杨振宁
的三篇学位论文

Chen Ning Yang's Theses With Commentary

朱邦芬　阮东　编

清華大學出版社
北　京

图书在版编目(CIP)数据

杨振宁的三篇学位论文/朱邦芬, 阮东编.—北京：清华大学出版社，2022.9
ISBN 978-7-302-61649-8

Ⅰ.①杨…　Ⅱ.①朱…②阮…　Ⅲ.①物理学－文集　Ⅳ.①O4-53

中国版本图书馆 CIP 数据核字(2022)第 145518 号

责任编辑：魏贺佳
封面设计：何凤霞
责任校对：王淑云
责任印制：曹婉颖

出版发行：清华大学出版社
网　　址：http://www.tup.com.cn, http://www.wqbook.com
地　　址：北京清华大学学研大厦 A 座　　邮　　编：100084
社 总 机：010-83470000　　邮　　购：010-62786544
投稿与读者服务：010-62776969, c-service@tup.tsinghua.edu.cn
质量反馈：010-62772015, zhiliang@tup.tsinghua.edu.cn
印 装 者：北京博海升彩色印刷有限公司
经　　销：全国新华书店
开　　本：185mm × 260mm　　印　　张：11.5　　字　　数：254 千字
版　　次：2022 年 9 月第 1 版　　印　　次：2022 年 9 月第 1 次印刷
定　　价：98.00 元

产品编号：078177-01

谨以本书庆贺
杨振宁先生百岁寿辰

西南联大硕士研究生毕业时的杨振宁 (1944)

芝加哥大学博士研究生毕业时的杨振宁 (1948)

目　　录

Comments on My Three Theses……Chen Ning Yang / 1

Group Theory and the Vibration of Polyatomic Molecules
（学士学位论文，1942 年）……5

Investigations in the Statistical Theory of Superlattices
（硕士学位论文，1944 年）……43

On the Angular Distribution in Nuclear Reactions and Coincidence Measurements
（博士学位论文，1948 年）……103

杨振宁在国立西南联合大学期间的学业……146

杨振宁的国立西南联合大学及芝加哥大学成绩单……149

西南联大的人才培养和杨振宁先生的学术起步
——代后记……朱邦芬 / 153

Comments on My Three Theses

Chen Ning Yang

Professors B.F. Zhu and R. Dong proposed to publish the 3 theses which I wrote in the 1940s. Each of the three was very important for my later research work, important in very different ways. So I thought detailing their respective influence on me may be useful for graduate students at the beginning of their careers.

1942 thesis for my BSc degree

At the Southwestern Associated University where I matriculated, 1938 to 1942, students were required to submit a thesis for their BSc degree. I had taken a course in quantum theory from Professor T. Y. Wu (吴大猷). So I went to him asking him to be my supervisor. What then happened I had described, as follows[1]:

With Professor Wu, 1982 in StonyBrook.

> *he gave me a copy of an article by J. E. Rosenthal and G. M. Murphy in the 1936 volume of Reviews of Modern Physics. It was a review paper on group theory and molecular spectra. I was thus introduced to group theory in physics. In retrospect I am deeply grateful to Wu for this introduction, since it had a profound effect on my subsequent development as a physicist.*

The importance to me of this early introduction to group theory's role in physics cannot be overemphasized. And I was deeply aware of this already in the early 1950s. In October 1957, one day after I learned I was to receive the Nobel Prize jointly with T.D. Lee, I wrote to Professor Wu[2]:

---

[1] C.N. Yang. Selected Papers 1945–1980 With Commentary. Freeman and Company (1983), p. 5.

[2] C.N. Yang. Selected Papers 1945–1980 With Commentary. Freeman and Company (1983), p. 41.

> *At this moment of great excitement, that also calls for deep personal reflection, it is my privilege to express to you my deep gratitude for your having initiated me into the field of symmetry laws and group theory in the spring of 1942. A major part of my subsequent work, including the parity problem, is traceable directly or indirectly to the ideas that I learned with you that spring fifteen years ago. This is something that I have always had an urge to tell you, but today is a particularly appropriate moment.*

1944 thesis for my MSc degree

This thesis was written under the direction of Professor J. S. Wang (王竹溪). It consisted of two papers on the specific heat of alloys, using approximations which were very popular at the time. The main ideas of such approximations are now called mean field theory. Neither paper made any impact in the field, but they did introduce me to statistical mechanics, in two important ways:


Professor Wang, early 1980s in StonyBrook.

(1) I was deeply impressed by Gibbs. I still remember vividly today reading his papers on the phase rule in an obscure journal published in Connecticut. [It is amazing that such an obscure journal was in the library of LianDa.] More important, his little book *Elementary Principles of Statistical Mechanics* converted me to an ardent admirer of his. I was to write in 1963[3]:

> *The beauty of his Elementary Principles of Statistical Mechanics is sheer poetry.*

(2) Statistical mechanics became one of my two major areas of research work. I still remember today Professor Wang excitedly telling me, one day in 1945, the breakthrough Onsager had made in the Ising model. I tried to understand this breakthrough, first in Kunming, later in Chicago in 1947, both without success. But finally in 1949, in a station wagon ride, I learned from Luttinger of a new paper by Kaufman and Onsager. And that led to my lifelong interest in statistical mechanics.

---

[3] C.N. Yang. Selected Papers 1945–1980 With Commentary. Freeman and Company (1983), p. 71.

1948 thesis for my PhD degree

During the first 9 month of 1946 I worked closely with Professor Teller. He had around 6 or 7 graduate students, and met us every week or two for lunch, to discuss our research. He also asked me to grade exercise papers of his students. So I had ample opportunity to observe his style of doing physics. He had very good physical intuition, especially about symmetries in atomic, molecular and nuclear physics. But he lacked patience to fill in the logical steps behind his intuition. For example as early as 1941 he had made, in a paper with Critchfield, statements about complexities in nuclear reactions involving particles with spin, but did not give any proofs. I began to think about how to supply complete proofs.

With Professor Teller,
1982 in Brookhaven National Laboratory.

In the late 1940s low energy nuclear physics was one of the most active fields. In particular there began experiments about correlations, such as β–γ and γ–γ correlations. Theoretical calculations were published about such correlations, showing *very surprisingly* that the final results often were very simple after *unexpected* cancellations. I did some of these calculations and soon realized that the cancellations *must be mathematical consequences of the spherical symmetry of nuclear physics.* But to substantiate such a statement required detailed mathematical analysis. This I succeeded in doing after a few weeks of analysis, and that was how my PhD thesis came about.

With Professor Teller, 1990s.

This thesis greatly increased my appreciation of the power of symmetry considerations in understanding natural laws. It happened that to study the newly discovered “strange particles” it was necessary to first determine their spin, parity and other quantum umbers, i.e. their symmetry properties. Thus I was able, one year after my thesis, to publish a paper on the spin and parity of the π^0 meson. This paper made me famous because it was in direct competition with L. Landau.

I should mention here that this π^0 paper used heavily *field theory*, which I had learned, very thoroughly in 1943–1945, from Professor S. T. Ma (马仕俊).


With Professors Wu and Ma, 1949.

In Chicago I was interested in using symmetry considerations not only on experiment related problems, such as those in my PhD thesis, but also on a more fundamental problem: the basic equations governing interactions between particles. Thus in 1947 I tried to generalize Weyl's gauge symmetry to non-Abelian groups. This effort met with smooth sailing at the beginning, but soon got into messy technical problems, and I had to give up. Fortunately I did return to it in 1954, at Brookhaven with Robert Mills. We succeeded in overcoming the technical difficulty and published a short paper on it[4]. That paper has now become one of the most important papers in physics after WWⅡ.

---

[4] C.N. Yang. Selected Papers 1945–1980 With Commentary. Freeman and Company (1983), p. 19.

Group Theory and the Vibration of Polyatomic Molecules

1942 thesis for BSc degree

GROUP THEORY AND THE VIBRATION OF POLYATOMIC MOLECULES

Cheng-Ning Yang (楊振寧)

INTRODUCTION

Informations about the structure of molecules can always be drawn from the analysis of their vibrational spectra, but owing to the mathematical difficulties involved in the theoretical calculation, only very simple types of molecules can be studied. The method developed by Bethe[1] in 1929, and then more completely by Wigner[2], however, removed considerably this difficulty. It is our purpose here to present the method together with some of the developments after them. A new method of finding the symmetrical coordinates is given (§4), in which the symmetry is preserved from step to step in spite of the existence of redundant coordinates. The theorem in §5 which renders the calculation of the degree of degeneracy very simple is also believed to be new.

