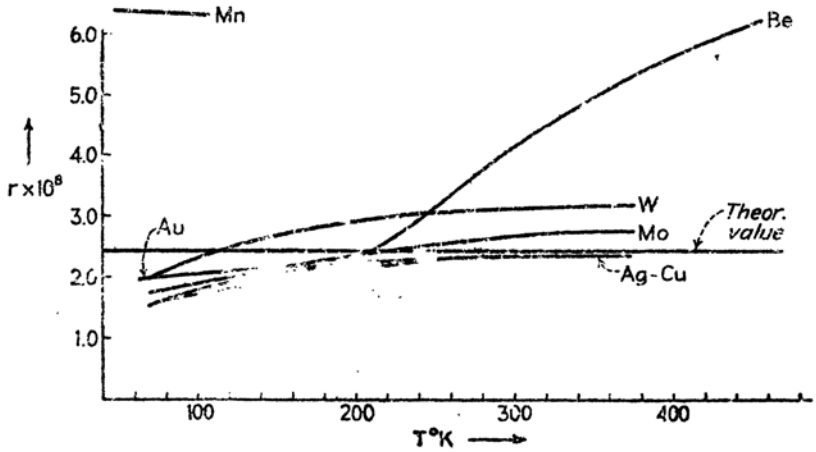

FIG. 11.—The Wiedemann-Franz ratio for several metals. The ratio $r(r = \kappa/T\sigma)$ is given in units in which σ is measured in ohms^{-1} cm^{-1}, κ is measured in watts deg^{-1} cm^{-1}, and T is the absolute temperature. The theoretical value is 2.45×10^{-8} watt-ohm/deg^2

We may eliminate E_x from (16*b*) by substituting this quantity from (16*a*), after setting $I_x = 0$. The result is

$$\kappa = -\frac{C_x}{dT/dx} = -\frac{1}{3}\sqrt{\frac{2}{m}}\left(\frac{K_3K_1 - K_2^2}{K_1T}\right). \tag{18}$$

According to Eq. (29), Sec 26,

$$K_i = -\frac{C}{V}\left\{\epsilon_0'^i l(\epsilon_0') + \frac{\pi^2}{6}k^2T^2\left[\frac{d^2(l\epsilon^i)}{d\epsilon^2}\right]_{\epsilon=\epsilon_0'}\right\}. \tag{19}$$

It is obvious that the total contribution to (18) from the first term of Eq. (19) is zero. The principal contribution from the second term may be reduced to

$$\kappa = \frac{\pi^2}{3}\frac{k^2n_0l(\epsilon_0')}{mv(\epsilon_0')}T \tag{20}$$

by straightforward manipulation. We find that

$$\frac{\kappa}{T\sigma} = \frac{\pi^2}{3}\left(\frac{k}{e}\right)^2, \tag{21}$$

upon comparing this expression with the one for the conductivity. In other words, the free-electron theory predicts that $\kappa/T\sigma$ should be a universal constant. A relationship of this type was observed first by Wiedemann and Franz;[1] for this reason, the ratio $\kappa/T\sigma$ is called the Wiedemann-Franz ratio. Experimental values usually differ somewhat from the theoretical ones. Figure 11 shows plots of the observed[2] ratio for several metals. Silver, copper, and gold obey the theoretical relation fairly well at high temperatures but do not at low temperatures. On the other hand, beryllium and β manganese show large deviations.[3]

33. Electrothermal Effects*.—The rate at which heat accumulates in a unit volume of a wire that is carrying both electrical and thermal currents in the x direction is

$$\frac{dH}{dt} = I_x \mathsf{E}_x - \frac{\partial C_x}{\partial x}. \tag{1}$$

The first term in this equation is the electrical work done, and the second is the divergence of the heat current which, according to the equation of continuity, is the rate at which heat flows into the volume. We may solve Eq. (16a) of the previous section for E_x thus:

$$\mathsf{E}_x = -\frac{3}{e^2}\sqrt{\frac{m}{2}}\frac{1}{\mathrm{K}_1}I_x - \frac{T}{e}\frac{d}{dT}\left(\frac{\epsilon'}{T}\right)\frac{dT}{dx} - \frac{\mathrm{K}_2}{\mathrm{K}_1}\frac{1}{Te}\frac{dT}{dx}. \tag{2}$$

The expression (16b) for C_x then becomes

$$C_x = -\frac{\mathrm{K}_2}{\mathrm{K}_1}\frac{1}{e}I_x + \frac{1}{3}\sqrt{\frac{2}{m}}\left(\frac{\mathrm{K}_2\mathrm{K}_1 - \mathrm{K}_2^2}{T\mathrm{K}_1}\right)\frac{dT}{dx} \tag{3}$$

when E_x is replaced by (2). If we now substitute Eqs. (2) and (3) in Eq. (1), we obtain

$$\frac{dH}{dt} = \frac{I_x^2}{\sigma} + \frac{I_x T}{e}\frac{d}{dx}\left(\frac{\mathrm{K}_2}{T\mathrm{K}_1} - \frac{\epsilon'}{T}\right) + \frac{d}{dx}\left(\kappa\frac{dT}{dx}\right) \tag{4}$$

where σ and κ are the electrical and thermal conductivities, which are given, respectively, by Eqs. (13) and (20) of the preceding section.

[1] WIEDEMANN and FRANZ, *Ann. Physik*, **89**, 497 (1853).

[2] Taken from Landolt-Bornstein. The measurements were made principally by W. Meissner.

[3] These deviations are most probably associated with the assumption that the electrons are perfectly free.

The first term in Eq. (4) is the Joule heat, and the third is that arising from ordinary flow of heat. Both these terms are independent of the relative directions of the electrical current and the thermal gradient. The second term, on the other hand, represents an accumulation of heat that depends upon both the direction of the current and the direction of the temperature gradient. The influence of this term is usually considered in three separate cases which we shall now discuss.

a. The Thomson Effect.—Suppose that we have a wire carrying a current I_x and having a uniform temperature gradient dT/dx. Then, in accordance with the second term in (4), heat will be produced at a rate

$$\frac{dH}{dt} = \frac{I_x}{e} T \frac{d}{dT}\left(\frac{K_2}{TK_1} - \frac{\epsilon'}{T}\right)\frac{dT}{dx}. \tag{5}$$

This is known as the Thomson heat, and the negative of the coefficient of $I_x dT/dx$, namely,

$$\sigma_T = -\frac{T}{e}\frac{d}{dT}\left(\frac{K_2}{TK_1} - \frac{\epsilon'}{T}\right), \tag{6}$$

is called the Thomson coefficient. The quantity in the parenthesis of (6) may be evaluated by use of Eq. (19) of the preceding section. The first-order terms cancel and the second-order terms lead to the result

$$\frac{1}{T}\left(\frac{K_2}{K_1} - \epsilon'\right) = k^2 T \frac{\pi^2}{6}\left[\frac{1}{l\epsilon'}\frac{d^2(l\epsilon'^2)}{d\epsilon'^2} - \frac{1}{l}\frac{d^2(l\epsilon')}{d\epsilon'^2}\right] = k^2 T \frac{\pi^2}{3}\left[\frac{1}{\epsilon_0'} + \frac{l'(\epsilon_0')}{l(\epsilon_0')}\right]. \tag{7}$$

Hence, σ_T is

$$\sigma_T = -\frac{k^2 T \pi^2}{e}\frac{1}{3}\left[\frac{1}{\epsilon_0'} + \frac{l'(\epsilon_0')}{l(\epsilon_0')}\right]. \tag{8}$$

The quantity $l'(\epsilon_0')/l(\epsilon_0')$ cannot be determined without a theory of the mean free path. We shall treat this as an unknown and shall use measured values of σ_T to evaluate it. Values of $\epsilon_0' l'(\epsilon_0')/l(\epsilon_0')$ and of σ_T/T are given in Table XLIII for the alkali metals.[1] It should be noticed that the observed values are similar for all cases except Li which has a different sign and is extremely large. Whether this difference is real or is due to experimental error remains to be seen. The negative value of σ_T for most of the alkalies shows that the electrons carry heat from hot to cold regions of the metal.

b. Peltier Effect.—When a current passes from one metal to another that is at the same temperature (for example, from 1 to 2), some heat usually is evolved or absorbed at the junction. This is known as the Peltier heat: the Peltier coefficient $\pi_{1\to 2}$ is defined by the relation

[1] See SOMMERFELD and FRANK, *op. cit.*

$$\pi_{1\to 2} = -\frac{dH}{dT} I_x$$

where dH/dT is the heat evolved per unit area of the junction. According to the second term of Eq. (4),

$$\pi_{1\to 2} = -\frac{T}{e}\int_1^2 \frac{d}{dx}\left(\frac{K_2}{TK_1} - \frac{\epsilon'}{T}\right)dx$$
$$= -\frac{T}{e}\left[\left(\frac{K_2}{TK_1} - \frac{\epsilon'}{T}\right)_2 - \left(\frac{K_2}{TK_1} - \frac{\epsilon'}{T}\right)_1\right]. \tag{9}$$

Hence, comparing this with Eq. (6), we see that

$$\frac{d}{dT}\left(\frac{\pi_{1\to 2}}{T}\right) = \frac{\sigma_{T,2} - \sigma_{T,1}}{T}. \tag{10}$$

This equation may be derived on the basis of purely thermodynamical reasoning.[1] Thus, the agreement between the observed and computed values of (9) should be neither better nor worse than the agreement between observed and computed values of σ_T.

TABLE XLIII

	σ_T/T observed, microvolts/deg^2	$\epsilon_0' l'(\epsilon_0')/l(\epsilon_0')$
Li	+0.40	−76
Na	−0.0282	2.69
K	−0.0275	1.29
Rb	−0.069	4.05
Cs	−0.062	2.88

c. The Seebeck Effect.—When the junctions of two metals 1 and 2 that are connected in series to form a closed loop are kept at different temperatures T' and T'', an emf acts within the circuit. The Seebeck emf F_S is defined as the value of the total force when the current is zero. F_S may be computed by integrating the expression (2) for E_x around the circuit

$$F_S = -\oint E_x dx = \frac{1}{e}\oint\left[T\frac{d}{dx}\left(\frac{\epsilon'}{T}\right) + \frac{K_2}{K_1 T}\frac{dT}{dx}\right]dx$$
$$= \frac{1}{e}\oint\left(-\frac{\epsilon'}{T}\frac{dT}{dx} + \frac{K_2}{K_1 T}\frac{dT}{dx}\right)dx. \tag{11}$$

[1] These effects and others are discussed from the thermodynamical viewpoint by P. W. Bridgman, *The Thermodynamics of Electrical Phenomena in Metals* (The Macmillan Company, New York, 1934).

We may now change the variable of integration from x to T and divide the integral (11) into integrals over each metal. Then,

$$F_S = -\frac{1}{e}\int_{T'}^{T''}\left[\left(\frac{K_2}{K_1 T} - \frac{\epsilon'}{T}\right)_2 - \left(\frac{K_2}{K_1 T} - \frac{\epsilon'}{T}\right)_1\right]dT.$$

Hence, according to (9),

$$F_S = \int_{T'}^{T''} \frac{\pi_{1\to 2}}{T}\, dT. \tag{12}$$

This equation also may be derived on a purely thermodynamical basis.

34. The Isothermal Hall Effect*.—Let us consider a metal that has electric fields in the x and y directions and a magnetic field in the z direction. The equation of state then is Eq. (24), Sec. 31, in which χ_1 and χ_2 satisfy Eqs. (28) of the same section. It may be seen that the electrical and thermal currents in the x and y directions are

$$I_x = -\frac{4\pi e}{3}\int v^2\chi_1 d\sigma = \frac{4\pi e}{3}\int vl\frac{f_1 - sf_2}{s^2+1}v^2dv, \tag{1a}$$

$$I_y = -\frac{4\pi e}{3}\int v^2\chi_2 d\sigma = \frac{4\pi e}{3}\int vl\frac{f_2 + sf_1}{s^2+1}v^2dv, \tag{1b}$$

$$C_x = \frac{4\pi m}{6}\int v^4\chi_1 d\sigma = -\frac{4\pi m}{6}\int v^3 l\frac{f_1 - sf_2}{s^2+1}v^2dv, \tag{1c}$$

$$C_y = \frac{4\pi m}{6}\int v^4\chi_2 d\sigma = -\frac{4\pi m}{6}\int v^3 l\frac{f_2 + sf_1}{s^2+1}v^2dv. \tag{1d}$$

We shall consider the emf induced in the y direction, when there is a current flowing in the x direction and when the metal is at constant temperature. This transverse field is known as the Hall emf. The conditions describing this physical situation are

$$\frac{\partial f_n^0}{\partial x} = \frac{\partial f_n^0}{\partial y} = 0 \qquad \text{and} \qquad I_y = 0. \tag{2}$$

Then Eqs. (1a) and (1b) are

$$\left.\begin{aligned} I_x &= -\frac{4\pi e^2}{3}\left(\mathsf{E}_x\int vl\frac{1}{s^2+1}\frac{\partial f_n^0}{\partial\epsilon}v^2dv - \mathsf{E}_y\int vl\frac{s}{s^2+1}\frac{\partial f_n^0}{\partial\epsilon}v^2dv\right), \\ 0 &= -\frac{4\pi e^2}{3}\left(\mathsf{E}_y\int vl\frac{1}{s^2+1}\frac{\partial f_n^0}{\partial\epsilon}v^2dv + \mathsf{E}_x\int vl\frac{s}{s^2+1}\frac{\partial f_n^0}{\partial\epsilon}v^2dv\right). \end{aligned}\right\} \tag{3}$$

We shall change the variable of integration to ϵ and replace f_n^0 by the Fermi-Dirac function. These equations then become

$$\left.\begin{aligned} I_x &= -\frac{e^2}{3}\sqrt{\frac{2}{m}}(\mathsf{E}_x L_1 - \mathsf{E}_y L_2), \\ 0 &= -\frac{e^2}{3}\sqrt{\frac{2}{m}}(\mathsf{E}_y L_1 + \mathsf{E}_x L_2), \end{aligned}\right\} \tag{4}$$

where

$$L_1 = \frac{C}{V}\int_0^\infty \frac{el}{s^2+1}\frac{\partial f}{\partial \epsilon}d\epsilon, \qquad L_2 = \frac{C}{V}\int_0^\infty \frac{els}{s^2+1}\frac{\partial f}{\partial \epsilon}d\epsilon, \tag{5}$$

and f is the Fermi-Dirac function.

Solving Eqs. (4) for E_x and E_y, we find

$$\left.\begin{aligned} \mathsf{E}_x &= -\frac{3}{e^2}\sqrt{\frac{m}{2}}\frac{L_1}{L_1^2+L_2^2}I_x, \\ \mathsf{E}_y &= \frac{3}{e^2}\sqrt{\frac{m}{2}}\frac{L_2}{L_1^2+L_2^2}I_x. \end{aligned}\right\} \tag{6}$$

The first equation shows that the electrical conductivity in a magnetic field is

$$\sigma(\mathsf{H}_z) = -\frac{e^2}{3}\sqrt{\frac{2}{m}}\frac{L_1^2+L_2^2}{L_1} \tag{7}$$

Now,

$$L_2 = \frac{el(\epsilon_0')}{mv(\epsilon_0')c}\mathsf{H}_z L_1, \tag{8}$$

whence the second of Eqs. (6) may be written

$$\mathsf{E}_y = -\frac{el(\epsilon_0')}{mv(\epsilon_0')c}\frac{\mathsf{H}_z I_x}{\sigma(\mathsf{H}_z)}.$$

The coefficient of $\mathsf{H}_z I_x$ in this is the Hall constant R:

$$R = -\frac{el(\epsilon_0')}{mv(\epsilon_0')c\sigma(\mathsf{H}_z)}, \tag{9}$$

which, when $\sigma(\mathsf{H}_z)$ is equal to the expression for zero magnetic field, reduces to

$$R = -\frac{1}{n_0 ec} \tag{10}$$

[*cf.* Eq. (13) of Sec. 32] where n_0 is the number of electrons per unit volume.[1] It may be shown by tracing through the preceding computation that the negative sign of Eq. (10) arises because the conductivity is related to an electronic current. The sign would be reversed if the carriers were positively charged.