The Symmetry of a Molecule

§1 MATHEMATICAL EXPRESSION OF SYMMETRY

There are reasons to suppose that the nuclei in a molecule arrange themselves in symmetrical positions when in equilibrium; i.e. some operations (Consisting of reflections and rotations) bring the molecule into itself. (For molecules containing isotopes this statement must be slightly modified. cf. §18) If we choose a set of rectangular coordinate axes with the origin at the centre of mass of the molecule in equilibrium, each covering operation C can be represented by an orthogonal matrix Γ_c (order: 3×3) so that the point $\begin{pmatrix} X \\ Y \\ Z \end{pmatrix}$ is brought to $\Gamma_c \begin{pmatrix} X \\ Y \\ Z \end{pmatrix}$ by the operation. Let $\boldsymbol{r}_1$, $\boldsymbol{r}_2$, $\cdots$, $\boldsymbol{r}_n$ be a set of coordinates specifying the relative positions of the nuclei (e.g. the distances between the nuclei and the angles between the bonds) in the molecule. When the nuclei vibrate about their positions of equilibrium, these $\boldsymbol{r}$'s vary (cf. §6). Let R_1, R_2, $\cdots$, R_n be their increments. Further, let $x_1, y_1, z_1, x_2, \cdots, x_N, y_N, z_N$ be the increments of the rectangular coordinates of the N nuclei. For small vibrations the R's are linear in the x's, y's and z's:

$$R = \begin{pmatrix} R_1 \\ R_2 \\ \vdots \\ R_n \end{pmatrix} = B\mathscr{C}, \qquad \text{where } \mathscr{C} = \begin{pmatrix} x_1 \\ y_1 \\ \vdots \\ z_N \end{pmatrix} \tag{1}$$

B being a constant matrix of n rows and $3N$ colomns. Now after the operation C, the molecule is indistinguishable from its original self, and we have a new equation obtained by writing (1) down for the new molecule:

$$\overset{c}{R} = B\overset{c}{\mathscr{C}} \tag{2}$$

Here $\overset{c}{R}$'s are the coordinates of the molecule which will be brought into coincidence with R's by the operation C, and

$$\overset{c}{\mathscr{C}} = \begin{pmatrix} \Gamma_c \begin{pmatrix} x_{C^{-1}1} \\ y_{C^{-1}1} \\ z_{C^{-1}1} \end{pmatrix} \\ \vdots \\ \Gamma_c \begin{pmatrix} x_{C^{-1}N} \\ y_{C^{-1}N} \\ z_{C^{-1}N} \end{pmatrix} \end{pmatrix} \tag{3}$$

where $C^{-1}i$ is the nucleus which will become after the operation C the nucleus i. Let Z_c be a square matrix of order $3N$ with the elements

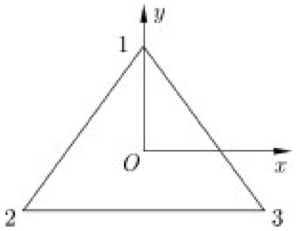
$${}_{ix}(Z_c)_{jx} = \delta_{j,C^{-1}i}, \quad {}_{ix}(Z_c)_{jy} = 0 \quad \text{etc.} \quad i, j = 1, 2, \cdots, N,$$

and let P_c stand for $\begin{pmatrix} \Gamma_c & & & \\ & \Gamma_c & & \\ & & \ddots & \\ & & & \Gamma_c \end{pmatrix}$, then (2) and (3) give

$$\overset{c}{R} = BP_cZ_c\mathscr{C}. \tag{4}$$

This equation holds for every operation C and is the mathematical expression of the symmetry of the molecule.

EXAMPLE Consider three equivalent nuclei forming an equilateral triangle. Let C be the operation: Rotation counterclockwise through 120° about O. Then



$$C^{-1}1 = 3, \quad C^{-1}2 = 1, \quad C^{-1}3 = 2,$$

and

$$Z_c = \left(\begin{array}{c|c|c} 0 & 0 & \begin{matrix}1 & & 0\\ & 1 & \\ 0 & & 1\end{matrix} \\ \hline \begin{matrix}1 & & 0\\ & 1 & \\ 0 & & 1\end{matrix} & 0 & 0 \\ \hline 0 & \begin{matrix}1 & & 0\\ & 1 & \\ 0 & & 1\end{matrix} & 0 \end{array}\right), \quad \Gamma_c = \begin{pmatrix} -\frac{1}{2} & -\frac{\sqrt{3}}{2} & 0 \\ \frac{\sqrt{3}}{2} & -\frac{1}{2} & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Let R_1 be the increment of the distance $\overline{12}$.

$$R_1 = \frac{1}{2}x_1 + \frac{\sqrt{3}}{2}y_1 - \frac{1}{2}x_2 - \frac{\sqrt{3}}{2}y_2.$$

Then $\overset{c}{R}_1$ is that of $\overline{31}$, and

$$\overset{c}{R}_1 = -\frac{1}{2}x_1 + \frac{\sqrt{3}}{2}y_1 + \frac{1}{2}x_3 - \frac{\sqrt{3}}{2}y_3.$$

Thus

$$B = \begin{pmatrix} \frac{1}{2} & \frac{\sqrt{3}}{2} & 0 & -\frac{1}{2} & -\frac{\sqrt{3}}{2} & 0 & 0 & 0 & 0 \end{pmatrix}.$$

(4) becomes the identity

$$-\frac{1}{2}x_1 + \frac{\sqrt{3}}{2}y_1 + \frac{1}{2}x_3 - \frac{\sqrt{3}}{2}y_3 = \begin{pmatrix} \frac{1}{2} & \frac{\sqrt{3}}{2} & 0 & -\frac{1}{2} & -\frac{\sqrt{3}}{2} & 0 & 0 & 0 & 0 \end{pmatrix} \cdot$$

$$\begin{pmatrix} \Gamma_c & 0 & 0 \\ 0 & \Gamma_c & 0 \\ 0 & 0 & \Gamma_c \end{pmatrix} Z_c \begin{pmatrix} x_1 \\ y_1 \\ \vdots \\ z_3 \end{pmatrix}.$$

§2 FUNDAMENTAL RELATIONSHIP

In some instances the coordinates R_1, R_2, $\cdots$, R_n are sufficient to determine $\overset{c}{R}_1$, $\overset{c}{R}_2$, $\cdots$, $\overset{c}{R}_n$ for all covering operations C. This is the case if (i) the R's contain only complete sets of equivalent coordinates (e.g. in the last example when $\begin{matrix}R_1\\R_2\\R_3\end{matrix} =$ increment of $\begin{matrix}3\widehat{1}2\\1\widehat{2}3\\2\widehat{3}1\end{matrix}$); or if (ii) the R's are all that are necessary to describe the internal structure of the molecule. In both cases we have for small vibrations $\overset{c}{R}=A_cR$, where A_c is in case (i) an orthogonal matrix having as elements 0 or 1, and in case (ii) a matrix of order $n \times n$. By (4),

$$BP_cZ_c\mathscr{C} = \overset{c}{R} = A_cR = A_cB\mathscr{C}.$$

But $\mathscr{C}$ is arbitrary (cf. §6), hence

$$BP_cZ_c = A_cB. \tag{5}$$

This is the fundamental relationship on which all the following deductions are based.

§3 GROUP PROPERTIES

To make further developments we notice that the covering operations C form a group and that the P_c's, Z_c's and A_c's each form a group isomorphic[*] with it. The group is known as the "point group". They are of such importance that their properties have been investigated in detail.[3]

Choice of Internal Coordinates

§4 INDEPENDENT REDUCED COORDINATES

We first choose the coordinates R_1, R_2, $\cdots$, R_n so that they contain only complete sets of equivalent internal coordinates, and such that they are more than necessary for the determination of the structure of the molecule. The simplest way is to choose the increments of the internuclear distances and the bond angles as the R's. In the example

---

[*] Let $C'C$ be the resultant operation of first operating C and then C', we have

$$P_{c'c} = P_{c'}P_c, \quad Z_{c'c} = Z_{c'}Z_c, \quad \text{but } A_{c'c} = A_cA_{c'}.$$

of §1 we may take the increments of the bonds $\overline{12}$, $\overline{23}$ and $\overline{31}$ as R_1, R_2 and R_3; or those of the lengths $\overline{O1}$, $\overline{O2}$, $\overline{O3}$ and the angles $1\widehat{O}2$, $2\widehat{O}3$ and $3\widehat{O}1$ as the R's. The matrix B can now be determined (§§11, 12). Evidently our choice belongs to the case (i) of §2, so that the A_c's are orthogonal and have as elements 0 or 1. It is plain that ${}_i(A_c)_j = 0$ if R_i and R_j are not equivalent. We shall make use of the following theorem in group theory[5]:

If A_c form a group of orthogonal matrices, and $W_c^\alpha (\alpha = 1, 2, \cdots, k)$ are the irreducible orthogonal representations of the group, there exists an orthogonal matrix M such that $W_c = MA_cM^{-1}$ is of the form

$$\begin{pmatrix} W_c^1 & & & & & & 0 \\ & W_c^1 & & & & & \\ & & \ddots & & & & \\ & & & W_c^1 & & & \\ & & & & W_c^2 & & \\ & & & & & \ddots & \\ 0 & & & & & & W_c^k \end{pmatrix}. \tag{6}$$

We define* $Q = \begin{pmatrix} Q_1 \\ Q_2 \\ \vdots \\ Q_n \end{pmatrix} = MR$ as the "reduced coordinates"[4]. Evidently

$$\overset{c}{Q} = M\overset{c}{R} = MA_cR = MA_cM^{-1}Q = W_cQ.$$

Now not all the Q's are independent. To select out the independent ones we need the following theorem:

THEOREM It is always possible to drop out some of the Q's so that

(i) the remaining ones are all independent,

(ii) the dropped ones depend on the remaining ones,

and (iii) the remaining ones belong to complete blocks of the group of matrices W_c.

Because of the properties (i) and (ii), the remaining coordinates $\mathscr{R} = \begin{pmatrix} \mathscr{R}_1 \\ \mathscr{R}_2 \\ \vdots \\ \mathscr{R}_{3N-6} \end{pmatrix}$ are called the "independent reduced coordinates", and from (iii), $\overset{c}{\mathscr{R}} = \mathfrak{A}_c\mathscr{R}$ where $\mathfrak{A}_c$ is

---

[*] We assume here that Q is real. There is probably no difficulty in handling complex coordinates, but they are not necessary unless molecules of symmetry C_n, C_{nh} or S_n $(n > 2)$ are studied.

obtained from W_c by dropping some blocks. These coordinates may also be obtained in other ways, and have been given various names by different authors. They are the "internal symmetry coordinates" of Wilson[7] and the "geometrical symmetry coordinates" of Rosenthal and Murphy[8]. The relation between $\mathscr{R}$ and the "symmetry coordinates" of Howard and Wilson[11,9] will be given in §14.

PROOF OF THE THEOREM Since Q is not entirely arbitrary, there exists a matrix H of order $\nu \times n$ so that

$$HH' = I, \quad HQ = 0, \tag{7}$$

where I is the unit matrix and the $'$ stands for "transposed". We may prove that the rank of H is ν and that there exists a matrix H_1 of order $(n-\nu) \times n$ such that $\begin{pmatrix} H \\ H_1 \end{pmatrix}$ is orthogonal. Put $\begin{pmatrix} H \\ H_1 \end{pmatrix} Q = \begin{pmatrix} 0 \\ Q_1 \end{pmatrix}$, then

$$Q = (H' \ H_1') \begin{pmatrix} 0 \\ Q_1 \end{pmatrix} = H_1' Q_1. \tag{8}$$

We may suppose that $\nu = n - (3N-6)$ so that Q_1 is arbitrary. The equ. $HQ = 0$ must be invariant under an operation C:

$$0 = H\overset{c}{Q} = HW_cQ = HW_cH_1'Q_1.$$

Hence

$$HW_cH_1' = 0.$$

Thus

$$\begin{pmatrix} H \\ H_1 \end{pmatrix} W_c(H' \ H_1') = \begin{pmatrix} HW_cH' & 0 \\ H_1W_cH' & H_1W_cH_1' \end{pmatrix}.$$

This matrix is orthogonal (for $\begin{pmatrix} H \\ H_1 \end{pmatrix}$ and W_c are both orthogonal). Hence

$$\begin{pmatrix} H \\ H_1 \end{pmatrix} W_c(H' \ H_1') = \begin{pmatrix} HW_cH' & 0 \\ 0 & H_1W_cH_1' \end{pmatrix}. \tag{9}$$

Thus HW_cH' is orthogonal and forms a group isomorphic with the one formed by C; i.e.

$$(HW_{c'}H')(HW_cH') = HW_{cc'}H'.$$

By the theorem quoted above there exists an orthogonal matrix J such that $JHW_cH'J' = \omega_c$ is of the form (6); i.e. if the "untermatrices" of W_c are

$$_{\alpha\beta}(W_c)_{\alpha'\beta'} = \delta_{\alpha\alpha'}\delta_{\beta\beta'}W_c^{\alpha}, \quad \alpha = 1, 2, \cdots, k, \quad \beta = 1, 2, \cdots, n_\alpha,$$

those of ω_c must be

$$_{ab}(\omega_c)_{a'b'} = \delta_{aa'}\delta_{bb'}W_c^{a}, \quad a = 1, 2, \cdots, k, \quad b = 1, 2, \cdots, \mu_a.$$

By (9), the irreducible blocks of ω_c must all be that of W_c, hence $\mu_a \leq n_\alpha$. Now

$$(JH)W_c = JHW_c(H'\ H_1')\begin{pmatrix} H \\ H_1 \end{pmatrix} = (JHW_cH' \quad 0)\begin{pmatrix} H \\ H_1 \end{pmatrix} = \omega_c JH \tag{10}$$

i.e.

$$_{ab}(JH)_{\alpha\beta}\, W_c^{\alpha} = W_c^{a}\ {}_{ab}(JH)_{\alpha\beta}.$$

But W_c^{α} is irreducible, hence[6]

$$_{ab}(JH)_{\alpha\beta} = \delta_{a\alpha}\ {}_b\lambda_\beta^\alpha I, \tag{11}$$

where ${}_b\lambda_\beta^\alpha$ is an ordinary number. Since JH is of rank ν, the submatrix $\left({}_b\lambda_\beta^\alpha\right)$ $(b = 1, 2, 3, \cdots, \mu_\alpha;\ \beta = 1, 2, 3, \cdots, n_\alpha)$ is of rank μ_α. Thus for every α there exists a set K_α of μ_α integers all $\leq n_\alpha$, such that the square matrix $\left({}_b\lambda_\beta^\alpha\right)$ $(b = 1, 2, \cdots, \mu_\alpha;\ \beta \text{ in } K_\alpha)$ is nonsingular. Hence JH may be divided into two untermatrices, one (of order $\nu \times \nu$) composed of those untermatrices (11) for which β is in the set k_α, and the other those for which β is not in K_α. The former is evidently nonsingular. Let H_2 and H_3 be the corresponding untermatrices of $H = J^{-1}(JH)$. Evidently a transposition of columns may bring H into the form $(H_2 \quad H_3)$. Now the constraint on Q is $HQ = 0$. Hence a corresponding transposition of the rows of Q bring it to $\begin{pmatrix} Q_2 \\ Q_3 \end{pmatrix}$ so that

$$H_2Q_2 + H_3Q_3 = 0.$$

Since H_2 is nonsingular, Q_3 may be chosen as the independent variable, Q_2 being dependent on it. Thus the latter can be dropped and the conditions (i), (ii) and (iii) are satisfied.

§5 CALCULATION OF SPUR ($\mathfrak{A}_c$)

Suppose that the constraints on the R's are given by

$$\mathscr{D}R = 0 \quad \text{where } \mathscr{D}\mathscr{D}' = I. \tag{12}$$

<u>THEOREM</u>

$$\mathrm{Spur}(\mathscr{A}_c) = \mathrm{Spur}(A_c)' - \mathrm{Spur}(\mathscr{D}A_c\mathscr{D}') \tag{13}$$

This theorem makes the calculation of $\mathrm{Spur}(\mathscr{A}_c)$ very simple (because the elements of A_c are 0 or 1), and before the transformation from R the Q is carried out. Also it enables us to calculate the contributions of the different constraints separately.

<u>PROOF</u> The constraints on $R_1, R_2, \cdots, R_n$ are $\mathscr{D}R = 0$. Hence we may take the matrix H of (7) to be $\mathscr{D}M'$. With the notation used there we have

$$\overset{c}{Q}_1 = H_1\overset{c}{Q} = H_1W_cQ = H_1W_cH_1'Q_1. \tag{14}$$

Since the $\mathscr{R}$'s are all independent, there exists M_1 such that $Q_1 = M_1\mathscr{R}$. Hence

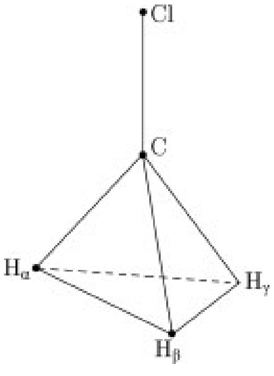
$$\overset{c}{Q}_1 = M_1\overset{c}{\mathscr{R}} = M_1\mathscr{A}_c\mathscr{R} = M_1\mathscr{A}_cM_1{}^{-1}Q_1.$$

On comparison with (14) it follows that $H_1W_cH_1' = M_1\mathscr{A}_cM_1{}^{-1}$. Thus

$$\begin{aligned}\mathrm{Spur}(\mathscr{A}_c) &= \mathrm{Spur}(H_1W_cH_1') \\ &= \mathrm{Spur}\left[\begin{pmatrix} H \\ H_1 \end{pmatrix} W_c\big(H'\ H_1'\big)\right] - \mathrm{Spur}(HW_cH') \\ &= \mathrm{Spur}(A_c) - \mathrm{Spur}(\mathscr{D}A_c\mathscr{D}').\end{aligned}$$

<u>EXAMPLE</u> Consider the molecule $\mathrm{CH_3Cl}$. Take

$$\begin{aligned} R_1, R_2, R_3, R_4 &= \text{increments of the distances } \overline{\mathrm{CCl}},\ \overline{\mathrm{CH}_\alpha},\ \overline{\mathrm{CH}_\beta},\ \overline{\mathrm{CH}_\gamma}, \\ R_5, R_6, R_7 &= \text{increments of the angles } \mathrm{Cl}\widehat{\mathrm{C}}\mathrm{H}_\alpha,\ \mathrm{Cl}\widehat{\mathrm{C}}\mathrm{H}_\beta,\ \mathrm{Cl}\widehat{\mathrm{C}}\mathrm{H}_\gamma, \\ R_8, R_9, R_{10} &= \text{increments of the angles } \mathrm{H}_\beta\widehat{\mathrm{C}}\mathrm{H}_\gamma,\ \mathrm{H}_\gamma\widehat{\mathrm{C}}\mathrm{H}_\alpha,\ \mathrm{H}_\alpha\widehat{\mathrm{C}}\mathrm{H}_\beta. \end{aligned}$$