[1] We may see from Eq. (9) that the product $R\sigma$ is the mobility (see footnote 2, p. 68).

TABLE XLIV.—COMPARISON OF OBSERVED HALL CONSTANTS WITH THOSE COMPUTED ON THE BASIS OF THE FREE-ELECTRON THEORY
(In volts/cm-abamp-gauss. Experimental values refer to room temperature.)

Metal	$R \cdot 10^{12}$	
	Observed	Calculated
Cu	− 5.5	− 7.4
Ag	− 8.4	−10.4
Au	− 7.2	−10.5
Li	− 17.0	−13.1
Na	− 25.0	−24.4
Be	+ 24.4	− 2.5
Zn	+ 3.3	− 4.6
Cd	+ 6.0	− 6.5
Al	− 3.0	− 3.4
Fe	+100	
Co	+ 24	
Ni	− 60	
Bi	$H \perp \sim -1{,}000$ $H \parallel \sim +\ 300$	$\sim -\ 4.1$

Experimental and theoretical values[1] of R appear in Table XLIV. The computed numbers are the same order of magnitude as the observed ones for the monovalent metals. The signs are opposite for the divalent metals beryllium, zinc, and cadmium and for iron and cobalt. The magnitude of the observed coefficient is one hundred times larger than

TABLE XLV.—THE ELECTRON MOBILITIES OF SEVERAL METALS AS DERIVED FROM THE PRODUCT OF THE CONDUCTIVITY AND THE HALL CONSTANT
(In cm²/volt-sec)

Metal	μ	Metal	μ
Cu	34.8	Zn	5.8
Ag	56.3	Cd	7.9
Au	29.7	Al	10.1
Li	19.1	Bi	$\perp$ 9.1
Na	48.0		$\parallel$ 2.1
Be	44.4		

[1] See, for example, the compilations of Landolt-Bornstein and the International Critical Tables.

the theoretical one for bismuth. Both these types of anomaly have been explained semiquantitatively by the band theory of solids.

The product of the Hall constant and the electrical conductivity is the electron mobility μ. Values of this quantity for several metals are given in Table XLV. It may be seen that the mobility differs less from metal to metal than do the conductivity and the Hall constant. Using Eq. (13) of Sec. 32 and these mobilities, we find that $l \sim 5 \cdot 10^{-7}$ *cm* for the best conductors.

By a straightforward manipulation with Eqs. (29), Sec. 31, we may reduce the expression (7) for $\sigma(H_z)$ to

$$\sigma(H_z) = \sigma(0)\left(1 - \frac{BH_z^2}{1 + CH_z^2}\right) \tag{11}$$

where

$$B = \frac{\pi^2}{3}\left[\frac{ekTl(\epsilon_0')}{m^2v^2(\epsilon_0')}\right]^2 \tag{12}$$

and

$$C = \sigma^2R^2, \tag{13}$$

in which R is the Hall constant (10).

Equation (11) predicts that σ should decrease quadratically with H_z for weak fields and should asymptotically approach a constant value $\sigma_0\left(1 - \frac{B}{C}\right)$ for large field strengths. The observed values of C usually agree in order of magnitude with those determined from directly measured values of the Hall coefficients. Table XLVI contains values of R for copper, silver, and gold, determined directly as well as from measured

TABLE XLVI.—A COMPARISON OF THE MEASURED HALL CONSTANTS WITH THOSE DETERMINED FROM THE CONDUCTIVITY IN MAGNETIC FIELDS
(In volts/cm-abamp-gauss)

Metal	$R \cdot 10^{12}$	
	Observed	From $\sigma(H_z)$
Cu	−5.5	−7.5
Ag	−8.4	−12.5
Au	−7.2	−17.5
Zn	3.3	∼35
Cd	6.0	∼100
Al	−3.0	13
Sb	∼2,000	∼1,000
Bi	⊥ −1,000	∼ −8,000

values of $\sigma(H_z)$.[1] Figure 12 illustrates the observed dependence of resistance upon H_z for a specimen of copper. The saturation effect predicted by Eq. (11) does not appear for the range of field strengths that has been employed in these metals. Table XLVI shows values of R for zinc, cadmium, aluminum, antimony, and bismuth which were also determined from $\sigma(H_z)$ as well as by direct measurement. The agreement is not so good as for the monovalent metals. The computed values of B are about 10^4 times smaller than the observed ones at ordinary temperatures, a fact showing that a more exact treatment of the problem is necessary.

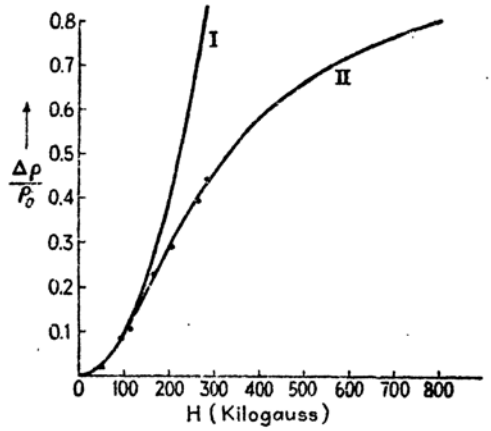

FIG. 12.—Variation of the resistivity of copper in a magnetic field. The ordinate is $\Delta\rho/\rho_0$ where ρ_0 is the resistivity in the absence of a field and $\Delta\rho = \rho(H) - \rho_0$. Curve I is obtained by fitting the points in the quadratic region near the origin with Eq. (11) when C is zero. Curve II is obtained by adjusting both B and C. The value of C obtained in this way agrees closely with the theoretical value, whereas the observed B is about 10^4 times larger than the theoretical value. (*After Sommerfeld and Frank.*)

Qualitative explanations of this discrepancy have been given by Sommerfeld and Frank and by Sommerfeld and Bethe,[2] and more accurate treatments have been developed by Jones and Zener,[3] and by Davis.[4] It may be shown that B depends upon the fluctuations in velocity of an electron in the metal. In the derivation of the relation (12), it is explicitly assumed that the electronic energy is an isotropic function of velocity, so that the only fluctuations that occur result from the fact that the Fermi-Dirac distribution function has a tail of width kT. In an actual metal it turns out, as we shall see in the following chapters, that the energy versus velocity function is not usually isotropic, because of the interactions between electrons and the lattice. The reader is

[1] Measurements by P. Kapitza, *Proc. Roy. Soc.*, **123**, 292, 342 (1929).

[2] SOMMERFELD and FRANK, *op. cit.* SOMMERFELD and BETHE, *op. cit.*

[3] H. JONES and C. ZENER, *Proc. Roy. Soc.*, **145**, 268 (1934).

[4] L. DAVIS, *Phys. Rev.*, **56**, 93 (1939).

referred to the paper by Davis for a more complete discussion of this topic.

B. SEMI-CONDUCTORS

35. A Simple Model of a Semi-conductor.—We shall discuss the theory of semi-conductors on the basis of a simple model[1] that is adequate for understanding most of the characteristic features of these substances. We shall assume that there are n_b bound electronic states per unit volume having energy $-\Delta\epsilon$ and that these levels are completely occupied at absolute-zero temperature by n_e electrons. In addition, we shall assume that above these bound states there is a conduction band of levels that

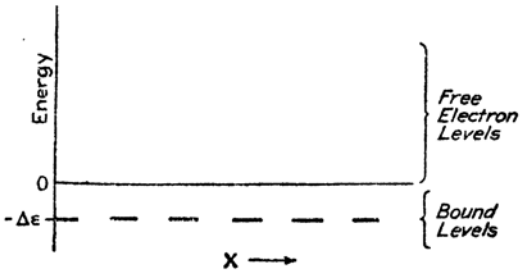

FIG. 13.—Model of a semi-conductor. The ordinate is energy and the abscissa is a symbolical positional coordinate x. The region above zero energy is quasi-continuously dense with free-electron levels. The levels at $-\Delta\epsilon$ correspond to electrons that are bound at particular positions in the lattice.

are completely unoccupied at absolute zero (see Fig. 13). The density $g_c(\epsilon)$ of conduction levels will be taken as

$$g_c(\epsilon)d\epsilon = C\sqrt{\epsilon}d\epsilon \qquad \epsilon > 0 \tag{1}$$

where for a unit volume,

$$C = \frac{4\pi(2m^*)^{\frac{3}{2}}}{h^3} \tag{2}$$

[*cf.* Eq. (9), Sec. 26]. We shall assume that it is permissible to speak of velocity, momentum, etc., for electrons in the conduction band in the conventional way.

Some electrons are thermally excited from the bound states to the conduction band at temperatures above the absolute zero. The number of electrons $n(\epsilon)$ per unit volume having energy ϵ is

$$n(\epsilon) = \frac{g(\epsilon)}{e^{\frac{\epsilon-\epsilon'}{kT}} + 1} \tag{3}$$

[1] A model similar to this was proposed by A. H. Wilson, *Proc. Roy. Soc.*, **133**, 458 (1931); **134**, 277 (1931).

where $g(\epsilon)$ is the density of levels and ϵ' should be determined from the usual condition

$$\int_{-\infty}^{\infty} n(\epsilon)d\epsilon = n_b. \tag{4}$$

The density function $g(\epsilon)$ is given by Eq. (1) when ϵ is greater than zero, and it is zero for all negative values of ϵ except $-\Delta\epsilon$. According to our assumptions about the density of bound states, the density must be so great at $\epsilon = -\Delta\epsilon$ that

$$\int_{-\infty}^{0} g(\epsilon)d\epsilon = n_b. \tag{5}$$

The discontinuous function defined in this way may be approximated as closely as we please by a continuous function. The only one of its properties we shall use is the relation

$$\int_{-\infty}^{0} f(\epsilon)g(\epsilon)d\epsilon = n_b f(-\Delta\epsilon) \tag{6}$$

where $f(\epsilon)$ is any continuous function of ϵ.

We shall divide the integral in Eq. (4) into two integrals, one extending from $-\infty$ to 0 and the other extending from 0 to ∞. The first integral is

$$\int_{-\infty}^{0} g(\epsilon)\frac{1}{e^{\frac{\epsilon-\epsilon'}{kT}}+1}d\epsilon = n_b\frac{1}{e^{\frac{-\Delta\epsilon-\epsilon'}{kT}}+1}, \tag{7}$$

as we may see from Eq. (6). The second integral is

$$C\int_0^{\infty}\frac{\sqrt{\epsilon}d\epsilon}{e^{\frac{\epsilon-\epsilon'}{kT}}+1}. \tag{8}$$

This evidently is equal to the total number of excited electrons, which, to begin with, we shall assume small compared with n_b. It is obvious from the form of the Fermi-Dirac distribution function that the quantity (8) is small only when $-\epsilon' \gg kT$, that is, when $e^{\frac{\epsilon'}{kT}} \ll 1$. We may use this fact to expand (8) in terms of $e^{\frac{\epsilon'}{kT}}$. If we retain only the first term, the result is

$$C\int_0^{\infty}\frac{\sqrt{\epsilon}d\epsilon}{e^{\frac{\epsilon-\epsilon'}{kT}}+1} \cong Ce^{\frac{\epsilon'}{kT}}\int_0^{\infty} e^{-\frac{\epsilon}{kT}}\sqrt{\epsilon}d\epsilon = Ce^{\frac{\epsilon'}{kT}}(kT)^{\frac{3}{2}}\frac{\sqrt{\pi}}{2}. \tag{9}$$

Thus, Eq. (4) is

$$Ce^{\frac{\epsilon'}{kT}}(kT)^{\frac{3}{2}}\frac{\sqrt{\pi}}{2} + n_b\frac{1}{e^{-\frac{\Delta\epsilon+\epsilon'}{kT}}+1} = n_b,$$

or

$$Ce^{\frac{\epsilon'}{kT}}(kT)^{\frac{1}{2}}\frac{\sqrt{\pi}}{2} = n_b\frac{e^{-\frac{\Delta\epsilon+\epsilon'}{kT}}}{e^{-\frac{\Delta\epsilon+\epsilon'}{kT}}+1}. \tag{10}$$

The quantity on the right-hand side of this equation is the number of electrons that have been excited from the bound states. The condition that must be satisfied if this is to be small compared with n_b is that ϵ' should be much larger than $-\Delta\epsilon$, so that the exponential in the denominator may be dropped. We have, as a result,

$$e^{\frac{2\epsilon'}{kT}}C(kT)^{\frac{1}{2}}\frac{\sqrt{\pi}}{2} = n_b e^{-\frac{\Delta\epsilon}{kT}}, \tag{11}$$

or

$$\epsilon' = -\frac{\Delta\epsilon}{2} + \frac{kT}{2}\log\frac{n_b 2}{C(kT)^{\frac{1}{2}}\sqrt{\pi}}. \tag{12}$$

The second quantity on the right is of the order of magnitude kT for ordinary densities of bound electrons. Thus ϵ' is very nearly equal to $-\Delta\epsilon/2$ when $\Delta\epsilon$ is greater than kT (Fig. 14).

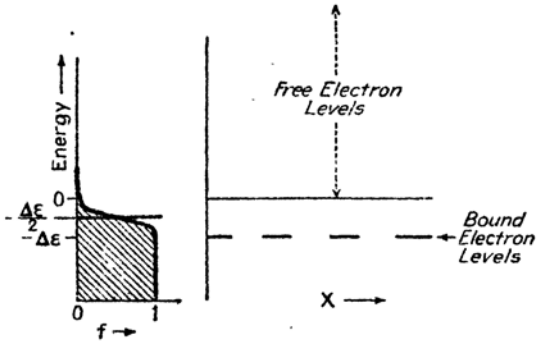

FIG. 14.—The position of the distribution function $f(\epsilon)$ relative to the energy level diagram of Fig. 13. The point $\epsilon = \epsilon_0$ of the distribution function occurs at the point $-\Delta\epsilon/2$ when $\Delta\epsilon >> kT$.

Substituting (12) into Eq. (3) and neglecting small quantities, we find that

$$n(\epsilon) = n_b^{\frac{1}{2}}\left[\frac{2C}{(kT)^{\frac{1}{2}}\sqrt{\pi}}\right]^{\frac{1}{2}} e^{-\frac{\Delta\epsilon}{2kT}}e^{-\frac{\epsilon}{kT}}\sqrt{\epsilon}. \tag{13}$$

This becomes identical with the classical distribution (13), Sec. 26, if we say that the number of free electrons per unit volume n_f is given by the temperature-dependent quantity

$$n_f = n_b^{\frac{1}{2}}\left[\frac{\sqrt{\pi}}{2}(kT)^{\frac{1}{2}}C\right]^{\frac{1}{2}} e^{-\frac{\Delta\epsilon}{2kT}}. \tag{14}$$

If kT is large compared with $\Delta\epsilon$, it is evident that practically all the electrons will have evaporated from the bound levels and that n_f may then be set equal to n_b.