The constraint is

$$[R_5 + R_6 + R_7]f + [R_8 + R_9 + R_{10}] = 0,$$

where f is a constant. Thus

$$\mathscr{D} = \sqrt{\frac{1}{3f^2+3}}(0\ 0\ 0\ 0\ f\ f\ f\ 1\ 1\ 1).$$

The theorem leads to

Operation C		Spur(A_c)	Spur$(\mathscr{D}A_c\mathscr{D})^*$	Spur$(\mathfrak{A}_c)$
No Motion	: C_1	10	1	9
Rotation about $\overline{\mathrm{CCl}}$ through	120° : C_2	1	1	0
	240° : C_3			
Reflection about the plane	Cl CH$_\alpha$: C_4	4	1	3
	Cl CH$_\beta$: C_5			
	Cl CH$_\gamma$: C_6			

The Kinetic and the Potential Energies

§6 CHOICE OF AXES

Suppose that $\begin{pmatrix} R_1 \\ R_2 \\ \vdots \\ R_{3N-6} \end{pmatrix} = R$ are the increments of $3N-6$ independent internal coordinates. For small vibrations, $R = B\mathscr{C}$, where $\mathscr{C} = \begin{pmatrix} x_1 \\ \vdots \\ z_N \end{pmatrix}$ is defined in §1. Suppose that the equilibrium positions of the nuclei are, in rectangular coordinates, $(X_1,\ Y_1,\ Z_1),(X_2,\ Y_2,\ Z_2),\cdots,(X_N,\ Y_N,\ Z_N)$. Let $m_1, m_2, \cdots, m_N$ be their masses. Write

$$F = \begin{pmatrix} m_1 & 0 & 0 & m_2 & 0 & 0 & \cdots & 0 \\ 0 & m_1 & 0 & 0 & m_2 & 0 & \cdots & 0 \\ 0 & 0 & m_1 & 0 & 0 & m_2 & \cdots & m_N \\ 0 & m_1Z_1 & -m_1Y_1 & 0 & m_2Z_2 & -m_2Y_2 & \cdots & -m_NY_N \\ -m_1Z_1 & 0 & m_1X_1 & -m_2Z_2 & 0 & m_2X_2 & \cdots & m_NX_N \\ m_1Y_1 & -m_1X_1 & 0 & m_2Y_2 & -m_2X_2 & 0 & \cdots & 0 \end{pmatrix}$$

---

* From this it is immediately seen that any constraint having the same coefficient for equivalent coordinates contributes 1 to Spur $(\mathscr{D}A_c\mathscr{D}')$.

Then the first column of $B\begin{pmatrix} \frac{1}{m_1} & & \\ & \ddots & \\ & & \frac{1}{m_N} \end{pmatrix} F'$ is the matrix $B\begin{pmatrix} 1 \\ 0 \\ 0 \\ 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$ which is the value of R when all the nuclei are displaced by 1 unit of length along the X-axis. But the R's are the increments of internal variables, hence $B\begin{pmatrix} 1 \\ 0 \\ 0 \\ 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix} = 0$. In the same way we can show that

$$B\begin{pmatrix} \frac{1}{m_1} & & \\ & \ddots & \\ & & \frac{1}{m_N} \end{pmatrix} F' = 0.$$

We have thus far described the the molecule in a certain system of rectangular coordinates. But when the molecule moves and rotates in space as well as vibrates, there remains an arbitrariness in fixing the coordinate axes to the molecule. The following method of choosing these axes is, however, the most preferable. The $3N-6$ internal variables $R_1, R_2, \cdots, R_{3N-6}$ is determined uniquely from the structure of the molecule. We define $x_1, y_1, \cdots, z_N$ by

$$\mathscr{C} = \begin{pmatrix} x_1 \\ \vdots \\ z_N \end{pmatrix} = \begin{pmatrix} B \\ F \end{pmatrix}^{-1} \begin{pmatrix} R \\ 0 \end{pmatrix}, \tag{15}$$

so that

$$F\mathscr{C} = 0. \tag{16}$$

From the definition of B, we conclude that a system of axes of reference can be found so that the positions of the nuclei are $(X_1+x_1, Y_1+y_1, Z_1+z_1)$, $(X_2+x_2, Y_2+y_2, Z_2+z_2)$, $\cdots$, $(X_N+x_N, Y_N+y_N, Z_N+z_N)$ when the molecule is not very much distorted from its equilibrium structure. We have thus 6 external variables specifying the position and orientation of the axes in space and $3N-6$ variables $R_1, R_2, \cdots, R_{3N-6}$ specifying by means of (15) the positions of the nuclei with respect to these axes.

§7 THE KINETIC ENERGY IN TERMS OF *R*

Equ. (16) expresses the facts that the origin of our moving axes is at any time the centre of mass of the molecule, and that the moment of momentum of the molecule in this system of reference is of the order of $mx\dot{x}$. Thus if the molecule rotates in space with an angular velocity ω about its centre of mass, its kinetic energy is

$$T_{\text{centre of mass}} + T_{\text{rotation}} + \frac{1}{2}\sum_i m_i(\dot{x}_i^2+\dot{y}_i^2+\dot{z}_i^2) + \text{term} \sim mx\dot{x}\omega.$$

But

$$T_{\text{rotation}} = (\text{K.E. of rotation if } \mathscr{C}=0) + \text{term} \sim m\omega^2 xX.$$

Now $T_{\text{centre of mass}} +$ (K.E. of rotation if $\mathscr{C}=0$) depends on the external variables only (together, of course, with their time derivatives). And in a gas, due to thermal agitation, $\omega X \sim \dot{x}$.

Thus in the first approximation the equ. of motion is to be derived from (cf. §§15, 16)

$$T = \frac{1}{2}\sum_i m_i(\dot{x}_i^2+\dot{y}_i^2+\dot{z}_i^2) = \frac{1}{2}(\dot{R}'\ 0)\begin{pmatrix} B \\ F \end{pmatrix}^{-1\prime}\begin{pmatrix} m_1 & & \\ & \ddots & \\ & & m_N \end{pmatrix}\begin{pmatrix} B \\ F \end{pmatrix}^{-1}\begin{pmatrix} \dot{R} \\ 0 \end{pmatrix}.$$

But

$$\begin{pmatrix} B \\ F \end{pmatrix}\begin{pmatrix} \frac{1}{m_1} & & \\ & \ddots & \\ & & \frac{1}{m_N} \end{pmatrix}\begin{pmatrix} B' & F' \end{pmatrix} = \begin{pmatrix} B\begin{pmatrix} \frac{1}{m_1} & & \\ & \ddots & \\ & & \frac{1}{m_N} \end{pmatrix}B' & 0 \\ 0 & F\begin{pmatrix} \frac{1}{m_1} & & \\ & \ddots & \\ & & \frac{1}{m_N} \end{pmatrix}F' \end{pmatrix}.$$

Hence writing

$$G = B\begin{pmatrix} \frac{1}{m_1} & & \\ & \ddots & \\ & & \frac{1}{m_N} \end{pmatrix}B', \tag{17}$$

we have

$$2T = \dot{R}'G^{-1}\dot{R}. \tag{18}$$

§8 THE KINETIC ENERGY IN TERMS OF $\mathscr{R}$

If we take the independent reduced coordinates $\mathscr{R}$ (§4) to be the R's of the last section, the results may be summarized:

$$\mathscr{R} = \mathscr{L}\mathscr{C}, \qquad \mathscr{G} = \mathscr{L}\begin{pmatrix}\frac{1}{m_1} & & \\ & \ddots & \\ & & \frac{1}{m_N}\end{pmatrix}\mathscr{L}', \qquad 2T = \dot{\mathscr{R}}'\mathscr{G}^{-1}\dot{\mathscr{R}}. \tag{19}$$

Now by (5),

$$\mathscr{L}P_cZ_c = \mathfrak{A}_c\mathscr{L} \qquad (\mathfrak{A}_c \text{ stands here for } A_c, \text{ cf. §4})$$

so that

$$\mathscr{G} = \mathscr{L}\begin{pmatrix}\frac{1}{m_1} & & \\ & \ddots & \\ & & \frac{1}{m_N}\end{pmatrix}\mathscr{L}' = \mathfrak{A}_c'\mathscr{L}P_cZ_c\begin{pmatrix}\frac{1}{m_1} & & \\ & \ddots & \\ & & \frac{1}{m_N}\end{pmatrix}Z_c'P_c'\mathscr{L}'\mathfrak{A}_c.$$

But $\begin{pmatrix}\frac{1}{m_1} & & \\ & \ddots & \\ & & \frac{1}{m_N}\end{pmatrix}$, P_c and Z_c commute[*] with each other, hence

$$\mathscr{G} = \mathfrak{A}_c'\mathscr{L}\begin{pmatrix}\frac{1}{m_1} & & \\ & \ddots & \\ & & \frac{1}{m_N}\end{pmatrix}\mathscr{L}'\mathfrak{A}_c = \mathfrak{A}_c'\mathscr{G}\mathfrak{A}_c.$$

Suppose

$${}_{\alpha\beta}(\mathfrak{A}_c)_{\alpha'\beta'} = \delta_{\alpha\alpha'}\delta_{\beta\beta'}W_c^{\alpha}, \quad \alpha = 1, 2, \cdots, k, \; \beta = 1, 2, \cdots, n_\alpha, \tag{20}$$

where W_c^α is irreducible. (cf. §4) Since $\mathfrak{A}_c\mathscr{G} = \mathscr{G}\mathfrak{A}_c$, we have

$$W_c^\alpha \; {}_{\alpha\beta}(\mathscr{G})_{\alpha'\beta'} = {}_{\alpha\beta}(\mathscr{G})_{\alpha'\beta'} \; W_c^{\alpha'}.$$

Hence[6]

$${}_{\alpha\beta}(\mathscr{G})_{\alpha'\beta'} = \delta_{\alpha\alpha'} \; {}_{\beta}g_{\beta'}^{\alpha} I. \tag{21}$$

If W_c^α is of order d_α, this shows that the nonvanishing elements of $\mathscr{G}$ are in $\sum\limits_{\alpha=1}^{k} d_\alpha$ diagonal blocks of which d_α are identical of dimension $n_\alpha \times n_\alpha$.

---

[*] cf. §18.

§9 THE POTENTIAL ENERGY IN TERMS OF $\mathscr{R}$

The potential energy depends on the internal coordinates only. For small vibrations, it is approximately equal to $\frac{1}{2}\mathscr{R}'\mathfrak{V}\mathscr{R}$, where $\mathfrak{V}$ is a positive symmetrical matrix, because when in equilibrium the molecule has a minimum potential energy. The covering operations leave the potential energy unchanged:

$$\frac{1}{2}\mathscr{R}'\mathfrak{V}\mathscr{R} = \frac{1}{2}\overset{c}{\mathscr{R}}'\mathfrak{V}\overset{c}{\mathscr{R}} = \frac{1}{2}\mathscr{R}'\mathfrak{A}'_c\mathscr{R}.$$

But $\mathscr{R}$ is arbitrary, hence

$$\mathfrak{V} = \mathfrak{A}'_c\mathfrak{V}\mathfrak{A}_c.$$

Thus we have, just as we had (21),

$$ {}_{\alpha\beta}(\mathfrak{V})_{\alpha'\beta'} = \delta_{\alpha\alpha'}\ {}_{\beta}v_{\beta'}{}^{\alpha}I, \quad \alpha = 1, 2, \cdots, k, \quad \beta = 1, 2, \cdots, n_\alpha. \tag{22}$$

There are therefore totally $\sum_{\alpha=1}^{k} n_a^2$ independent constants in the potential energy, which are usually unknowns.[11]

Calculation of the Kinetic Energy

§10 THE TRANSFORMATION MATRIX M[7]

In §4 the orthogonal matrix M was introduced to reduce A_c. Since

$$A_c = \begin{pmatrix} A_{c1} & & 0 \\ & A_{c2} & \\ 0 & & \ddots \end{pmatrix}$$

where each A_{cp} belongs to a set of equivalent coordinates, (§4)

$$M = \begin{pmatrix} M_1 & & 0 \\ & M_2 & \\ 0 & & \ddots \end{pmatrix}$$

where M_p is orthogonal. Hence $W_{cp} = M_p A_{cp} M_p^{-1}$ is of the form (6), i.e.

$$W_{cp} = \begin{pmatrix} W_{cp}^1 & & & \\ & W_{cp}^1 & & \\ & & \ddots & \\ & & & W_{cp}^k \end{pmatrix}.$$

Suppose that W_{cp}^1 is of dimension $d \times d$. Denote the first d rows of M_p by

$$L = \begin{pmatrix} {}_1(M_p)_1 & {}_1(M_p)_2 & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots \\ {}_d(M_p)_1 & \cdots & \cdots & {}_d(M_p)_{g_p} \end{pmatrix},$$

where $g_p =$ no. of equivalent coordinates R of the p-th set.

Now

$$W_{cp} M_p = M_p A_{cp}.$$

Hence

$$W_{cp}^1 L = L A_{cp}.$$

Thus

$$L'L = A'_{cp} L' L A_{cp}, \quad \text{i.e.} \quad (L'L)A_{cp} = A_{cp}(L'L).$$

Hence

$$\sum_{t=1}^{g_p} {}_1(L'L)_t \, {}_t(A_{cp})_2 = \sum_{t=1}^{g_p} {}_1(A_{cp})_t \, {}_t(L'L)_2.$$

There exists an operation C which brings the second coordinate of the p-th set to the first, for which $\overset{c}{R}_1 = R_2$, so that ${}_1(A_{cp})_t = \delta_{t2}$, ${}_t(A_{cp})_2 = \delta_{t1}$. Hence ${}_1(L'L)_1 = {}_2(L'L)_2$. Thus all diagonal elements of $L'L$ are equal, and

$$\sum_{\gamma=1}^{d} {}_\gamma(M_p)_t^2 = {}_t(L'L)_t = \frac{1}{g_p}\text{Spur}(L'L) = \frac{1}{g_p}\text{Spur}(LL') = \frac{d}{g_p}.$$

Now d can be determined from the values of Spur $(\mathfrak{A}_c)$, (cf. equ. (39)) so that this relation facilitates the calculation of M.

§11 VECTORIAL NOTATION [7]

The matrix $\mathscr{G} = \mathscr{L} \begin{pmatrix} \frac{1}{m_1} & & \\ & \ddots & \\ & & \frac{1}{m_N} \end{pmatrix} \mathscr{L}'$ is calculated by first computing $\mathscr{L}$. Now $\mathscr{R}$ is a submatrix of Q. Hence

$$\mathscr{L} = M_1 B, \qquad \text{where } M_1 = {}_{\text{some rows}}(M)_{\text{all columns}}. \tag{23}$$

Thus we have to find B first.

Now

$$R = B\mathscr{C}.$$

Write

$$(x_t \quad y_t \quad z_t) = \overrightarrow{\mathscr{C}_t}, \qquad ({}_kB_{tx} \quad {}_kB_{ty} \quad {}_kB_{tz}) = \overrightarrow{{}_kS_t}.$$

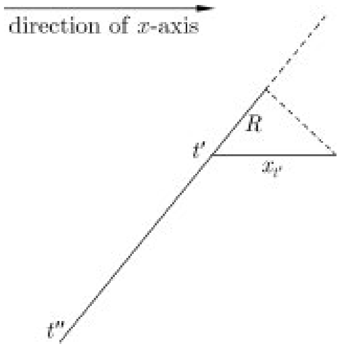
Then $R_k = \sum\limits_t \overrightarrow{{}_kS_t} \cdot \overrightarrow{\mathscr{C}_t}$. In a similar way we shall write

$$
{}_{\alpha\beta\gamma}(\mathscr{L})_t = \sum_{k=1}^{n} {}_{\alpha\beta\gamma}(M_1)_k \, \overrightarrow{{}_kS_t} = {}_{\alpha\beta\gamma}\overrightarrow{\mathbb{S}_t}, \quad \gamma = 1, 2, \cdots, d_\alpha, \quad t = 1, 2, \cdots, N. \quad \text{(cf. (21))} \tag{24}
$$

§12 EXPLICIT EXPRESSION OF $\overrightarrow{{}_kS_t}$ [7]

The use of the vectors $\overrightarrow{S}$ is advantageous because they are independent of the coordinate axes. Two kinds of R's are commonly used:

(i) $R =$ the increment of the bond between the nuclei t' and t'' in length. Evidently when $\overrightarrow{\mathscr{C}_{t'}} = \overrightarrow{\mathscr{C}_{t''}} = 0$, $R = 0$. Hence



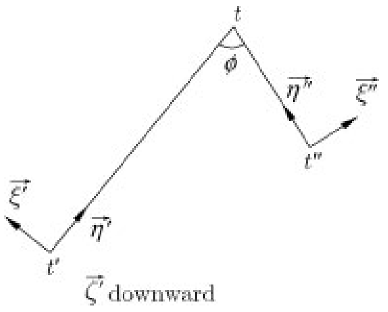
$$\overrightarrow{S_t} = 0 \qquad \text{for all } t \text{ except } t' \text{ or } t''.$$

If $0 = y_{t'} = z_{t'} = x_{t''} = y_{t''} = z_{t''}$, we have $R = x_{t'} \cos\theta$, so that $B_{t'x} = \cos\theta$. Hence

$$
\begin{cases}
\overrightarrow{S_{t'}} = \text{unit vector from } t'' \text{ to } t'. \\
\overrightarrow{S_{t''}} = -\overrightarrow{S_{t'}}. \\
\text{all other } \overrightarrow{S}_t \text{ are } 0.
\end{cases} \tag{25}
$$