We shall proceed to discuss the conductivity, the thermoelectric effects, and the Hall effect in semi-conductors and shall use Eq. (10) in the same way that we used the Fermi-Dirac distribution in metals. For convenience, we shall write Eq. (13) in the form

$$n(\epsilon) = ae^{-\frac{\epsilon}{kT}}\sqrt{\epsilon} \tag{15}$$

where

$$a = n_f \frac{2}{\sqrt{\pi}} \frac{1}{(kT)^{\frac{3}{2}}}.$$

36. Electrical Conductivity*.—According to Eq. (5), Sec. 32, the electrical conductivity σ is

$$\sigma = \frac{I_x}{E_x} = -\frac{4\pi e^2}{3}\int_0^\infty v^3 l \frac{\partial f_n^0}{\partial \epsilon} dv \tag{1}$$

where $f_n^0 dv_x dv_y dv_z$ is the number of free electrons per unit volume having velocity components in the range from v_x to $v_x + dv_x$, etc. The relation between ϵ and v is

$$\epsilon = \frac{m^* v^2}{2}$$

in our model of a semi-conductor. Hence, f_n^0 is related to $n(\epsilon)$ in Eq. (12) of the preceding section by the equation

$$4\pi f_n^0 v^2 dv = n(\epsilon) d\epsilon. \tag{2}$$

Thus,

$$\begin{aligned} 4\pi f_n^0 &= \frac{n(\epsilon)}{\sqrt{\epsilon}} 2\left(\frac{m^*}{2}\right)^{\frac{3}{2}} \\ &= 2\left(\frac{m^*}{2}\right)^{\frac{3}{2}} a e^{-\frac{\epsilon}{kT}}, \end{aligned} \tag{3}$$

and

$$\sigma_c = -\frac{e^2}{3}\sqrt{\frac{2}{m^*}} a \int_0^\infty l\epsilon^{\frac{1}{2}} \frac{\partial}{\partial \epsilon}(e^{-\frac{\epsilon}{kT}}) d\epsilon. \tag{4}$$

A simpler way of obtaining the same result is to recall that the distribution function in the present case is the limit of the Fermi-Dirac function when ϵ' is negative and is equal to the quantity in Eq. (12) of the preceding section. Then the equation for the conductivity (and for the thermoelectric effects) may be derived by replacing the Fermi-Dirac function f

in all the equations of Secs. 32, 33, and 34 by the following approximate value:

$$f \cong e^{\frac{\epsilon'}{kT}} e^{-\frac{\epsilon}{kT}} = \left[\frac{n_b 2}{C(kT)^{\frac{3}{2}}\sqrt{\pi}}\right]^{\frac{1}{2}} e^{-\frac{\Delta\epsilon}{2kT}} e^{-\frac{\epsilon}{kT}}$$

$$= \frac{a}{C} e^{-\frac{\epsilon}{kT}}. \tag{4a}$$

Equation (9), Sec. 32, then leads to Eq. (4).

We shall assume that l is a constant l_0 for the range of velocities over which the integrand is appreciable. Then,

$$\sigma_e = \frac{e^2}{3}\sqrt{\frac{2}{m^*}} a l_0 kT = \frac{4 n_f l_0 e^2}{3\sqrt{2\pi m^* kT}}. \tag{5}$$

This equation was derived first by Lorentz,[1] who used it for metals under the assumption that n_f is the number of free electrons per unit volume. In passing, let us compare Eq. (5) with σ in Eq. (13), Sec. 32. We shall call the latter σ_q and shall set $n_f = n_0$. The ratio of the two conductivities is

$$\frac{\sigma_q}{\sigma_e} = \frac{3}{4} \frac{\sqrt{2\pi m^* kT}}{m v(\epsilon_0')} \frac{l(\epsilon_0')}{l_0}. \tag{6}$$

If we assume that $l(\epsilon_0')/l_0$ is of the order of unity, this ratio is of the order $\sqrt{kT/\epsilon_0'}$, which is about 10^{-1} at room temperature for ordinary metals. Thus, σ_e agrees with experiment at room temperature only if we assume that either n_f or l_0 is about 10^{-1} times as large as the corresponding quantities that appear in σ_q. Thus, the classical mean free path must be of the order of 10^{-8} cm if n_f is assumed to be equal to n_0. Then, contrary to experiment, the contribution to the electronic heat must be taken as $3R/2$ per mol. On the other hand, if we assume that the classical mean free path is the same as the quantum mean free path, we must assume that

$$n_f \cong n_0 \sqrt{\frac{kT}{\epsilon_0'}}. \tag{7}$$

There is no a priori reason for making this assumption on purely classical grounds. Actually, the use of quantum statistics is equivalent to eliminating all but a fraction of the free electrons, as we have pointed out in previous sections.

[1] See footnote 2, p. 139.

Returning to semi-conductors, for which n_f is expressed by Eq. (14) of the preceding section, we see that

$$\sigma = n_b^{\frac{1}{2}} \frac{4\sqrt{2}}{3} \frac{e^2 l_0}{h^{\frac{1}{2}}} (2\pi m^* kT)^{\frac{1}{4}} e^{-\frac{\Delta\epsilon}{2kT}}. \tag{8}$$

In the range of temperature in which $kT \ll \Delta\epsilon/2$, the coefficient of $e^{-\frac{\Delta\epsilon}{2kT}}$ varies so slowly with temperature, compared with the exponential $e^{-\frac{\Delta\epsilon}{2kT}}$ that (8) agrees within experimental error with the observed law

$$\sigma = Ae^{-\frac{E}{kT}} \tag{9}$$

(*cf.* Sec. 6). We conclude that the observed E and A are related to quantities in Eq. (9) by the equations

$$\begin{aligned} E &= \frac{\Delta\epsilon}{2}, \\ A &= n_b^{\frac{1}{2}} \frac{4\sqrt{2}}{3} \frac{e^2 l_0}{h^{\frac{1}{2}}} (2\pi m^* kT)^{\frac{1}{4}} \\ &= 0.024 l_0 n_b^{\frac{1}{2}} T^{\frac{1}{4}} \text{ ohm}^{-1} \text{ cm}^{-1}. \end{aligned} \tag{10}$$

The numerical value of A has been determined by setting m^* equal to the actual electronic mass.

37. Thermoelectric Effects and the Hall Effect in Semi-conductors.—According to Eq. (6), Sec. 33, the Thomson coefficient σ_T is

$$\sigma_T = -\frac{T}{e} \frac{d}{dT} \left(\frac{K_2}{TK_1} - \frac{\epsilon'}{T} \right) \tag{1}$$

where the quantities K_2 and K_1 must be evaluated by use of the classical value of f in the manner described in Sec. 26. We find

$$\frac{K_2}{TK_1} = k \frac{\int_0^\infty x^2 e^{-x} dx}{\int_0^\infty x e^{-x} dx} = 2k \tag{2}$$

if we assume that l is constant. The expression (2) contributes nothing to (1), since it is independent of temperature. Hence,

$$\sigma_T = \frac{T}{e} \frac{d}{dT} \left(\frac{\epsilon'}{T} \right) = \frac{\Delta\epsilon}{2T} - \frac{3k}{4} \log T. \tag{3}$$

The Peltier coefficient $\pi_{1\to 2}$ and the Seebeck emf F_s are related to σ_T by Eqs. (10) and (12) of Sec. 33, namely:

$$\frac{d}{dT} \left(\frac{\pi_{1\to 2}}{T} \right) = \frac{\sigma_{T,2} - \sigma_{T,1}}{T}, \tag{4}$$

$$F_s = \int_{T'}^{T''} \frac{\pi_{1\to 2}}{T} dT. \tag{5}$$

The value of the Hall constant R is

$$R = -\frac{L_2}{L_1}\frac{1}{\mathsf{H}_z\sigma(\mathsf{H}_z)} \tag{6}$$

[*cf.* Eq. (6), Sec. 34], where the ratio L_2/L_1 now is

$$\frac{L_2}{L_1} = \frac{e\mathsf{H}_z}{mc}l_0\frac{\int_0^\infty \frac{\epsilon}{v}\frac{\partial f}{\partial \epsilon}d\epsilon}{\int_0^\infty \frac{\partial f}{\partial \epsilon}d\epsilon} = \frac{e\mathsf{H}_z}{4mc}l_0\sqrt{\frac{2\pi m}{kT}}. \tag{7}$$

Hence,

$$R\sigma(\mathsf{H}_z) = -\frac{e}{\sqrt{2\pi mkT}}\frac{\pi l_0}{2c}. \tag{8}$$

As in Sec. 32, we shall assume that $\sigma(\mathsf{H})$ is equal to $\sigma(0)$ for ordinary fields. Then σ is given by Eq. (5), Sec. 36, and

$$R = -\frac{3\pi}{8}\frac{1}{n_f ec}. \tag{9}$$

All the electrical quantities in these equations are expressed in electrostatic units.

Let us apply these equations to the case of zinc oxide, which is a typical semi-conductor. Its properties, which have been measured by Fritsch,[1] are strongly dependent upon such factors as thermal treatment and oxygen vapor pressure. This behavior is characteristic of most semiconductors (*cf.* Chap. I). The temperature dependence of the conductivity of a particular specimen is shown in Fig. 15*a*. These results may be fitted by the function

$$\sigma = Ae^{-\frac{E}{kT}}$$

where A is 3.72 ohm^{-1} cm^{-1} and E is 0.013 ev. This specimen was then kept at 900°C for 30 hours in an atmosphere of oxygen at 120 atmospheres. As a result of this treatment, A changed to 2.1 ohm^{-1} cm^{-1}, and E changed to 0.38 ev. Thus, the room-temperature conductivity fell by a factor of about 10^6. The fact that A was nearly the same before and after the heat treatment indicates, according to Eq. (10), that the product $ln_b^{\frac{1}{3}}$ did not change by a large factor during heating. The Hall constant of this specimen was not measured, but it was measured on another specimen before and after heating in the high-pressure furnace. In this

[1] O. FRITSCH, *Ann. Physik*, **22**, 375 (1935).

case, A changed from 5.63 to 0.62 ohm^{-1} cm^{-1}, and E changed from 0.012 to 0.063 ev after 120 hours of heating. It is interesting to note that the change induced in E in this case was not nearly so large as in the case cited above, although the heating in oxygen was extended about four times as long. Apparently erratic results of this kind are characteristic of semi-conductors. The room-temperature Hall constants, before and after heating, were -9×10^{-8} and -380×10^{-8} volt/cm-amp-gauss, respectively. From these results and Eq. (9), we may conclude that n_f dropped from $8 \cdot 10^{17}$ to $2 \cdot 10^{16}$ electrons per cubic centimeter as a

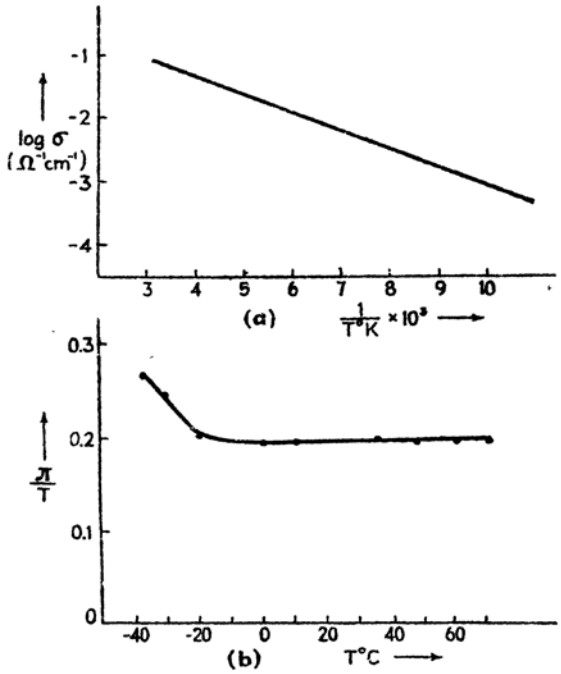

FIG. 15.—*a*, the log σ versus $1/T$ plot for a specimen of zinc oxide; *b*, the Seebeck emf per degree for the same specimen. The ordinates are in units of millivolt/deg.

result of heating. The product $R\sigma$ at room temperature, was $-30 \cdot 10^{-8}$ gauss^{-1} before heating and remained practically unchanged. This, according to Eq. (8), implies that l_0 was about $1.7 \cdot 10^{-7}$ cm.

Figure 15*b* shows the Seebeck emf per degree as a function of temperature for a copper-zinc oxide system. These measurements were made on the second specimen discussed in the previous paragraph after it was heated. The Seebeck emf per degree should be equal to the derivative of (5) with respect to T''; that is, it should be equal tc

$$\frac{\pi_{1\to 2}}{T}. \tag{10}$$

This expression may be related to the Thomson coefficients of copper and of zinc oxide by means of Eq. (4). We shall assume that the coefficient for copper is negligible compared with that for zinc oxide, basing this assumption on the fact that the Seebeck emf for metals is about 10^{-3} times as large as that for semi-conductors. The quantity (10) is then

$$\frac{\pi_{1\to 2}}{T} \cong \frac{\epsilon'}{Te} \tag{11}$$

because of (3) and (4). Substituting the value of ϵ' derived in Sec. 35 [*cf.* Eqs. (12) and (14)], we find

$$\frac{\pi_{1\to 2}}{T} \cong -\frac{1}{e}\left(\frac{\Delta\epsilon}{T} + k \log \frac{n_f}{n_b}\right). \tag{12}$$

Since the measured value of $\Delta\epsilon$ is $2 \cdot 0.063$ ev, Eq. (12) predicts an effect of the order of magnitude $0.4 \cdot 10^{-3}$ volt per degree, which agrees roughly with the observed effect. The precise variation, however, does not seem to lie within the descriptive power of the simple theory.

This manner of correlating the properties of a semi-conductor may be applied to a large number of other cases. It should be emphasized that the simple theory we have used applies only to those substances for which the Hall effect is negative. There are, however, a large number of substances, such as copper iodide, for which the Hall constant has a positive sign, as though positive charges carry the current. We shall return to a discussion of these substances after developing the band theory of solids.

CHAPTER V

QUANTUM MECHANICAL FOUNDATION

In Part A of this chapter we shall consider the principles and theorems of quantum mechanics that have particular use in the theory of the solid state. Although some of this material may be well known to the reader,[1] we shall present it here in order to place it in a form that is consistent with the treatment of later chapters. In Part B, we shall discuss the theory of radiation.

PART A

38. Elementary Postulates of the Theory.—The development of the nonrelativistic form of quantum mechanics, with which we shall be solely concerned, brought with it a revision of the logical and mathematical discipline of mechanics. This revision was necessary in order to include principles that are applicable in a larger domain of the physical world than that in which the classical laws are valid. There is a correspondence between the classical and the quantum mechanical laws, for the domain of the former is contained in the domain of the latter. Thus, there is a quantity in quantum mechanics corresponding to each quantity in classical mechanics, and the quantum laws reduce to the classical ones when Planck's constant may be regarded as a very small quantity.

One of the primary features of quantum mechanics is the introduction of a state function which is said to describe a given dynamical system completely when the system is in a given state of motion. This function, as we shall use it, is an ordinary function of Cartesian coordinates. All available information may be derived from the state function by the proper use of certain operators that correspond, individually, to measurable quantities, such as position coordinates, momentum, energy, etc. Neither state functions nor the operators have immediate physical significance; only certain quantities derived by proper juxtaposition of the two are measurable numbers.

[1] General references: P. A. M. DIRAC, *The Principles of Quantum Mechanics* (Oxford University Press, New York, 1935). S. DUSHMAN, *Elements of Quantum Mechanics* (John Wiley & Sons, Inc., New York, 1938). E. C. KEMBLE, *The Fundamental Principles of Quantum Mechanics* (McGraw-Hill Book Company, Inc., New York, 1937). V. ROJANSKY, *Introductory Quantum Mechanics* (Prentice-Hall, Inc., New York, 1938).