(ii) $R =$ the increment of the angle $t' - t - t''$. Evidently all $\overrightarrow{S}$ are 0 except $\overrightarrow{S_t}$, $\overrightarrow{S_{t'}}$ and $\overrightarrow{S_{t''}}$. Suppose that

$$\overrightarrow{S_{t'}} = x'\overrightarrow{\xi'} + y'\overrightarrow{\eta'} + z'\overrightarrow{\zeta'}.$$



When $\overrightarrow{\mathscr{C}_t} = \overrightarrow{\mathscr{C}_{t''}} = 0$, $\overrightarrow{\mathscr{C}_{t'}} = \overrightarrow{\xi'}$, we have, since t' alone is displaced, $R = \frac{1}{\epsilon}$. Hence $x' = \frac{1}{\epsilon}$. Similarly by considering the case when $\overrightarrow{\mathscr{C}_t} = \overrightarrow{\mathscr{C}_{t''}} = 0$, $\overrightarrow{\mathscr{C}_{t'}} = \overrightarrow{\eta'}$, we get $y' = 0$. Proceeding in this way we arrive finally at the first line of (26). Now R vanishes when $\overrightarrow{\mathscr{C}_{t'}} = \overrightarrow{\mathscr{C}_t} = \overrightarrow{\mathscr{C}_{t''}}$. Thus $(\overrightarrow{S_{t'}} + \overrightarrow{S_{t''}} + \overrightarrow{S_t}) = 0$. So that

$$
\begin{cases}
\overrightarrow{S_{t'}} = \dfrac{1}{\epsilon}\overrightarrow{\xi'}, \quad \overrightarrow{S_{t''}} = \dfrac{1}{\gamma}\overrightarrow{\xi''} \\
\overrightarrow{S_t} = -(\overrightarrow{S_{t'}} + \overrightarrow{S_{t''}}) \\
\text{all other } \overrightarrow{S} = 0
\end{cases} \tag{26}
$$

From (25) and (26) we can write down, for all k and t, the vector $\overrightarrow{{}_kS_t}$ from which $\overrightarrow{\mathbb{S}}$ may be calculated through (24).

§13 KINETIC ENERGY IN TERMS OF THE $\overrightarrow{\mathbb{S}}$'S [7]

We shall show however, that not all the $\overrightarrow{\mathbb{S}}$'s are needed for the calculation of $\mathscr{G}$.

Take (5)

$$\boldsymbol{\mathscr{L}} P_c Z_c = \mathfrak{A}_c \mathscr{L}. \tag{27}$$

By (20)

$$\sum_{t'=1}^{N} {}_{\alpha\beta}\mathscr{L}_{t'} \Gamma_c \delta_{t,c^{-1}t'} = W_c^{\alpha}\, {}_{\alpha\beta}\mathscr{L}_t, \quad (\Gamma_c \text{ is defined in §2})$$

i.e.

$$ {}_{\alpha\beta}\mathscr{L}_{ct}\Gamma_c = W_c^{\alpha}\, {}_{\alpha\beta}\mathscr{L}_t.$$

Hence for any β and β',

$$\Gamma_c'\, {}_{\alpha\beta}\mathscr{L}_{ct}'\, {}_{\alpha\beta'}\mathscr{L}_{ct}\, \Gamma_c = {}_{\alpha\beta}\mathscr{L}_t'\, {}_{\alpha\beta'}\mathscr{L}_t,$$

so that

$$\mathrm{Spur}({}_{\alpha\beta}\mathscr{L}_{ct}'\, {}_{\alpha\beta'}\mathscr{L}_{ct}) = \mathrm{Spur}({}_{\alpha\beta}\mathscr{L}_t'\, {}_{\alpha\beta'}\mathscr{L}_t). \tag{28}$$

Now by (21),

$$ {}_{\beta}g_{\beta'}^{\alpha} I = {}_{\alpha\beta}\mathscr{G}_{\alpha\beta'} = \sum_p {}_{\alpha\beta}\mathscr{L}_p \left(\frac{1}{m_p} I\right)_{\alpha\beta'} \mathscr{L}_p', \tag{29}$$

where each p refer to a set of equivalent nuclei.

From (27)

$$ {}_{\alpha\beta}(\mathscr{L})_p\, {}_p(P_c)_p\, {}_p(Z_c)_p = W_c^{\alpha}\, {}_{\alpha\beta}(\mathscr{L})_p.$$

Hence

$$ {}_{\alpha\beta}(\mathscr{L})_p\, {}_{\alpha\beta'}(\mathscr{L})_p' = W_c^{\alpha}\, {}_{\alpha\beta}\mathscr{L}_p\, {}_{\alpha\beta'}\mathscr{L}_p',$$

i.e.

$$({}_{\alpha\beta}\mathscr{L}_p\, {}_{\alpha\beta'}\mathscr{L}_p') W_c^{\alpha} = W_c^{\alpha}({}_{\alpha\beta}\mathscr{L}_p\, {}_{\alpha\beta'}\mathscr{L}_p').$$

Thus[6]

$$ {}_{\alpha\beta}\mathscr{L}_p\, {}_{\alpha\beta'}\mathscr{L}_p' = {}_{\beta}\boldsymbol{\mathscr{b}}_{\beta'}^{\alpha p} I, \tag{30}$$

where ${}_{\beta}\boldsymbol{\mathscr{b}}_{\beta'}^{\alpha p}$ is an ordinary number. From (29) we have,

$$ {}_{\beta}g_{\beta'}^{\alpha} I = \sum_p \frac{1}{m_p}\, {}_{\beta}\boldsymbol{\mathscr{b}}_{\beta'}^{\alpha p} I. \tag{31}$$

Hence

$$ {}_{\beta}\boldsymbol{\mathscr{b}}_{\beta'}^{\alpha p} d_\alpha = \mathrm{Spur}({}_{\alpha\beta}\mathscr{L}_p\, {}_{\alpha\beta'}\mathscr{L}_p') = \mathrm{Spur}({}_{\alpha\beta'}\mathscr{L}_p'\, {}_{\alpha\beta}\mathscr{L}_p) = \sum_{t \text{ in } p} \mathrm{Spur}({}_{\alpha\beta'}\mathscr{L}_t'\, {}_{\alpha\beta}\mathscr{L}_t). \tag{32}$$

Now consider the sum of (28). Since ct goes over all values of t in p when c goes over all the operations of the group,

$$\mathrm{Spur}({}_{\alpha\beta}\mathscr{L}'_{ct}\ {}_{\alpha\beta'}\mathscr{L}_{ct}) = \mathrm{Spur}({}_{\alpha\beta}\mathscr{L}'_{ct'}\ {}_{\alpha\beta'}\mathscr{L}_{ct'}),$$

if t and t' are in the same p-th set. Thus (32) becomes

$${}_{\beta}\boldsymbol{b}^{\alpha p}_{\beta'}\, d_\alpha = g_p \mathrm{Spur}({}_{\alpha\beta}\mathscr{L}_t\ {}_{\alpha\beta'}\mathscr{L}'_t),$$

where t represents any nucleus of the p-th set. Substitute into (31) and make use of (24). We obtain:

$${}_{\beta}g^{\alpha}_{\beta'} = \frac{1}{d_\alpha}\sum_p \frac{g_p}{m_p}\sum_{\gamma=1}^{d_\alpha} {}_{\alpha\beta\gamma}\overrightarrow{\mathbb{S}_t} \cdot {}_{\alpha\beta'\gamma}\overrightarrow{\mathbb{S}_{t'}}. \tag{33}$$

Thus in calculation $\mathscr{G}$ (${}_{\alpha\beta}\mathscr{G}_{\alpha'\beta'} = \delta_{\alpha\alpha'}\ {}_{\beta}g^{\alpha}_{\beta'}I$) we have only to know ${}_{\alpha\beta\gamma}\overrightarrow{\mathbb{S}_t}$ for all α ($= 1, 2, \cdots, k$), β ($= 1, 2, \cdots, n_d$) and γ ($= 1, 2, \cdots, d_\alpha$); but only for $\underline{\text{one}}$ t in each p. From (24) we see that not all $\overrightarrow{S}$ are needed.

Secular Equation and Degeneracy

§14 NORMAL COORDINATES

In terms of the independent reduced coordinates $\mathscr{R}$ defined in §4, the kinetic energy and the potential energy are $\frac{1}{2}\dot{\mathscr{R}}'\mathscr{G}^{-1}\dot{\mathscr{R}}$ and $\frac{1}{2}\mathscr{R}'\mathfrak{V}\mathscr{R}$, (§§8, 9) where $\mathscr{G}$ is positive definite. Let $\lambda_1, \lambda_2, \cdots, \lambda_{3N-6}$ be the roots of the equ.

$$\left|\lambda\mathscr{G}^{-1} - \mathfrak{V}\right| = 0, \quad \text{i.e.} \quad |\lambda I - \mathfrak{V}\mathscr{G}| = 0. \tag{34}$$

Then there exists a matrix L such that

$$L\mathscr{G}L' = I, \qquad L^{-1\prime}\mathfrak{V}L^{-1} = \begin{pmatrix} \lambda_1 & & & 0 \\ & \lambda_2 & & \\ & & \ddots & \\ 0 & & & \lambda_{3N-6} \end{pmatrix} = \varLambda.$$

Put

$$\mathfrak{N} = \begin{pmatrix} \mathfrak{N}_1 \\ \vdots \\ \mathfrak{N}_{3N-6} \end{pmatrix} = L\mathscr{R},$$

we get immediately

$$2\times\text{kinetic energy} = \dot{\mathscr{R}}'\mathscr{G}^{-1}\dot{\mathscr{R}} = \dot{\mathscr{N}}'\mathscr{N} = \sum_{i=1}^{3N-6} \dot{\mathscr{N}}_i^2, \tag{35}$$

$$2\times\text{potential energy} = \mathscr{R}'\mathscr{V}\mathscr{R} = \mathscr{N}'\Lambda\mathscr{N} = \sum_i \lambda_i \mathscr{N}_i^2. \tag{36}$$

These coordinates $\mathscr{N}_i$ are known as the "normal coordinate" 's. They are obtained by first solving (34), thus getting λ_i; and then determining L from

$$L\mathscr{G}\mathscr{V} = \Lambda L. \tag{37}$$

This L must be normalized by

$$L\mathscr{G}L' = I. \tag{38}$$

The normal coordinates are then calculated from $\mathscr{N} = L\mathscr{R}$.

Now we shall be able to see the role that symmetry plays in simplifying the calculations. Equs. (21) and (22) show the secular equ. (34) is factored into $\sum_{\alpha=1}^{k} d_\alpha$ equations of which d_α are identical and are of the n_α-th degree. The labor of solving for the λ's is considerably saved. Moreover, the no. of unknown constants in $\mathscr{V}$ is reduced.

The normal coordinates are also a special form of the "symmetry coordinates" introduced by Howard and Wilson.[21] The most general form of the symmetry coordinates $\mathscr{M}$ is given by $\mathscr{N} = \mathscr{L}\mathscr{M}$ where

$$_{\alpha\beta}\mathscr{L}_{\alpha'\beta'} = \delta_{\alpha\alpha'}\, {}_{\beta}\ell^{\alpha}_{\beta'} U^{\alpha}, \quad \alpha = 1, 2, \cdots, k, \ \beta = 1, 2, \cdots, n_a,$$

in which U^α is orthogonal and ${}_{\beta}\ell^{\alpha}_{\beta'}$ form an orthogonal matrix when β and β' range over $1, 2, \cdots, n_\alpha$.

§15 SOLUTION OF THE PHYSICAL PROBLEM—CLASSICAL THEORY

The Lagrangian is, from (35) and (36),

$$\frac{1}{2}\sum_{i=1}^{3N-6}(\dot{\mathscr{N}}_i^2 - \lambda_i \mathscr{N}_i^2).$$

The equs. of motion are therefore

$$\ddot{\mathscr{N}}_i + \lambda_i \mathscr{N}_i = 0, \quad i = 1, 2, \cdots, 3N-6.$$

Hence

$$\mathscr{N}_i = (\mathscr{N}_i)_0 \cos(\sqrt{\lambda_i}t + \phi_i).$$

The frequencies of vibration are thus

$$\frac{\sqrt{\lambda_1}}{2\pi}, \frac{\sqrt{\lambda_2}}{2\pi}, \cdots, \frac{\sqrt{\lambda_{3N-6}}}{2\pi}.$$

This is correct only in the first approximation; but it serves to give almost all our present knowledge about the forces within the molecules.

§16 SOLUTION OF THE PHYSICAL PROBLEM—QUANTUM THEORY

Now

$$T = \frac{1}{2}\sum_i \dot{\mathfrak{N}}_i^2, \quad V = \frac{1}{2}\sum_i \lambda_i \mathfrak{N}_i^2.$$

The wave equation is most easily obtained from the variational formulation of the problem:

$$\delta \int \left\{ \sum \hbar^2 g^{ij} \frac{\partial \psi^*}{\partial \mathfrak{N}_i} \frac{\partial \psi}{\partial \mathfrak{N}_j} + (V - E)\psi^* \psi \right\} \sqrt{g}\, \mathrm{d}\mathfrak{N}_1 \cdots \mathrm{d}\mathfrak{N}_{3N-6} = 0,$$

where g^{ij} is given by $T = g_{ij}\dot{x}^i \dot{x}^j$. Hence

$$\sum_i \frac{1}{2}\left(-\hbar^2 \frac{\partial^2}{\partial \mathfrak{N}_i^2} + \lambda_i \mathfrak{N}_i^2\right)\psi = E\psi.$$

The electric moment in any direction is, in the first approximation

$$D_0 + \sum_{i=1}^{3N-6} D_i \mathfrak{N}_i$$

where D_i is immediately calculable from (15). Thus the wave function and selection rules, hence the frequencies in the vibrational spectra of the molecule, are the same as those of $3N-6$ independent harmonic oscillators with coordinates $\mathfrak{N}_1, \mathfrak{N}_2, \cdots, \mathfrak{N}_{3N-6}$. The frequencies are therefore exactly those obtained classically.

§17 THE DEGREE OF DEGENERACY

We have seen that the secular equ. (34) is factorized into d_1 identical equs. of the n_1-th degree, d_2 identical equs. of the n_2-th degree, $\cdots$. There must therefore be n_1 frequencies each corresponding to d_1 different normal moods of vibration (d_1-fold degenerate). The nos. d_1, d_2, $\cdots$, are accordingly called “the degrees of degeneracy”.

It is interesting that they can be determined together with n_1, n_2, $\cdots$, without carrying out the calculation of the normal coordinates. For, by (20)

$$\sum_\alpha n_\alpha \mathrm{Spur}(W_c^\alpha) = \mathrm{Spur}(\mathfrak{A}_c).$$

But

$$\sum_c \mathrm{Spur}(W_c^\alpha)\mathrm{Spur}(W_c^\beta) = \delta_{\alpha\beta} h,$$

where h = order of the group. Hence

$$n_\alpha = \frac{1}{h}\sum_c \mathrm{Spur}(\mathfrak{A}_c)\mathrm{Spur}(W_c^\alpha). \tag{39}$$

Now both $\mathrm{Spur}(W_c^\alpha)$ and d_α can be found^3 from a table of characters of the point group, and $\mathrm{Spur}(\mathfrak{A}_c)$ can be calculated through the theorem in §5, so that d_α and n_α are easily calculable from (39).

EXAMPLE We consider again the molecule CH_3Cl. (§5) From the left table of characters we get $n_1 = 3$, $n_2 = 0$, $n_3 = 3$.It follows that the secular equ. is factored into 3 cubic equs. 2 of which are identical.

Representation	d_α	$\mathrm{Spur}(W_c)$		
		$c = c_1$	$c = c_2, c_3$	$c = c_4, c_5, c_6$
W^1	1	1	1	1
W^2	1	1	1	-1
W^3	2	2	-1	0

We conclude that there are totally 6 vibrational frequencies, 3 of which are doubly degenerate.

The no. of independent constants in the potential energy is (§9) $\sum_\alpha n_a^2 = 18$.

Other Developments

§18 ISOTOPE RULE [7]

In the above example the no. of unknown constants in the potential energy is much greater than the no. of observable frequencies, as is usually the case. The situation is, however, not so hopeless, because the frequencies of the molecule with some of the atoms replaced by their isotopes serve as additional data. The new molecule has evidently the same equilibrium configuration and potential energy as the original one. But in calculating its kinetic energy it must be remembered that all the results of §§1 to 7 are correct, (because they do not concern the masses of the nuclei) but that since now $\begin{pmatrix} \frac{1}{m_1} & & \\ & \ddots & \\ & & \frac{1}{m_N} \end{pmatrix}$ does not necessarily commute with P_c, (21) does not hold. The secular equ. is therefore not factored as before. The vectorial method is still convenient.

If the new kinetic energy matrix is denoted by $\mathscr{G}'$, from (34)

$$\Pi\lambda_i = |\mathfrak{V}\mathscr{G}| = |\mathfrak{V}||\mathscr{G}|,$$
$$\Pi\lambda_i' = |\mathfrak{V}\mathscr{G}'| = |\mathfrak{V}||\mathscr{G}'|.$$

Hence

$$\frac{\Pi\lambda_i}{\Pi{\lambda_i}'} = \frac{|\mathscr{G}|}{|\mathscr{G}'|} = \frac{\left|\mathscr{L}\begin{pmatrix} \frac{1}{m_1} & & \\ & \ddots & \\ & & \frac{1}{m_N} \end{pmatrix}\mathscr{L}'\right|}{\left|\mathscr{L}\begin{pmatrix} \frac{1}{m_1}' & & \\ & \ddots & \\ & & \frac{1}{m_N}' \end{pmatrix}\mathscr{L}'\right|}.$$

§19 SPLITTING OF HIGH FREQUENCIES [7]

Sometimes it is known that certain force constants are considerably larger than the others, so that the corresponding frequencies of vibration are much greater than the rest. To solve for the low frequencies we may make the approximation of putting the large force constants equal to infinity.