The most convenient definition of an operator, from our standpoint, is the following: An operator is a quantity symbolizing a process in which a given function is changed into another function. For example, the process of taking the square root of a function defines an operator. If we designate this operator by $\sqrt{\ }$ and the operating process by a dot, we have

$$\sqrt{\ } \cdot f = \sqrt{f};$$

that is, $\sqrt{\ }$ operating on f gives $\sqrt{f}$. Similarly, we may regard the ordinary differential symbol $\partial/\partial x$ as an operator for which

$$\frac{\partial}{\partial x} \cdot f(x) = \frac{\partial f}{\partial x}.$$

Likewise, the multiplication of the functions $f(x)$ and $v(x)$ defines an operator since the product $v(x)f(x)$ is a new function of x.

Complete description in quantum theory does not imply precise knowledge of all measurable quantities at all instants of time as it does in classical mechanics. Quantum mechanics is primarily a statistical theory; its results tell us the mean or expectation values of measurements. Thus repetition of the procedure for making a precise measurement of a given quantity usually should not lead to a repetition of results even when the system on which measurements are made is in the same state at the beginning of each measurement. This principle is to be contrasted with the principles of classical theory, according to which one should expect precisely repeatable results under identical experimental conditions.

The statistical formulation of quantum theory is believed to be ultimate, in contrast with the formulation of classical statistical mechanics in which probability is introduced only as a convenient tool. Thus, it is believed that the limitation of description contained in quantum theory can be verified by direct experiment. The electron-diffraction experiments[1] of Davisson and Germer, Thomson, and Rupp and the molecular-beam experiments of Rabi[2] and his coworkers have gone a long way toward providing this verification. Even without this direct experimental check of the uncertainty relations, however, there is overwhelming experimental evidence of other kinds to justify the use of quantum mechanics for the types of problem that are considered in this book.

[1] C. DAVISSON and L. H. GERMER, *Phys. Rev.*, **30**, 705 (1927). G. P. THOMSON, *Proc. Roy. Soc.*, **117**, 600 (1928); **119**, 651 (1928). E. RUPP, *Ann. Physik*, **85**, 981 (1928).

[2] I. RABI, *Phys. Rev.*, **49**, 324 *ff.* (1936).

Although experimental results usually should not be precisely repeatable, there is one case in which duplicate observations should give identical results. Suppose that α is the operator corresponding to a given observable quantity and that f is the state function of a system for a particular state of motion. Precisely the same values of the measured quantity associated with α should be obtained when the system is in the state f if

$$\alpha \cdot f = af \tag{1}$$

where a is constant. Moreover, the number a should be the result of each measurement. A function f that satisfies an equation of this type is by definition an *eigenfunction* of the operator α. a is called the eigenvalue of f.

From a practical standpoint, the problems in which we are primarily interested are to determine (1) the possible forms that may be given to the state function and to the dynamical operators and (2) the dynamical laws of the theory. The solutions may be placed in many possible forms, just as in classical theory. A serviceable form for our purposes is the one based on Schrödinger's scheme. It may be summarized in the following way:

a. The operators that correspond to the Cartesian coordinates x_i, y_i, z_i of the ith particle of a system and to the time variable t are taken to be the variables themselves. The state function f is then chosen as a function of these variables. Then,

$$f = f(x_1, y_1, z_1, \cdots, x_n, y_n, z_n, t).$$

b. The operators of the variables that are conjugate to these variables in the classical sense, namely, the corresponding momenta p_{x_1}, p_{y_1}, p_{z_1}, $\cdots$, p_{z_n}, and the negative of the Hamiltonian function H are taken in the form

$$\frac{\hbar}{i}\frac{\partial}{\partial x_1}, \quad \frac{\hbar}{i}\frac{\partial}{\partial y_1}, \cdots, \frac{\hbar}{i}\frac{\partial}{\partial z_n}, \quad \frac{\hbar}{i}\frac{\partial}{\partial t}, \tag{2}$$

respectively. In classical dynamics, $-H$ is the conjugate of t only in the special case in which it is independent of time. The assignment of the operator $-\frac{\hbar}{i}\frac{\partial}{\partial t}$ to H is assumed to hold, however, even when H is dependent on time.

c. The operator corresponding to any classical dynamical variable that is a function of the x, the p, and t may be obtained by replacing the p by the differential operator introduced in part *b*. In particular, the operator corresponding to the Hamiltonian function, which generally has the form,

$$H(x_1, \cdots, z_n, p_{x_1}, \cdots, p_{z_n}, t)$$

is

$$H\left(x_1, \cdots, z_n, \frac{\hbar}{i}\frac{\partial}{\partial x_1}, \cdots, \frac{\hbar}{i}\frac{\partial}{\partial z_n}, t\right). \tag{3}$$

This operator must be equal to $-\frac{\hbar}{i}\frac{\partial}{\partial t}$, if the identification presented in part *b* is also correct. Since this relation is not an identity, it is necessary to assume that the theory is concerned only with those state functions that satisfy the relation

$$H\left(x_1, \cdots, z_n, \frac{\hbar}{i}\frac{\partial}{\partial x_1}, \cdots, \frac{\hbar}{i}\frac{\partial}{\partial z_n}, t\right)f = -\frac{\hbar}{i}\frac{\partial f}{\partial t}. \tag{4}$$

This is the fundamental dynamical equation of quantum mechanics, which is used to determine the state function. We shall refer to it as the first Schrödinger equation.

d. The states that are eigenfunctions of the energy operator satisfy the equation

$$Hf = -\frac{\hbar}{i}\frac{\partial f}{\partial t} = Ef. \tag{5}$$

Hence, they must have the form

$$f = \Psi(x_1, \cdots, z_n)e^{-i\frac{E}{\hbar}t} \tag{6}$$

where Ψ satisfies the equation

$$H\Psi = E\Psi. \tag{7}$$

This is the second Schrödinger equation.

e. It was mentioned in connection with Eq. (1) that the observed value of α is always a if a system is in a state f that satisfies the equation

$$\alpha f = af.$$

If the state function is not an eigenfunction of α, the mean value $\bar{a}$ of the measured value is

$$\bar{a} = \frac{\int f^*\alpha \cdot f d\tau}{\int f^* f d\tau} \tag{8}$$

where the integration is to be extended over all values of the x and t; f^* is the complex conjugate of f; and $d\tau$ is the product of dt and the volume element in $3n$-dimensional configuration space. Whenever $\int f^* f d\tau$ is infinite, $\bar{a}$ is the limit to which the ratio (8) approaches as the volume of integration is gradually increased to include the entire space. Since

all measured quantities are real, the allowed operators α must satisfy the relation

$$\int f^*(\alpha \cdot f)d\tau = \int f(\alpha \cdot f)^* d\tau. \tag{9}$$

Such operators are called Hermitian.

f. One may legitimately ask for the operator that corresponds to the measurement of whether or not the system has the coordinates $x_1', \cdots, z_n', t'$. It is difficult to formulate this operator in a mathematically rigorous fashion in the Schrödinger theory. This difficulty may be avoided by defining the operator to be one for which the integral (8) has the value

$$f^*f(x_1', \cdots, z_n', t'). \tag{10}$$

Thus, $f^*f(x_1', \cdots, z_n', t')$ is interpreted as the probability that the system has the coordinates $x_1', \cdots, z_n', t'$. If f is an eigenfunction of the Hamiltonian operator, it has the form (6), and the quantity $f^*f = \Psi^*\Psi$ is independent of time. States of this type are said to be stationary.

g. It is natural to expect that f^*f should be an integrable function if f^*f is interpreted as the relative probability that the system should occupy the coordinates $x_1', \cdots, z_n', t'$. This is a general restriction on the state function. Ordinarily, this condition is fulfilled by demanding that f should be finite everywhere; the more accurate condition, however, is that the integral of f^*f over any finite volume should exist and be finite.

It may easily be shown[1] that classical mechanics and quantum mechanics lead to identical results in the ordinary large-scale domain in which classical mechanics is ordinarily applied.

39. Auxiliary Theorems.—There are several additional theorems that are frequently used in conjunction with the preceding principles because they are useful in applying the theory to specific problems. They may be listed in the following way.

a. The purely spatial part of an eigenfunction (6) of a time-independent Hamiltonian operator may be used without the time-dependent factor $e^{-\frac{iEt}{\hbar}}$ when the mean value of any time-independent operator is computed, since these factors cancel out of the integrands in Eq. (8). Since the eigenfunctions of other operators usually do not satisfy the time-dependent Schrödinger equation (4) and hence cannot be written in the form (6), their purely spatial eigenfunctions usually cannot be used in the same way. We shall mention two useful theorems concerning the spatial eigenfunctions of Hermitian operators.

1. The space eigenfunctions $\psi_1, \psi_2, \cdots$ of any time-independent Hermitian operator form a complete orthogonal set; therefore, any space

[1] See KEMBLE, *op. cit.*, p. 49.

function φ, such as an eigenfunction of a Hamiltonian operator, may be expanded in terms of them in a Fourier-series fashion. Thus,

$$\varphi = \sum_i a_i \psi_i \tag{1}$$

where

$$a_i = \frac{\int \varphi \psi_i^* d\tau(x_1, \cdots, z_n)}{\int |\psi_i|^2 d\tau}. \tag{2}$$

2. The condition that must be satisfied in order that two operators α and β may have all eigenfunctions in common is that they should commute, that is, that they must satisfy the condition

$$(\alpha\beta - \beta\alpha)\psi = 0 \tag{3}$$

for arbitrary ψ. This theorem is particularly useful when we have operators that commute with the Hamiltonian operator of a system, for then we may choose the stationary states to be eigenfunctions of all these operators.

b. Normalizing Conditions.—Since it is possible to interpret $|f(x_1, \cdots, z_n, t)|^2$ as the relative probability that the system has the coordinates $x_1, \cdots, z_n$, at time t, it is reasonable to ask that f should satisfy the equation

$$\int |f(x_1, \cdots, z_n, t)|^2 d\tau(x_1, \cdots, z_n) = 1 \tag{4}$$

at each instant of time, where the integration extends over all space. This is known as the normalizing condition. Since the relation

$$\frac{d}{dt}\int |f|^2 d\tau(x_1, \cdots, z_n) = 0 \tag{5}$$

may be proved by use of the Schrödinger equation and the Hermitian condition on H, a state function that is normalized at one instant of time remains normalized at all later times. As a rule, normalized functions are used because they give absolute rather than relative probabilities.

c. The Variational Theorem.—One of the most powerful tools for obtaining solutions of the Schrödinger equation by means of approximate methods is furnished by the variational theorem of quantum mechanics. This theorem states that functions $\Psi(x_1, \cdots, z_n)$, for which the variation of the mean value

$$\bar{a} = \frac{\int \Psi^* \alpha \Psi d\tau(x_1, \cdots, z_n)}{\int \Psi^* \Psi d\tau} \tag{6}$$

is zero, satisfy the relation

$$\alpha\Psi = a\Psi.$$

The converse of this is also true.

We shall prove this theorem for the case in which α is a time-independent Hermitian operator and in which integration extends over all space.

We have, as the condition on $\bar{a}$,

$$\delta\bar{a} = \delta\frac{\int\Psi^*\alpha\Psi d\tau}{\int\Psi^*\Psi d\tau} = 0 \tag{8}$$

where $\delta\bar{a}$ indicates the variation in $\bar{a}$ that is to be associated with a variation $\delta\Psi$ in Ψ. If we set

$$\begin{aligned} A &= \int\Psi^*\alpha\Psi d\tau, \\ B &= \int\Psi^*\Psi d\tau, \end{aligned} \tag{9}$$

$\delta\bar{a}$ becomes

$$\delta\bar{a} = \delta\left(\frac{A}{B}\right) = \frac{B\delta A - A\delta B}{B^2} = 0. \tag{10}$$

Hence, we must have

$$\frac{A}{B}\delta B - \delta A = 0,$$

or

$$\bar{a}\delta B - \delta A = 0, \tag{11}$$

since

$$\bar{a} = \frac{A}{B}.$$

Now,

$$\begin{aligned} \delta A &= \int\delta\Psi^*\alpha\Psi d\tau + \int\Psi^*\alpha\delta\Psi d\tau, \\ \delta B &= \int\delta\Psi^*\Psi d\tau + \int\Psi^*\delta\Psi d\tau, \end{aligned}$$

whence (11) may be written in the form

$$\int\delta\Psi^*(\bar{a} - \alpha)\Psi d\tau + \int\Psi^*(\bar{a} - \alpha)\delta\Psi d\tau = 0$$

or in the form

$$\int\delta\Psi^*(\bar{a} - \alpha)\Psi d\tau + \int\delta\Psi(\bar{a} - \alpha)\Psi^* d\tau = 0, \tag{12}$$

since α is Hermitian. If we express Ψ in terms of its real and imaginary parts, Ψ_r and Ψ_i, respectively, (11) becomes

$$\int\delta\Psi_r(\bar{a} - \alpha)\Psi_r d\tau + \int\delta\Psi_i(\bar{a} - \alpha)\Psi_i d\tau = 0.$$

The necessary and sufficient conditions for this equality are that

$$(\bar{a} - \alpha)\Psi_r = 0,$$
$$(\bar{a} - \alpha)\Psi_i = 0,$$

since both $\delta\Psi_r$ and $\delta\Psi_i$ are arbitrary. Hence, the necessary and sufficient condition for the validity of Eq. (8) is that Ψ should satisfy the equation

$$\alpha\Psi = \bar{a}\Psi. \tag{13}$$

It is not possible without further investigation to say whether a particular Ψ satisfying (13) gives $\bar{a}$ a maximum value, a minimum value, or just an inflection point.

If we apply this theorem to the Hamiltonian operator H, there generally is a lowest value of the mean value corresponding to the stationary state of lowest energy. States of higher energy are invariably orthogonal to this, as we mentioned in part *a*. Hence, the state Ψ_2, just above the lowest Ψ_1, may be specified by the two conditions

$$\delta\frac{\int\Psi_2^* H\Psi_2 d\tau}{\int\Psi_2^*\Psi_2 d\tau} = 0, \tag{14}$$
$$\int\Psi_2^*\Psi_1 d\tau = 0. \tag{15}$$

The necessary and sufficient condition for these equations is that Ψ_2 should satisfy the relation

$$H\Psi_2 = (E - \lambda)\Psi_2 = E'\Psi_2$$

where the Lagrangian parameter λ may be determined by the condition (15). Higher discrete states may be defined in a similar way by the condition (14) and by additional conditions of the type (15) which express the fact that the higher state is orthogonal to all lower ones.

The variational theorem shows that the accuracy of the mean value of H for a given approximate function f is usually greater than the accuracy of the mean value of other quantities. This fact may be shown directly as follows. Suppose that f is expressed in the form

$$f = \Psi + \alpha\Phi \tag{16}$$

where Ψ is the exact eigenfunction, which f represents approximately, and $\alpha\Phi$ is the part of f orthogonal to Ψ. We shall assume that Ψ and Φ are normalized and that α is a small number. The mean value of H for the function (16) is

$$\frac{\int(\Psi^* + \alpha\Phi^*)H(\Psi + \alpha\Phi)d\tau}{(1 + \alpha^2)} = \frac{(E + \alpha^2\int\Phi^* H\Phi d\tau)}{(1 + \alpha^2)}$$

where E is the eigenvalue of Ψ. Thus, the fractional accuracy of Ψ is the square root of that of E. Since the mean values of other operators

involve terms in the first power of α, their fractional accuracy is also of the order of the square root of the accuracy of the energy.

40. Electron Spin*.[1]—A significant feature of the preceding formulation of quantum mechanics is that it contains an operator for each classical variable. It has been found necessary to introduce other operators that do not correspond to classically measurable quantities, in order to explain certain experimental observations. Most important among these operators are those associated with electron spin. Historically, they find their origin in an attempt of Goudsmit and Uhlenbeck to explain certain features of atomic spectra that had not been previously interpreted. These workers were led to assume that an electron possesses a spin about an axis passing through its center and that the total spin angular momentum is equal to $\hbar/2$. This condition on the angular momentum may be expressed in the form

$$|\boldsymbol{\sigma}| = \sqrt{\sigma_x^2 + \sigma_y^2 + \sigma_z^2} = \frac{\hbar}{2}$$

where σ_x, σ_y, and σ_z are, respectively, the x, y, and z components of spin angular momentum. In addition, they found it necessary to assume that the magnetic moment $\boldsymbol{\mu}$, associated with electron spin, is related to the mechanical moment $\boldsymbol{\sigma}$, by means of the equation

$$\boldsymbol{\mu} = -\frac{e}{mc}\boldsymbol{\sigma}. \tag{2}$$

In contrast with this, the relationship between the orbital angular momentum $\mathbf{L}$ (that is, the angular momentum of an electron moving about an axis that does not pass through it) and the orbital magnetic moment $\mathbf{M}_0$ is

$$\mathbf{M}_0 = -\frac{e}{2mc}\mathbf{L}; \tag{5}$$

that is, there is an additional factor of 2 in the right-hand side of Eq. (2). There have been several attempts, based on classical electromagnetic theory, to prove that the mechanical and magnetic moments of a spinning spherical charge distribution actually do satisfy the relation (2). At present, this work is generally regarded as inapplicable, aside from being inconclusive, for the phenomenon of electron spin is viewed as lying outside the domain in which classical concepts have meaning.

The most complete theory of spin yet devised was discovered by Dirac[2] in a search for a form of quantum mechanics that satisfies the principles of relativity. Equation (2) arises as a by-product of other assumptions in Dirac's theory. We shall not be concerned with the

[1] See *ibid.*, p. 510.

[2] See ROJANSKY, *op. cit.*

details of this theory, for we are interested only in the nonrelativistic approximation. In this case, spin may be handled by a scheme developed by Pauli to which Dirac's treatment reduces when the velocities of electrons are small compared with the velocity of light.

The significance of spin in Pauli's theory is connected with the statement that electrons are particles for which the complete state function has two components instead of one. The two components are not scalar functions in the sense that they are independent of the choice of coordinate system in three-dimensional space, for they transform between one another in a complex way when coordinate axes are transformed. Thus, it is customary to introduce the concept of a two-dimensional spin space in which the two-component functions of the state function are represented as orthogonal components of a spinor. A transformation of coordinate axes in ordinary space induces a corresponding transformation in spin space. We shall not consider the details of spinor transformation theory because we shall have no explicit use for the transformation equations.[1] It should be mentioned, however, that the transformation characteristics of two-dimensional spinors are considerably different from those of two-dimensional vectors.

The introduction of spin in the quantum theory of electrons is analogous to the introduction of the concept of polarization into the theory of light. Suppose that we were acquainted with none of the polarization phenomena of optics. Then it would be possible to describe many optical experiments, such as those of interference or of energy transport, by assuming that the amplitude of a light wave is a scalar quantity, just as the amplitude of sound. As soon as experiments on polarization are performed, however, we are compelled to say that there is a vector character associated with the amplitude of a light wave. This vector character may be described by taking components of the amplitude in two orthogonal directions of polarization. The analogy between the electron and the light wave does not hold in a quantitative way, however, for the two independent directions of electron polarization are separated by 180 degrees rather than by 90 degrees, as they are for light. This fact marks the difference between vector character and spinor character.

We shall represent the coordinate variables in spin space by ζ_1 and ζ_2, respectively. A unit spinor in the direction of the ζ_1 axis has components $\zeta_1 = 1$, $\zeta_2 = 0$ and may be represented by the column matrix

$$\mathfrak{n}(1,0) = \begin{pmatrix} 1 \\ 0 \end{pmatrix}; \tag{4}$$

[1] See, for example, E. P. WIGNER, *Gruppentheorie* (Vieweg, Braunschweig, Germany, 1931).

a unit spinor along the ζ_2 axis may be represented by

$$\mathfrak{n}(0,1) = \begin{pmatrix} 0 \\ 1 \end{pmatrix}. \tag{5}$$

Pauli assumed that the components of spin angular momentum represented in spin space by the matrices

$$\sigma_x = \frac{\hbar}{2}\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \qquad \sigma_y = \frac{\hbar}{2}\begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \qquad \sigma_z = \frac{\hbar}{2}\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \tag{6}$$

The spinor $\mathfrak{n}(1,0)$ of Eq. (4) satisfies the relation

$$\sigma_z \cdot \mathfrak{n}(1,0) = \frac{\hbar}{2}\mathfrak{n}(1,0), \tag{7}$$

and $\mathfrak{n}(0,1)$ satisfies

$$\sigma_z \cdot \mathfrak{n}(0,1) = -\frac{\hbar}{2}\mathfrak{n}(0,1). \tag{8}$$

Hence, these two spinors are said to correspond respectively to the precise values $+\frac{1}{2}\hbar$ and $-\frac{1}{2}\hbar$ of the z component of spin angular momentum. A direct computation of the matrix

$$\sigma^2 = \sigma_x^2 + \sigma_y^2 + \sigma_z^2, \tag{9}$$

which corresponds to the total spin angular momentum, shows that

$$\sigma^2 = \tfrac{3}{4}\hbar^2\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}. \tag{10}$$

Thus any spinor $\mathfrak{n}(\zeta_1, \zeta_2)$ satisfies the equation

$$\sigma^2 \cdot \mathfrak{n}(\zeta_1, \zeta_2) = \tfrac{3}{4}\hbar^2 \cdot \mathfrak{n}(\zeta_1, \zeta_2), \tag{11}$$

and the precise value of the square of the total spin angular momentum is $3\hbar^2/4$.

In the coordinate system in spin space for which the matrices of σ_x, σ_y, and σ_z have the form (6), the two components of the state function may be labeled by a variable ζ_z which takes two values, namely, $+1$ for the coordinate going with the diagonal element $\hbar/2$ of σ_z and -1 for the coordinate going with the value $-\hbar/2$. In other words, we may label the axes in spin space with the eigenvalues of $2\sigma_z/\hbar$, instead of by the indices 1 and 2. The state function then has the form $f(x,y,z,\zeta_z)$, it being understood that $f(x,y,z,1)$ and $f(x,y,z,-1)$ are the components of a spinor

$$\begin{pmatrix} f(x,y,z,1) \\ f(x,y,z,-1) \end{pmatrix}. \tag{12}$$

All state functions (12) are eigenfunctions of the total square spin angular momentum (10). The reader may readily verify that the state functions

$$\begin{pmatrix} f(x,y,z) \\ 0 \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} 0 \\ f(x,y,z) \end{pmatrix}$$

are eigenfunctions of σ_z and that

$$\begin{pmatrix} f(x,y,z) \\ \pm f(x,y,z) \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} f(x,y,z) \\ \pm if(x,y,z) \end{pmatrix}$$

are eigenfunctions of σ_x and σ_y, respectively.

Classically, the components of angular momentum of a particle, relative to the origin of coordinates, are given by

$$\left.\begin{aligned} m_x &= yp_z - zp_y, \\ m_y &= zp_x - xp_z, \\ m_z &= xp_y - yp_x, \end{aligned}\right\} \tag{13}$$

where x, y, z are the spatial coordinates of the particle and p_x, p_y, p_z are the components of angular momentum. The quantum operators that correspond to (13) are

$$\left.\begin{aligned} m_x &= \frac{\hbar}{i}\left(y\frac{\partial}{\partial z} - z\frac{\partial}{\partial y}\right), \\ m_y &= \frac{\hbar}{i}\left(z\frac{\partial}{\partial x} - x\frac{\partial}{\partial z}\right), \\ m_z &= \frac{\hbar}{i}\left(x\frac{\partial}{\partial y} - y\frac{\partial}{\partial x}\right). \end{aligned}\right\} \tag{14}$$

It is easily shown that these obey the commutation rules

$$\left.\begin{aligned} m_y m_z - m_z m_y &= \hbar i m_x, \\ m_z m_x - m_x m_z &= \hbar i m_y, \\ m_x m_y - m_y m_x &= \hbar i m_z. \end{aligned}\right\} \tag{15}$$

Since the spin matrices (6) obey exactly the same commutation rules, they may be viewed as quantum mechanical angular-momentum operators even though this angular momentum cannot be measured directly, even in an idealized experiment.

Along with the spin operators, we may introduce the following operators:

$$\left.\begin{aligned} \mu_x &= \frac{e}{mc}\sigma_x, \\ \mu_y &= \frac{e}{mc}\sigma_y, \\ \mu_z &= \frac{e}{mc}\sigma_z, \end{aligned}\right\} \tag{16}$$

which we shall call the components of spin magnetic moment. The choice of coefficient in these relations is justified both by Dirac's theory and by experiment, as we have said previously.

When the mechanical system contains n electrons instead of 1, the state function must be viewed as a spinor in a 2^n-dimensional spin space. We may choose the coordinate system of this in such a way that the components of the state function are labeled by n variables $\zeta_{z_1}, \zeta_{z_2}, \cdots, \zeta_{z_n}$, each of which takes the two values ± 1; that is, we may write the state function in the form

$$f(x_1, \cdots, z_n, \zeta_{z_1}, \cdots, \zeta_{z_n}). \tag{17}$$

The function of $x_1, \cdots, z_n$ associated with each of the 2^n possible values of the ζ_z is a component of a 2^n-dimensional spinor. It is implied that the operators $\sigma_{z_i} (i = 1, \cdots, n)$, corresponding to the z components of spin of the n electrons, are 2^n-dimensional diagonal matrices in this coordinate system and that the diagonal element of $2\sigma_{z_i}/\hbar$ is $+1$ or -1, respectively, for the coordinate axes for which ζ_{z_i} is $+1$ or -1. Similarly, the matrices σ_{x_i} and σ_{y_i} are represented by 2^n-dimensional matrices that have the form (6) in the two-dimensional subspace that is associated with the two axes labeled by $\zeta'_1, \zeta'_2, \cdots, \zeta'_{i-1}, \pm 1, \zeta'_{i+1}, \cdots, \zeta'_n$. All other components are zero. It is easy to show that the σ satisfy the commutation rules

$$\left.\begin{aligned} \sigma_{y_i}\sigma_{z_j} - \sigma_{z_j}\sigma_{y_i} &= \delta_{ij}\hbar i\sigma_{x_i}, \\ \sigma_{z_i}\sigma_{x_j} - \sigma_{x_j}\sigma_{z_i} &= \delta_{ij}\hbar i\sigma_{y_i}, \\ \sigma_{x_i}\sigma_{y_j} - \sigma_{y_j}\sigma_{x_i} &= \delta_{ij}\hbar i\sigma_{z_i}. \end{aligned}\right\} \tag{18}$$

The matrices $\sigma_i^2 = \sigma_{x_i}^2 + \sigma_{y_i}^2 + \sigma_{z_i}^2$ are diagonal matrices, all diagonal elements of which are equal to $3\hbar^2/4$ [*cf.* Eq. (10)].

The three matrices, defined by the equations

$$\left.\begin{aligned} \Sigma_x &= \sum_{i=1}^{n} \sigma_{x_i} \\ \Sigma_y &= \sum_{i=1}^{n} \sigma_{y_i} \\ \Sigma_z &= \sum_{i=1}^{n} \sigma_{z_i} \end{aligned}\right\} \tag{19}$$

are called the operators of the components of the total spin angular momentum, and

$$\Sigma^2 = \Sigma_x^2 + \Sigma_y^2 + \Sigma_z^2 \tag{20}$$

is called the square of the total spin angular momentum. Σ_x, Σ_y, Σ_z

commute with $\mathbf{\Sigma}^2$ and obviously satisfy the commutation rules

$$\left.\begin{aligned} \Sigma_y\Sigma_z - \Sigma_z\Sigma_y &= i\hbar\Sigma_x, \\ \Sigma_z\Sigma_x - \Sigma_x\Sigma_z &= i\hbar\Sigma_y, \\ \Sigma_x\Sigma_y - \Sigma_y\Sigma_x &= i\hbar\Sigma_z, \end{aligned}\right\} \tag{21}$$

because of Eqs. (18). The one-electron operators σ_{xi}, σ_{yi}, and σ_{zi} commute with Σ_x, Σ_y, Σ_z, respectively; but the pairs σ_{xi} and Σ_y do not commute with one another. Moreover, the one-electron operators do not commute with $\mathbf{\Sigma}^2$ when there are two or more electrons in the system. Hence, the one-electron operators will not be in diagonal form if the coordinate axes are selected so that Σ_z and $\mathbf{\Sigma}^2$ are diagonal.

If the Hamiltonian operator does not contain any spin terms, as happens in many actual cases, it commutes with all spin operators. Then the stationary states of the system may be chosen to be eigenfunctions of any set of commuting spin operators, such as the set $\Sigma_z, \sigma_{z1}, \cdots, \sigma_{zn}$ or the set $\mathbf{\Sigma}^2, \Sigma_z$. The second set is particularly useful when the Pauli principle, which we shall describe in the next section, is properly taken into account.

41. The Pauli Principle and Related Restrictions.—The Hamiltonian operator of any system that contains at least two particles of a given kind remains invariant in form if the coordinates of like particles are permuted among themselves. It may be shown that for this reason the set of stationary states of the system that are associated with a given eigenvalue of the Hamiltonian transform among themselves in one of several different ways when the coordinate variables are permuted. The theory of groups of transformations is particularly concerned with this property of eigenfunctions. This topic need not interest us at present, however, because of the Pauli exclusion principle which states that the physically permissible solutions of Schrödinger's equations behave in definite ways when the coordinates of particles are permuted. In particular, the Pauli principle requires that the state function be antisymmetric under electron permutations; that is, the state function must transform into its negative under odd permutations of electrons and into itself under even ones. It is also true that the state function must be symmetric under the permutation of the coordinates of light quanta; that is, it must transform into itself under both even and odd permutations in this case.

The fundamental reason for these requirements is not completely understood at present. It seems reasonable, however, to expect that the exclusion principle will appear as the natural consequence of some general invariance principle, possibly unformulated as yet, just as the concept of spin arises out of the requirement of relativistic invariance in Dirac's theory.

Statistical theories, corresponding to classical statistical mechanics, that take the exclusion principle into account have been developed[1] by Fermi and Dirac for the antisymmetric case and by Bose and Einstein for the symmetric one.

It seems to be a general rule in nature that all elementary particles, except photons, obey Fermi-Dirac statistics. The statistical behavior of nuclei, as is determined from the analysis of band spectra, is explained completely by assuming that nuclei are composed of protons and neutrons. In order to remove the single exception, photons, it has been postulated that light quanta are composed of two elementary particles which are not observed separately in standard optical experiments. The evidence for this postulate, however, is not very conclusive.

The antisymmetric states are the most important ones from the standpoint of the electron theory of solids. Although the actual process of selecting such states will be discussed in detail in sections that deal with the approximate solutions of the Schrödinger equation, there are several points that should be brought out here.

Suppose that we have a system that contains n electrons. Let us designate them by integers ranging from 1 to n in order to establish a normal arrangement. There are $n!$ possible different permutations of these n indices, each of which relabels the electrons in different ways. The Hamiltonian operator is invariant under these permutations of indices. For each permutation, say the νth of the set of $n!$, we may introduce a permutation operator P_ν, which is defined in such a way that it permutes electrons in the manner described by the νth permutation. Any operator α that is invariant under the νth permutation satisfies the relation

$$P_\nu \alpha g = \alpha P_\nu g$$

for an arbitrary function of g. We may write the operator equation corresponding to this

$$P_\nu \alpha = \alpha P_\nu$$

and may say that α commutes with P_ν. It is quite clear that two permutation operators generally do not satisfy the relationship

$$P_\nu P_\nu g = P_\nu P_\nu g$$

for arbitrary g. The Pauli restriction on the allowable state functions, namely,

$$P_\nu f = (-1)^{p(\nu)} f,$$
$$P_\nu f = (-1)^{p(\nu)} f,$$

where $p(\nu)$ is the order of the νth permutation, implies, however, that

$$P_\nu P_\nu f = P_\nu P_\nu f$$

[1] See, for example, L. BRILLOUIN, *Quantenstatistik* (Julius Springer, Berlin, 1930).

for all allowable state functions. Thus, for our purposes, all permutation operators may be said to commute.

The one-electron spin operators do not commute with the permutation operators, since each spin operator refers to a specific electron. Hence, we cannot expect to find functions that are simultaneous eigenfunctions of H, of $P_{r1}, \cdots, P_{rn}$, and of the one-electron spin quantities. On the other hand Σ^2 and Σ_z, which commute with one another, do commute with the P_r since they contain all electron variables symmetrically. Hence, when H is independent of spin, we may expect to find stationary states that satisfy the Pauli principle and are eigenfunctions of Σ^2 and Σ_z. For this reason, the physically interesting states are eigenfunctions of Σ^2 and Σ_z rather than eigenfunctions of the one-electron operators.

We shall accept without proof the following theorems[1] concerning the eigenvalues of Σ^2 and Σ_z.

a. For a system of n electrons, the eigenvalues of Σ^2 have the value

$$S(S+1)\hbar^2$$

where S may range from $n/2$ down to 0 or $\frac{1}{2}$, depending respectively upon whether n is even or odd. S is called the total spin quantum number, and the eigenfunctions of Σ^2 are said to be states of definite multiplicity.

b. There are $2S+1$ degenerate states associated with each value of S. These states may be chosen in such a way that they are eigenfunctions of Σ_z and have eigenvalues ranging from $+S$ to $-S$, by integer steps. The number $2S+1$ is called the multiplicity of the degenerate level associated with S.

Rules for constructing eigenfunctions of definite multiplicity may be found in the references in footnote.[2]

PART B. THE INTERACTION BETWEEN MATTER AND RADIATION*

We shall develop the theory of radiation[3] in this part of the present chapter for use in discussing the optical properties of solids. The

[1] See, for example, E. U. CONDON and G. H. SHORTLEY, *The Theory of Atomic Spectra* (Cambridge University Press, 1935).

[2] Eigenfunctions of definite multiplicity are used extensively in the theory of molecular valence, in which they are called "bond functions." A discussion of the theory of these functions may be found in the following papers: H. Eyring and G. E. Kimball, *Jour. Chem. Phys.*, **1**, 239 (1933); G. Rumer, *Nachr. Gott.*, M. P. Klasse, 337 (1932); J. H. Van Vleck and A. Sherman, *Rev. Modern Phys.*, **7**, 167 (1935).

* This part is used primarily in Chap. XVII, which deals with the optical properties of solids, and may be omitted by a reader not immediately interested in this topic.

[3] General references: G. BREIT, *Rev. Modern Phys.*, **4**, 504 (1932); E. FERMI, *Rev. Modern Phys.*, **4**, 87 (1932); W. HEITLER, *The Quantum Theory of Radiation* (Oxford University Press, New York, 1936).

topics of principal interest are the theory of light quanta and the theory of the interaction between matter and light. We shall begin with a brief discussion of the classical theory and shall use this to develop the quantum equations.

42. The Classical Electromagnetic Equations.[1] *a. The Radiation Field.*—Maxwell's equations for free space are

$$\begin{aligned} \operatorname{div} \mathsf{E} &= 0, & \operatorname{div} \mathsf{H} &= 0, \\ \operatorname{curl} \mathsf{E} &= -\frac{1}{c}\frac{\partial \mathsf{H}}{\partial t}, & \operatorname{curl} \mathsf{H} &= \frac{1}{c}\frac{\partial \mathsf{E}}{\partial t}, \end{aligned} \tag{1}$$

where E and H are the electric and magnetic field intensities. The time-dependent solutions of these equations that correspond to light waves have the form

$$\left.\begin{aligned} \mathsf{E} &= \mathsf{E}_0 e^{2\pi i(\eta\cdot\mathbf{r}-\nu t)}, \\ \mathsf{H} &= \mathsf{H}_0 e^{2\pi i(\eta\cdot\mathbf{r}-\nu t)}, \end{aligned}\right\} \tag{2}$$

where E_0 and H_0 are constant vectors, $\mathbf{n}$ is the wave number vector, and ν is the frequency of the wave. These quantities are interrelated by the equations

$$\left.\begin{aligned} \mathsf{E}_0 \cdot \mathbf{n} &= 0, & \mathsf{H}_0 \cdot \mathbf{n} &= 0, \\ \frac{\mathbf{n}^2}{\nu^2} &= \frac{1}{c^2}, & \mathsf{H}_0 \cdot \mathsf{E}_0 &= 0, \\ & & \mathsf{H}_0^2 &= \mathsf{E}_0^2, \end{aligned}\right\} \tag{3}$$

from which it may be concluded that E_0 and H_0 are orthogonal to one another and to the direction of propagation of the light wave.

It is convenient to express Maxwell's equations and the solution (2) in terms of the vector and scalar potentials A and φ, which are related to E and H by the equations

$$\left.\begin{aligned} \mathsf{E} &= -\frac{1}{c}\frac{\partial \mathsf{A}}{\partial t} - \operatorname{grad} \varphi, \\ \mathsf{H} &= \operatorname{curl} \mathsf{A}. \end{aligned}\right\} \tag{4}$$

The equations for A and φ are

$$\begin{aligned} \Delta\varphi - \frac{1}{c^2}\frac{\partial^2 \varphi}{\partial t^2} &= 0, & \Delta \mathsf{A} - \frac{1}{c^2}\frac{\partial^2 \mathsf{A}}{\partial t^2} &= 0, \\ & & \operatorname{div} \mathsf{A} &= \frac{1}{c}\frac{\partial \varphi}{\partial t}. \end{aligned} \tag{5}$$

[1] M. ABRAHAM and R. BECKER, *Classical Electricity and Magnetism* (Blackie & Son, Ltd., London, 1932); R. BECKER, *Theorie der Elektrizität*, Vol. II (Julius Springer, Berlin, 1932).

In the case of light waves, we may set φ equal to zero and take A to be

$$\mathsf{A} = \frac{c}{2\pi i\nu}\mathsf{E}_0 e^{2\pi i(\eta r - \nu t)}. \tag{6}$$

It may be seen that Eqs. (4) then lead to Eqs. (2) and (3).

If we have a number of waves of different frequencies in a cubical space of volume V, the real vector potential of the system may be expressed in the form of a series of traveling waves

$$\mathsf{A} = \sum_{\eta,s} \mathbf{f}_s(\mathbf{n}) \frac{c}{[\nu_s(\mathbf{n})V]^{\frac{1}{2}}} \{A_s(\mathbf{n}) e^{2\pi i(\eta r - \nu_s(\eta)t)} + A_s^+(\mathbf{n}) e^{-2\pi i(\eta r - \nu_s(\eta)t)}\}. \tag{7}$$

The quantities that appear in this equation are very similar to those that appear in the Fourier resolutions of the atomic motions in lattice theory (*cf.* Sec. 22). The quantities $\mathbf{f}_s(\mathbf{n})(s = 1, 2)$ are real polarization vectors that satisfy the equations

$$\left.\begin{aligned} \mathbf{f}_s(\mathbf{n}) \cdot \mathbf{n} = 0, \qquad \mathbf{f}_1(\mathbf{n}) \cdot \mathbf{f}_2(\mathbf{n}) = 0, \\ \mathbf{f}_s \cdot \mathbf{f}_s = 1, \end{aligned}\right\} \tag{8}$$

and the $A_s(\mathbf{n})$ are complex constants that are proportional to the amplitudes of the wave of wave number $\mathbf{n}$ having E in the direction $\mathbf{f}_s(\mathbf{n})$. Since we shall deal with real quantities, we shall assume that

$$A_s^*(\mathbf{n}) = A_s^+(\mathbf{n}). \tag{9}$$

The summation of $\mathbf{n}$ extends over the allowed values of this variable, which are given by the equations

$$\eta_x = \frac{n_x}{L}, \qquad \eta_y = \frac{n_y}{L}, \qquad \eta_z = \frac{n_z}{L} \tag{10}$$

where n_x, n_y, and n_z are integers and L is the length of an edge of the box. It may be seen from this that the density of the allowed values of $\mathbf{n}$ in wave-number space is $L^3 = V$. The factor $c/(\nu V)^{\frac{1}{2}}$ is introduced into (7) in order to simplify results that we shall obtain below.

We shall now derive a Hamiltonian function for this radiation field using Eq. (7). In material systems, the Hamiltonian, when independent of time, is equal to the total energy. Hence, we should expect to obtain a Hamiltonian for the radiation field by computing the energy E_r, namely,

$$E_r = \frac{1}{8\pi}\int (\mathsf{E}^2 + \mathsf{H}^2) dV. \tag{11}$$

If Eqs. (4) are used to compute E and H from the vector potential, Eq. (11) becomes

$$E_r = \sum_{\eta,s} 2\pi\nu_s(\mathbf{n}) a_s(\mathbf{n}) a_s^*(\mathbf{n}) \tag{12}$$

where

$$\left.\begin{aligned} a_s(\mathbf{n}) &= A_s(\mathbf{n})e^{-2\pi i\nu_s(\eta)t} \\ a_s^*(\mathbf{n}) &= A_s^*(\mathbf{n})e^{+2\pi i\nu_s(\eta)t}. \end{aligned}\right\} \tag{13}$$

If we now regard $a_s(\mathbf{n})$ and $ia_s^*(\mathbf{n})$ as conjugate variables and (12) as the Hamiltonian of the system, we find that the Hamiltonian equations are

$$\begin{aligned} \dot{a}_s(\mathbf{n}) &= \frac{\partial H_R}{\partial i a_s^*(\mathbf{n})} = -2\pi i\nu_s(\mathbf{n})a_s(\mathbf{n}), \\ \dot{a}_s^*(\mathbf{n}) &= \frac{\partial H_R}{\partial a_s(\mathbf{n})} = 2\pi i\nu_s(\mathbf{n})a_s^*(\mathbf{n}). \end{aligned} \tag{14}$$

The solutions of these equations lead to the time dependence expressed by Eqs. (13). Hence, (11) actually is the Hamiltonian function of the system.

Since the variables $a_s(\mathbf{n})$ and $a_s^*(\mathbf{n})$ are not real, it is convenient to replace them by the real quantities

$$\begin{aligned} p_s(\mathbf{n}) &= \frac{1}{\sqrt{2}}(a_s + a_s^*), \\ q_s(\mathbf{n}) &= \frac{i}{\sqrt{2}}(a_s - a_s^*). \end{aligned} \tag{15}$$

The Hamiltonian, when expressed in terms of these variables, is

$$H_R = \frac{1}{2}\sum_{s,\eta} 2\pi\nu_s(\mathbf{n})[p_s^2(\mathbf{n}) + q_s^2(\mathbf{n})], \tag{16}$$

which is the same as for a system of harmonic oscillators. It is this fact that justifies the statement that a radiation field is equivalent to a system of harmonic oscillators. The Hamiltonian (16) may be used as the starting point for a discussion of the quantum theory of the radiation field.

If we express the vector potential A in terms of the a and the p and q, we find

$$\begin{aligned} \mathsf{A} &= \sum_{\eta,s} \mathfrak{f}_s(\mathbf{n})\frac{c}{[\nu_s(\mathbf{n})V]^{\frac{1}{2}}}[a_s(\mathbf{n})e^{2\pi i\eta\cdot\mathbf{r}} + a_s^*(\mathbf{n})e^{-2\pi i\eta\cdot\mathbf{r}}] \\ &= \sum_{\eta,s} \mathfrak{f}_s(\mathbf{n})\frac{c}{\sqrt{2}[\nu_s(\mathbf{n})V]^{\frac{1}{2}}}(p_s\cos 2\pi\mathbf{n}\cdot\mathbf{r} + q_s\sin 2\pi\mathbf{n}\cdot\mathbf{r}). \end{aligned} \tag{17}$$

b. The Interaction between Matter and Radiation.[1]—The force that acts on an electron of charge $-e$ that is moving with velocity $\dot{\mathbf{r}}$ in an electro-

[1] M. ABRAHAM and R. BECKER, *Classical Electricity and Magnetism* (Blackie & Son, Ltd., London, 1932); R. BECKER, *Theorie der Elektrizität*, Vol. II (Julius Springer, Berlin 1932).

magnetic field is given by Lorentz's equation

$$\mathbf{f} = -e\mathbf{E} - \frac{e}{c}\dot{\mathbf{r}} \times \mathbf{H}.$$

Hence, the Newtonian equation of motion is

$$m\ddot{\mathbf{r}} = -e\mathbf{E} - \frac{e}{c}\dot{\mathbf{r}} \times \mathbf{H}.$$

This equation may be derived from the Lagrangian function

$$L = \frac{1}{2}m\dot{\mathbf{r}}^2 - \frac{e}{c}\mathbf{A} \cdot \dot{\mathbf{r}} + e\varphi, \tag{18}$$

as may be verified by writing out the Eulerian equations associated with this function. The components of momenta are

$$p_x = \frac{\partial L}{\partial \dot{x}} = m\dot{x} - \frac{e}{c}A_x,$$

etc., so that the Hamiltonian $H = \mathbf{p} \cdot \dot{\mathbf{r}} - L$ is

$$\begin{aligned} H &= \frac{1}{2m}\left(\mathbf{p} + \frac{e}{c}\mathbf{A}\right)^2 - e\varphi \\ &= \frac{1}{2m}\mathbf{p}^2 + \frac{e}{mc}\mathbf{p} \cdot \mathbf{A} + \frac{e^2}{2mc^2}\mathbf{A}^2 - e\varphi. \end{aligned} \tag{19}$$

The terms

$$H_I = \frac{e}{mc}\mathbf{p} \cdot \mathbf{A} + \frac{e^2}{2mc^2}\mathbf{A}^2 \tag{20}$$

are of interest when the charge is in a radiation field because they give the interaction between the field and the particle. The other two terms constitute the Hamiltonian in the absence of a field.

It should now be clear that the total Hamiltonian of a system of n electrons in a radiation field is

$$\begin{aligned} H &= H_M + \sum_i \left[\frac{e}{mc}\mathbf{p}_i \cdot \mathbf{A}(\mathbf{r}_i) + \frac{e^2}{2mc^2}\mathbf{A}^2(\mathbf{r}_i)\right] \\ &= H_M + H_I \end{aligned} \tag{21}$$

where H_M is the Hamiltonian in the absence of a radiation field and H_I is the interaction term. If we desire to include the radiation field in the Hamiltonian, we must add to (21) the function H_R that was derived in Part A. The total Hamiltonian then is

$$H_T = H_M + H_I + H_R. \tag{22}$$

43. The Semiclassical Method of Treating Radiation.—Previous to the development of Dirac's theory of radiation,[1] which, at present, is the most accurate method of treating radiation problems, Schrödinger,[2] Gordon,[3] and Klein[4] developed a simpler theory which is still useful for obtaining some of the important equations in a simple way. In this scheme, the classical interaction term H_I in the Hamiltonian (21) of the preceding section is treated as a time-dependent perturbation in the Schrödinger equation. If the radiation field is zero in the classical sense that the system is in the dark, the system is unperturbed and cannot change its state. If, on the other hand, the field is finite, the system may change its state by emission, absorption, or scattering of light. The fact that the state can change only if radiation is present shows a fundamental weakness in the theory, for it cannot be used to treat the problem of spontaneous emission of light by an excited atom. In spite of this weakness, the method leads to many correct results relatively simply. An important reason for this simplicity is the fact that the Schrödinger-Gordon-Klein method can treat absorption, emission, and dispersion in an approximation in which the $e^2\mathbf{A}^2/2mc^2$ terms of Eq. (21) are neglected as small quantities, whereas the Dirac theory can treat dispersion only in an approximation in which all the terms of H_I are retained.

The semiclassical method will now be used to develop the equations for absorption, emission, and dispersion of radiation by an atom. We shall assume that the radiation field extends continuously over a finite range of frequencies and that the vector potential of the wave of wave number $\mathbf{n}$ is

$$\mathsf{A}(\mathbf{n}) = \frac{c}{2\pi i\nu}\mathsf{E}_0(\mathbf{n})[e^{2\pi i(\eta\cdot\mathbf{r}-\nu t)} - e^{-2\pi i(\eta\cdot\mathbf{r}-\nu t)}] \tag{1}$$

where E_0 is the amplitude of the electrostatic field. The interaction term in the Hamiltonian arising from this vector potential is

$$H_I' = \sum_i \frac{e}{2\pi i\nu m}\mathbf{p}_i\cdot\mathsf{E}_0[e^{2\pi i(\eta\cdot\mathbf{r}_i-\nu t)} - e^{-2\pi i(\eta\cdot\mathbf{r}_i-\nu t)}]. \tag{2}$$

We shall designate the Hamiltonian of the atom by H_M and the stationary states by Ψ_i where

$$H_M\Psi_i = E_i\Psi_i. \tag{3}$$

[1] See footnote 3, p. 210.
[2] E. SCHRÖDINGER, *Ann. Physik*, **81**, 134 (1926).
[3] W. GORDON, *Z. Physik*, **40**, 117 (1926).
[4] O. KLEIN, *Z. Physik*, **37**, 895 (1926).

It is now of interest to search for a solution of the time-dependent Schrödinger equation

$$(H_M + H_I')f = -\frac{\hbar}{i}\frac{\partial f}{\partial t} \tag{4}$$

that is equal to Ψ_0 at time $t = 0$. This solution may be expressed in terms of the Ψ_i in the following way:

$$f = a_0\Psi_0 e^{-\frac{i}{\hbar}E_0 t} + \sum_{i\neq 0} a_i\Psi_i e^{-\frac{i}{\hbar}E_i t} \tag{5}$$

where a_0 is unity at time zero and the a_i are small quantities that are zero at the same time. Substituting this in Eq. (4), we find after a simple reduction that employs Eq. (3)

$$-\frac{\hbar}{i}\sum_i \frac{\partial a_i}{\partial t}\Psi_i e^{-\frac{i}{\hbar}E_i t} = a_0 H_I'\Psi_0 e^{-\frac{i}{\hbar}E_0 t} \tag{6}$$

in which small terms involving the product of a_i and H_I' have been neglected. If this equation is multiplied by Ψ_j^* and the result is integrated over the electronic coordinates, it is found that

$$-\frac{\hbar}{i}\frac{\partial a_j}{\partial t} = a_0 H_{j0} e^{-\frac{i}{\hbar}(E_0 - E_j)t} \tag{7}$$

where

$$H_{j0} = \int \Psi_j^* H_I'\Psi_0 d\tau. \tag{8}$$

In deriving (7), we have used the relation

$$\int \Psi_j^*\Psi_i d\tau = \delta_{ij}. \tag{9}$$

We shall assume in the following work that a_0 is close to unity at all times in which we are interested. Since H_{j0} involves time, it is convenient to show this dependence explicitly by using Eq. (2). Equation (8) then becomes

$$H_{j0} = \mathsf{E}_0 \cdot (\mathbf{C}_{j0} e^{-2\pi i\nu t} - \mathbf{C}_{j0}^{+} e^{2\pi i\nu t}) \tag{10}$$

where

$$\mathbf{C}_{j0} = \frac{e}{2\pi i\nu m}\int \Psi_j^*\left(\sum_i \mathbf{p}_i e^{2\pi i\eta\cdot r_i}\right)\Psi_0 d\tau \tag{11}$$

and $\mathbf{C}_{j0}^{+}$ is the same quantity with the sign of η reversed. Thus, Eq. (7) is

$$-\frac{\hbar}{i}\frac{\partial a_j}{\partial t} = \mathsf{E}_0 \cdot [\mathbf{C}_{j0} e^{-\frac{i}{\hbar}(E_0 - E_j + h\nu)t} - \mathbf{C}_{j0}^{+} e^{-\frac{i}{\hbar}(E_0 - E_j - h\nu)t}].$$

The integral of this equation that vanishes at $t = 0$ is

$$a_j = -\mathsf{E}_0 \cdot \left[\mathsf{C}_{j0} \frac{1 - e^{-\frac{i}{\hbar}(E_0 - E_j + h\nu)t}}{E_0 - E_j + h\nu} - \mathsf{C}_{j0}^{+} \frac{1 - e^{-\frac{i}{\hbar}(E_0 - E_j - h\nu)t}}{E_0 - E_j - h\nu} \right]. \quad (12)$$

When either one of the relations

$$E_0 - E_j \pm h\nu = 0 \quad (13)$$

is not closely satisfied, a_j oscillates very rapidly with time about the value zero. In this case, it may be said that the atom remains in the state Ψ_0 and behaves like a system undergoing forced oscillations far from resonance. On the other hand, when either one of the relations (13) is satisfied we may say that the atom resonates and changes its state. We shall first discuss the case of resonance and return to the other later.

In the case of resonance, we may interpret $|a_j|^2$ as the probability that at time t the system is found in the state Ψ_j. Since $h\nu$ is positive, the case

$$E_0 - E_j = h\nu \quad (14)$$

can occur only when $E_j < E_0$, so that it corresponds to induced emission. The case

$$E_j - E_0 = h\nu, \quad (15)$$

on the other hand, corresponds to induced absorption. The probability that either process occurs in time t is, respectively,

$$P_\nu(t) = |\mathsf{E}_0 \cdot \mathsf{C}_{j0}|^2 \omega(E_0 - E_j \pm h\nu) \quad (16)$$

where the negative sign corresponds to (14) and the positive to (15), and

$$\omega(\epsilon) = \frac{2\left(1 - \cos \frac{\epsilon t}{\hbar}\right)}{\epsilon^2} = \frac{4 \sin^2 \frac{\epsilon t}{2\hbar}}{\epsilon^2}. \quad (17)$$

This varies quadratically with time at small time for any given value of ν. If the radiation is continuous and extends over a sufficiently broad range of the spectrum, however, the total probability varies linearly with time. Let us suppose that the energy of the radiation that lies in the range of frequency from ν to $\nu + d\nu$ and is polarized in the direction $\mathbf{n}_0$ is $\rho(\nu)d\nu$, where ρ is practically constant near the resonance frequency. Since the mean density is related to the amplitude E_0 of (1) by the equation

$$\rho = \frac{\mathsf{E}_0^2}{2\pi},$$

we may rewrite Eq. (16) in the form

$$P_\nu(t) = 2\pi\rho|\mathbf{n}_0 \cdot \mathbf{C}_{j0}|^2\omega(E_0 - E_j \pm h\nu). \tag{18}$$

The total probability of the transition then is

$$P(t) = \int_0^\infty P_\nu(t)d\nu = 2\pi\rho\int_0^\infty |\mathbf{n}_0 \cdot \mathbf{C}_{j0}|^2\omega(E_0 - E_j \pm h\nu)d\nu. \tag{19}$$

The function ω has a peak of half width $\Delta\nu = 1/t$ at the resonance frequency. For the optical absorption and emission time in which we shall be interested, t is of the order of 10^{-8} sec, so that $h\Delta\nu$ is of the order of 10^{-8} ev, which is very small compared with ordinary values of ν. Hence, we may replace ω by a delta function of the same area. Now,

$$\int_{-\infty}^{\infty} 4\frac{\sin^2\frac{\epsilon t}{2\hbar}}{\epsilon^2}d\epsilon = \frac{2t}{\hbar}\int_{-\infty}^{\infty}\frac{\sin^2 x}{x^2}dx = \frac{2\pi t}{\hbar}.$$

Hence, $\omega(\epsilon)$ may be replaced by

$$\frac{2\pi t}{\hbar}\delta(\epsilon) \tag{19a}$$

where $\delta(\epsilon)$ is a delta function that satisfies the equation

$$\int_{-\Delta\epsilon}^{\Delta\epsilon}\delta(\epsilon)d\epsilon = 1.$$

Thus, $P(t)$ is

$$P(t) = \frac{8\pi^3}{h^2}|\mathbf{n}_0 \cdot \mathbf{C}_{j0}|^2\rho_\nu t \tag{20}$$

where ρ_ν is the density at the frequency defined by Eq. (14) or (15).

Equation (20) shows that the induced emission and absorption probabilities are proportional to the radiation density and that the selection rules for transitions are determined by the matrix components of the operator $\sum_i \mathbf{p}_i e^{2\pi i\mathbf{q}\cdot\mathbf{r}_i}$. These matrix components can easily be expressed in terms of the matrix components of the atomic dipole moment

$$\mathbf{M} = -\sum_i e\mathbf{r}_i \tag{21}$$

in the case in which $\mathbf{n}\cdot\mathbf{r}$ varies so little over the atomic system that $e^{2\pi i\mathbf{q}\cdot\mathbf{r}}$ can be replaced by unity. If the Schrödinger equation

$$H\Psi_k = E_k\Psi_k$$

is multiplied by $\left(\sum_i \mathbf{r}_i\right)\Psi_j^*$ and the result is subtracted from the equation

$$H\Psi_j^* = E_j\Psi_j^*,$$

after the latter is multiplied by $\left(\sum_i \mathbf{r}_i\right)\Psi_k$, it is found that

$$(E_j - E_k)\int \Psi_j^*\left(\sum \mathbf{r}_i\right)\Psi_k d\tau = -\frac{\hbar^2}{2m}\int\left(\sum_i \mathbf{r}_i\right)\left[\Psi_k\left(\sum_i \Delta_i\right)\Psi_j^* - \Psi_j^*\left(\sum_i \Delta_i\right)\Psi_k\right]d\tau. \quad (22)$$

The right-hand side may, by use of Green's theorem, be transformed to

$$\frac{\hbar^2}{2m}\int \sum_i (\Psi_k \operatorname{grad}_i \Psi_j^* - \Psi_j^* \operatorname{grad}_i \Psi_k)d\tau. \quad (23)$$

It is assumed that Ψ_k vanishes at large distances, so that the surface integrals in Green's formula can be dropped. In a similar way, the integral of the second term in (23) may be shown to be equal to the first. Hence,

$$\sum_i \int \Psi_j^* \operatorname{grad}_i \Psi_k d\tau = -\frac{m}{\hbar^2}(E_j - E_k)\int\left(\sum_i \mathbf{r}_i\right)\Psi_j^*\Psi_k d\tau. \quad (24)$$

According to Eq. (11),

$$\mathbf{C}_{j0} = -\frac{e\hbar}{2\pi\nu m}\sum_i \int \Psi_j^*(\operatorname{grad}_i \Psi_0)d\tau$$

when $e^{2\pi i\eta\cdot\mathbf{r}}$ is unity, so that we find with the use of Eq. (24)

$$-\mathbf{C}_{j0} = \frac{E_j - E_0}{2\pi\nu\hbar}\int \Psi_j^*\left(\sum_i e\mathbf{r}_i\right)\Psi_0 d\tau = \pm\mathbf{M}_{j0}. \quad (25)$$

Here

$$\mathbf{M}_{j0} = -\int \Psi_j^*\left(\sum_i e\mathbf{r}_i\right)\Psi_0 d\tau, \quad (26)$$

and opposite signs are valid for the cases of Eqs. (14) and (15), respectively.

Equation (20) may now be written

$$P(t) = \frac{8\pi^3}{h^2}|\mathbf{M}_{j0}\cdot\mathbf{n}_0|^2\rho_\nu t. \quad (27)$$

The average value of the coefficient of $\rho_\nu t$ in this equation, namely,

$$\frac{8\pi^3}{3h^2}|\mathbf{M}_{j0}|^2, \tag{28}$$

is the Einstein B coefficient for induced transitions between two states Ψ_0 and Ψ_j. Here

$$|\mathbf{M}_{j0}|^2 = |M_{x,j0}|^2 + |M_{y,j0}|^2 + |M_{z,j0}|^2. \tag{29}$$

If the lowest atomic state is an S state,[1] which is spherically symmetrical in electron coordinates, the dipole matrix components are finite only if the excited state is a P function, which has the symmetry of a first-order surface harmonic. If the three degenerate P functions are chosen to have the symmetry of the functions $\sum_i x_i$, $\sum_i y_i$, and $\sum_i z_i$, respectively, and are labeled with indices x, y, z, it may be shown that

$$M_{x,x0} = M_{y,y0} = M_{z,z0},$$
$$M_{x,y0} = M_{x,z0} = \cdots = M_{y,x0} = \cdots = M_{z,y0} = 0.$$

Thus, light polarized in either the x, y, or z direction may induce a transition to one of the triply degenerate states.

In order to discuss by the semiclassical method the scattering of light when the frequency is not near the resonance frequency, it is necessary to compute the mean value of the atomic dipole moment $\mathbf{M}$ for the state f of Eq. (5). The result is a time-dependent function that contains terms which vary harmonically. We shall assume that the real coefficient of the term that varies as $e^{2\pi i\nu t}$ is the amplitude of a forced atomic oscillation of frequency ν. We shall also assume that $e^{2\pi i\eta\cdot r}$ may be replaced by unity.

The mean value of $\mathbf{M}$ for the state f is

$$\mathbf{M}' = \int f^*\mathbf{M}f\,d\tau$$
$$= \int \Psi_0^*\mathbf{M}\Psi_0\,d\tau + \sum_i{}' \left[a_i\mathbf{M}_{i0}^* e^{\frac{i}{\hbar}(E_0 - E_i)t} + a_i^*\mathbf{M}_{i0} e^{-\frac{i}{\hbar}(E_0 - E_i)t} \right] \tag{30}$$

in which terms involving squares of the a have been neglected. If it is recalled that a_i has the form

$$a_i = \mathsf{E}_0 \cdot \mathbf{M}_{i0} \frac{E_i - E_0}{h\nu} \left[\frac{1 - e^{-\frac{2\pi i(E_0 - E_i + h\nu)t}{\hbar}}}{E_0 - E_i + h\nu} - \frac{1 - e^{-\frac{2\pi i(E_0 - E_i - h\nu)t}{\hbar}}}{E_0 - E_i - h\nu} \right], \tag{31}$$

[1] We shall use the conventional notation of atomic spectra in which the states having total angular momentum 0, 1, 2, 3, 4, etc., in units of $\hbar$, are designated respectively by S, P, D, F, G, etc. See, for example, H. E. White, *Introduction to Atomic Spectra* (McGraw-Hill Book Company, Inc., New York, 1934).

it is found that the terms that oscillate with frequency ν are

$$\sum_k{}' \mathsf{E}_0 \cdot \mathbf{M}_{k0}\mathbf{M}_{k0}^* \frac{E_k - E_0}{h\nu}\left(\frac{1}{E_0 - E_k - h\nu} - \frac{1}{E_0 - E_k + h\nu}\right)(e^{-2\pi i\nu t} + e^{2\pi i\nu t})$$

$$= \sum_k \frac{\mathsf{E}_0 \cdot \mathbf{M}_{k0}\mathbf{M}_{k0}^*}{h} \frac{2\nu_{k0}}{\nu_{k0}^2 - \nu^2}(e^{-2\pi i\nu t} + e^{2\pi i\nu t}) \quad (32)$$

where

$$\nu_{k0} = \frac{(E_k - E_0)}{h}.$$

Since the electrostatic field intensity is

$$\mathsf{E} = \mathsf{E}_0(e^{-2\pi i\nu t} + e^{2\pi i\nu t}),$$

the atomic polarizability tensor α for unmodified scattering is

$$\alpha = \sum_k{}' \frac{\mathbf{M}_{k0}^*\mathbf{M}_{k0}}{h} \frac{2\nu_{k0}}{\nu_{k0}^2 - \nu^2}. \quad (33)$$

This tensor is a constant with the value

$$\sum_k{}' |\mathbf{M}_{0k}|^2 \frac{2\nu_{k0}}{\nu_{k0}^2 - \nu^2} \quad (34)$$

if the state Ψ_0 is an S state. The remaining terms in the mean value of $\mathbf{M}$ depend upon time through functions of the type

$$e^{-\frac{i}{h}(E_0 - E_i)t}.$$

These terms do not have significance as induced scattering terms and will not be considered further.

In order to discuss Raman, or modified, scattering of light, it is necessary to extend the preceding computation by considering matrix components of $\mathbf{M}$ between states f_0 and f_i. By a correspondence-principle argument[1] that is not very satisfactory, one may then arrive at the equation

$$\alpha_{\nu+\nu_{0n}} = \sum_k \left(\frac{\mathbf{M}_{0k}\mathbf{M}_{kn}}{\nu_{kn} + \nu} - \frac{\mathbf{M}_{kn}\mathbf{M}_{0k}}{\nu_{0k} + \nu}\right) \quad (35)$$

for the polarizability tensor associated with the absorption of frequency ν and the emission of frequency $\nu + \nu_{0n}$. The final frequency must, of course, be positive.

44. The Current Operator.—It is necessary to use the matrix components of the current operator in applying the semiclassical method to

[1] O. KLEIN, *Z. Physik*, **37**, 895 (1926).

solids, as we shall do in Chap. XVII. This operator may be derived in the following simple way. If we multiply the Schrödinger equation

$$-\frac{\hbar}{i}\frac{\partial}{\partial t}\Psi_k = -\sum_i \frac{\hbar^2}{2m}\Delta_i\Psi_k + \frac{\hbar e}{imc}\sum_i \mathsf{A}(\mathbf{r}_i)\cdot \operatorname{grad}_i \Psi_k + V\Psi_k \qquad (1)$$

by Ψ_k^* and subtract from it the equation for Ψ_k^* multiplied by Ψ_k, we obtain

$$\frac{\partial|\Psi_k|^2}{\partial t} = \frac{i\hbar}{2m}\sum_i (\Psi_k^*\Delta_i\Psi_k - \Psi_k\Delta_i\Psi_k^*) - \frac{e}{mc}\sum_i \operatorname{div}_i (\mathsf{A}(\mathbf{r}_i)\Psi_k^*\Psi_k), \qquad (2)$$

since div $\mathsf{A} = 0$ for the fields in which we shall be interested. Now,

$$\Psi_k^*\Delta_i\Psi_k - \Psi_k\Delta_i\Psi_k^* = \operatorname{div}_i (\Psi_k^* \operatorname{grad}_i \Psi_k - \Psi_k \operatorname{grad}_i \Psi_k^*),$$

so that Eq. (2) may be written

$$-e\frac{\partial|\Psi_k|^2}{\partial t} = -\sum_i \operatorname{div}_i\left[-\frac{e\hbar}{2mi}(\Psi_k^* \operatorname{grad}_i \Psi_k - \Psi_k \operatorname{grad}_i \Psi_k^*) - \frac{e^2}{mc}\mathsf{A}(\mathbf{r}_i)\Psi_k^*\Psi_k\right]. \qquad (3)$$

The quantity $-e|\Psi_k|^2$ may be regarded as the charge density ρ in $3N$-dimensional space. Hence, if we compare Eq. (3) with the equation of continuity, namely,

$$\frac{\partial\rho}{\partial t} = -\operatorname{div} \mathsf{J},$$

where J is the current, we obtain

$$\mathsf{J}_{kk} = \sum_i\left[-\frac{e\hbar}{2mi}(\Psi_k^* \operatorname{grad}_i \Psi_k - \Psi_k \operatorname{grad}_i \Psi_k^*) - \frac{e^2}{mc}\mathsf{A}(\mathbf{r}_i)\Psi_k^*\Psi_k\right]. \qquad (4)$$

We shall regard this as the diagonal matrix component of the $3N$-dimensional current operator of which the general element is

$$\mathsf{J}_{jk} = \sum_i\left[-\frac{e\hbar}{2mi}(\Psi_j^* \operatorname{grad}_i \Psi_k - \Psi_k \operatorname{grad}_i \Psi_j^*) - \frac{e^2}{mc}\mathsf{A}(\mathbf{r}_i)\Psi_j^*\Psi_k\right]. \qquad (5)$$

Now, the charge density $\rho(\mathbf{r}_1)$ of the first electron is

$$\rho(\mathbf{r}_1) = \int \rho(\mathbf{r}_1, \cdots, \mathbf{r}_n)\,d\tau' \qquad (6)$$

where the primed integral extends over the variables of all electrons except the first. Thus, the current associated with the first electron is

$$\mathbf{i}_{jk}(\mathbf{r}_1) = \sum \int' \Big[-\frac{e\hbar}{2mi}(\Psi_j^* \operatorname{grad}_i \Psi_k - \Psi_k \operatorname{grad}_i \Psi_j^*) - \frac{e^2}{mc}\mathbf{A}(\mathbf{r}_i)\Psi_j\Psi_k^* \Big] d\tau', \quad (7)$$

and the mean total current at the point $\mathbf{r}_1$ is

$$\mathbf{J}_{kk}(\mathbf{r}_1) = N \int' \Big[-\frac{e\hbar}{2mi}(\Psi_k^* \operatorname{grad}_i \Psi_k - \Psi_k \operatorname{grad}_i \Psi_k^*) - \frac{e^2}{mc}\mathbf{A}(\mathbf{r}_i)\Psi_k\Psi_k^* \Big] d\tau'. \quad (8)$$

45. Line Breadth.[1]—The absorption and emission lines of an atomic system that has discrete levels obviously should be infinitely narrow in the delta-function approximation of Eq. (19a), Sec. 43. The same result is obtained even if $\omega(\epsilon)$ is not replaced by a delta function, for the function

$$\omega(\epsilon) = \frac{\sin^2 \frac{\epsilon t}{2\hbar}}{\epsilon^2} \quad (1)$$

may be made as narrow as we please by making t large enough. Thus, after a long time the frequency distribution of the total radiation from a system of excited atoms should be very narrow. This result is a consequence of the approximations that were used in solving Eqs. (7), Sec. 43, and does not occur if the equations are solved more accurately by taking into account the fact that the coefficient $a_0(t)$ for the initial state is not a constant but varies with time.

It is possible to obtain a nearly self-consistent solution of Eq. (6) Sec. 43, in a number of important cases by assuming that

$$a_0 = e^{-2\pi\Gamma t} \quad (2)$$

where Γ is a constant. With this assumption, the equation for a_j is

$$-\frac{\hbar}{i}\frac{\partial a_j}{\partial t} = H_{j0}e^{-\frac{i}{\hbar}(E_0 - E_j - i\hbar\Gamma)t} \quad (3)$$

in the approximation in which terms other than a_0 may be neglected in the right-hand side. The solution of this equation is

$$a_j = -\mathbf{E}_0 \cdot \left[\mathbf{C}_{j0} \frac{1 - e^{-\frac{i}{\hbar}(E_0 - E_j - i\hbar\Gamma + h\nu)t}}{E_0 - E_j - i\hbar\Gamma + h\nu} - \mathbf{C}_{j0}^{+} \frac{1 - e^{-\frac{i}{\hbar}(E_0 - E_j - i\hbar\Gamma - h\nu)t}}{E_0 - E_j - i\hbar\Gamma - h\nu} \right]. \quad (4)$$

[1] See the survey article by V. Weisskopf, *Phys. Zeits.*, **34**, 1 (1933).

A value of Γ may be obtained by substituting (2) into the equation for a_0, namely,

$$-\frac{\hbar}{i}\frac{\partial a_0}{\partial t} = \sum_j H_{0j} a_j e^{-\frac{i(E_j - E_0)t}{\hbar}} \tag{5}$$

In the semiclassical theory, Γ is proportional to the energy density of radiation and is, as a consequence, very small when the radiation density is small. In the Dirac theory, however, it contains an additional constant term that is related to the probability coefficient for spontaneous emission from the upper to the lower member of the pair of levels between which the optical transitions are occurring.

The function $a_j(t)$ given by Eq. (4) reduces to $a_j(t)$ of Eq. (12), Sec. 43, when t is much smaller than $1/\Gamma$. Hence, the results of Sec. 43 are valid only for comparatively short times. At times long compared with $1/\Gamma$, the square of the absolute value of one of the terms in (4) reduces to

$$\frac{|\mathsf{E}_0 \cdot \mathsf{C}_{j0}|^2}{[(E_0 - E_j) \pm h\nu]^2 + h^2\Gamma^2} \tag{6}$$

a fact showing that the distribution of transition probability for different frequencies is governed by a function of the form

$$\frac{1}{[(E_0 - E_j) \pm h\nu]^2 + h^2\Gamma^2}, \tag{7}$$

which, as a function of ν, has a peak at

$$\nu = \frac{|E_0 - E_j|}{h} \tag{8}$$

of which the width at half maximum is

$$\Delta\nu = 2\Gamma. \tag{9}$$

It may be shown that Γ is of the order of magnitude 10^8 sec^{-1} for ordinary atomic transitions so that this *natural width* ordinarily is small compared with emission frequencies.

The fact that an emission line or an absorption line has a finite natural breadth does not imply that energy is not conserved. This breadth finds its origin in the fact that the energy of the excited atomic state is uncertain because the interaction between the atom and the radiation field, expressed by the term H_I, is uncertain. Thus the half width could be made as small as we please by making H_I sufficiently small.

Natural broadening is usually masked by one or more kinds of broadening that have a completely different origin. In the case of gases, for example, lines are broadened by the Doppler effect, since the atoms

move with different speeds, and by the interaction between atoms. Doppler broadening is not important in crystals, but the analogue of interaction broadening is important. We shall discuss it briefly here and in more detail in Chap. XVII.

If we have a large number of free atoms that are infinitely separated from one another and are stationary, their electronic levels are discrete

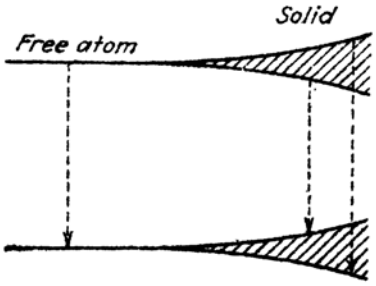

FIG. 1.—The discrete atomic energy levels may be broadened in the solid. If transitions between many levels in each band are allowed, the resulting emission "line" will be broad. This occurs in metals.

so that the energy levels of the entire system are discrete. Thus, the width of the emission lines is determined entirely by natural broadening. The atoms interact, however, if they are brought within a finite distance of one another, and this interaction broadens those levels of the entire system which were degenerate when the atoms were infinitely separated,

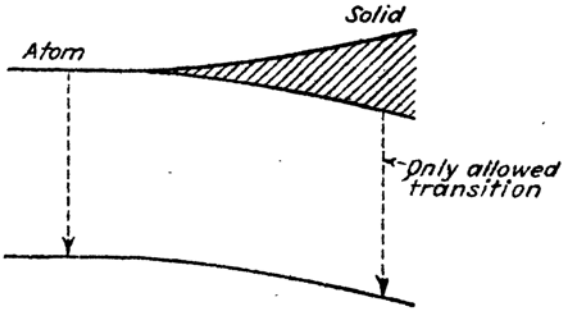

FIG. 2.—In this case the lowest level remains nondegenerate in passing from the system of free atoms to the solid. If selection rules forbid transitions from more than one of the excited states to the lower level, the emission line should be sharp. In actual cases of this type the line is broadened because of coupling between the electrons and the lattice.

as we shall see in Sec. 66. If transitions are allowed between many of the levels in two bands, the corresponding emission "line" should now have a breadth that is determined by the selection rules and the width of the bands. Thus the lines may broaden because the atomic levels are spread into bands (*cf.* Fig. 1).

It frequently happens, in the case of solids, that selection rules forbid transitions from more than one level of a band to a nondegenerate level

of the entire system (*cf.* Fig. 2). The corresponding line would have only the natural breadth, in this case, if only the electronic states of the system had to be taken into account. An assembly of atoms, however, has, in addition to its electronic states, a system of energy states that are associated with internuclear motion. In the case of solids, these states may be described in terms of lattice oscillations, as we have seen in Chap. III. It turns out that the vibrational states may be excited during transitions between electronic states. Since the range of allowed vibrational frequencies is continuous, the amounts of energy that can be given to the vibrational system are spread over a continuous range. Now, the energy that is not given to the vibrational system during spontaneous emission is radiated as light; hence, the emission lines have vibrational broadening. This vibrational broadening is one of the most important of the factors that determine the shape of absorption and emission bands of insulating solids at temperatures above absolute zero.

CHAPTER VI

APPROXIMATE TREATMENTS OF THE MANY-BODY PROBLEM

46. Introduction.—This chapter deals with some of the methods that have been devised to handle the Schrödinger equation for a mechanical system in which there are at least two interacting particles. In practically all these schemes, an attempt is made to select one member of a given set of admittedly approximate wave functions by use of the variational theorem (*cf.* Chap. V, Sec. 39). In one method—the variational method—the set of approximate state functions is obtained by writing down a definite function of the electronic variables that contains a number of parameters. The best function of the family is chosen by fixing these parameters so that the mean value of the Hamiltonian is a minimum. In another scheme—that of Hartree, Fock, and Slater—the starting set of functions is chosen as a combination of one-particle functions, that is, functions that involve the coordinates of only one particle. The one-particle functions are then determined by use of the variational theorem.

Since the exact solutions of the Schrödinger equation for any many-body system usually are very intricate functions of all variables and since the functions that may be manipulated are restricted to fairly simple types, even the best function obtained by one of the approximate treatments usually leads to an energy that differs appreciably from the experimental value. There are exceptions to this statement, such as the case of the normal state of helium, which we shall find useful for locating the cause of error in other problems.

We shall begin the discussion with a few remarks concerning the Hamiltonian operator that will be used in solids. These will be followed by a presentation of the two schemes that were described above.

47. The Hamiltonian Function and Its Mean Value.—In order to place the problem of determining the stationary states of solids upon a working basis, it is necessary to overlook certain terms in the Hamiltonian operator that are of secondary importance. First, we shall neglect the effects that arise from the motion of nuclei, assuming that the nuclei are at rest. The nuclear coordinates then enter into the Hamiltonian as parameters. In later chapters that deal with phase-changes, conductivity, and optical properties, we shall be interested both in the motion of nuclei and in the effect of nuclear motion on electrons. Second, we shall