Let

$$\mathscr{G} = \begin{pmatrix} \mathscr{G}_{11} & \mathscr{G}_{12} \\ \mathscr{G}_{21} & \mathscr{G}_{22} \end{pmatrix}, \qquad \mathfrak{V} = \begin{pmatrix} \mathfrak{V}_{11} & \mathfrak{V}_{12} \\ \mathfrak{V}_{21} & \mathfrak{V}_{22} \end{pmatrix}.$$

We want to find the λ's when $\mathfrak{V}_{11} \to \infty$. Now

$$\left|\lambda\mathscr{G}^{-1}\mathfrak{V}^{-1} - I\right| = 0,$$

and when $\mathfrak{V}_{11} \to \infty$, $\mathfrak{V}^{-1} \to \begin{pmatrix} 0 & 0 \\ 0 & \mathfrak{V}_{22}^{-1} \end{pmatrix}$. Since

$$\mathscr{G}^{-1} = \begin{pmatrix} (\mathscr{G}^{-1})_{11} & (\mathscr{G}^{-1})_{12} \\ (\mathscr{G}^{-1})_{21} & (\mathscr{G}_{22} - \mathscr{G}_{21}\mathscr{G}_{11}^{-1}\mathscr{G}_{12})^{-1} \end{pmatrix},$$

we get

$$\left|\lambda\begin{pmatrix} 0 & (\mathscr{G}^{-1})_{12}\mathfrak{V}_{22}^{-1} \\ 0 & (\mathscr{G}_{22} - \mathscr{G}_{21}\mathscr{G}_{11}^{-1}\mathscr{G}_{12})^{-1}\mathfrak{V}_{22}^{-1} \end{pmatrix} - I\right| = 0,$$

or

$$\left|\lambda(\mathscr{G}_{22} - \mathscr{G}_{21}\mathscr{G}_{11}^{-1}\mathscr{G}_{12})^{-1}\mathfrak{V}_{22}^{-1} - I\right| = 0,$$

i.e.

$$\left|\lambda(\mathscr{G}_{22} - \mathscr{G}_{21}\mathscr{G}_{11}^{-1}\mathscr{G}_{12})^{-1} - \mathfrak{V}_{22}\right| = 0.$$

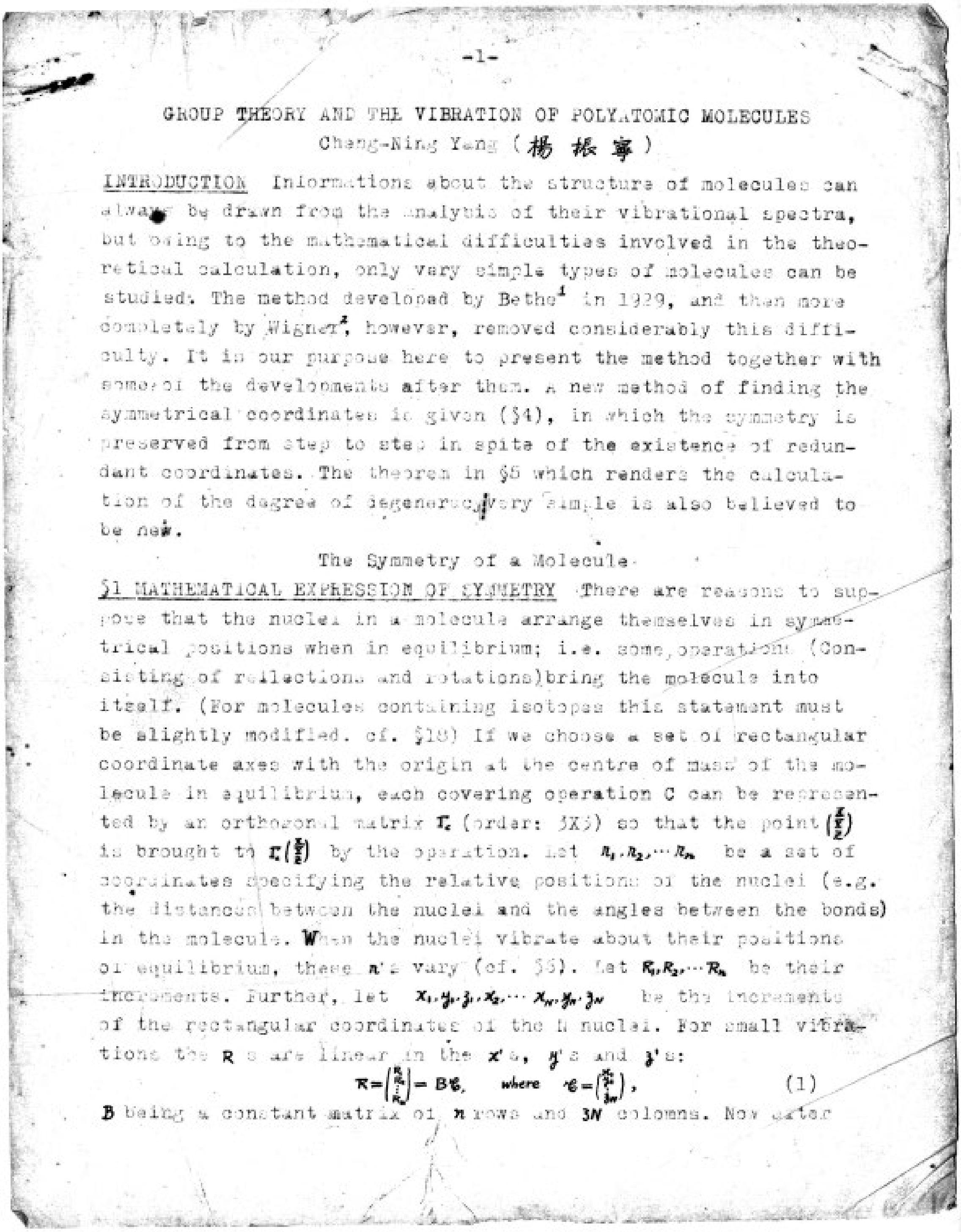
Acknowledgement

The author wishes to express his thanks to Prof. T. Y. Wu for his continual guidance.

++++++++++++++

References

1. von H. Bethe, Ann. d. Physik, 3 (1929), 133.
2. E. Wigner, Göttingen Nachrichton, (1930), p.133.
3. See 1. and 2., also Speiser: Theorie der Gruppen von Endlicher Ordnung (Berlin, 1927).
   and J. E. Rosenthal and G. M. Murphy, Rev. Mod. Phys., 8 (1936), 317.
   The degeneracy and other properties are tabulated by E. B. Wilson Jr. in, J. Chem. Phys., 2 (1934), 432. (complete to molecules containing 7 atoms)
   The irreducible representations are given in F. Seitz, Zeits. f. Krist., A88 (1934), 433.
4. E. Wigner, Gruppentheorie. (Braunschweig 1931) This book is always meant when Wigner is referred to.
5. Wigner, p.86.
6. ibid. p.83.
7. E. B. Wilson Jr., J. Chem. Phys., 9 (1941), 76.
8. J. E. Rosenthal and G. M. Murphy, Rev. Mod. Phys., 8 (1936), 317.
9. E. B. Wilson Jr. and B. L. Crawford Jr., J. Chem. Phys., 8 (1936), 223.
10. D. M. Dennison, Rev. Mod. Phys., 3 (1931), 280.
11. J. B. Howard and E. B. Wilson, Jr., J. Chem. Phys., 2 (1934), 630.



-1-

GROUP THEORY AND THE VIBRATION OF POLYATOMIC MOLECULES

Cheng-Ning Yang (楊振寧)

INTRODUCTION Informations about the structure of molecules can always be drawn from the analysis of their vibrational spectra, but owing to the mathematical difficulties involved in the theoretical calculation, only very simple types of molecules can be studied. The method developed by Bethe[1] in 1929, and then more completely by Wigner[2], however, removed considerably this difficulty. It is our purpose here to present the method together with some of the developments after them. A new method of finding the symmetrical coordinates is given (§4), in which the symmetry is preserved from step to step in spite of the existence of redundant coordinates. The theorem in §5 which renders the calculation of the degree of degeneracy very simple is also believed to be new.

The Symmetry of a Molecule

§1 MATHEMATICAL EXPRESSION OF SYMMETRY There are reasons to suppose that the nuclei in a molecule arrange themselves in symmetrical positions when in equilibrium; i.e. some operations (Consisting of reflections and rotations) bring the molecule into itself. (For molecules containing isotopes this statement must be slightly modified. cf. §18) If we choose a set of rectangular coordinate axes with the origin at the centre of mass of the molecule in equilibrium, each covering operation C can be represented by an orthogonal matrix T_c (order: 3X3) so that the point $\begin{pmatrix}x\\y\\z\end{pmatrix}$ is brought to $T_c\begin{pmatrix}x\\y\\z\end{pmatrix}$ by the operation. Let $r_1, r_2, \cdots r_n$ be a set of coordinates specifying the relative positions of the nuclei (e.g. the distances between the nuclei and the angles between the bonds) in the molecule. When the nuclei vibrate about their positions of equilibrium, these r's vary (cf. §6). Let $R_1, R_2, \cdots R_n$ be their increments. Further, let $x_1, y_1, z_1, x_2, \cdots x_N, y_N, z_N$ be the increments of the rectangular coordinates of the N nuclei. For small vibrations the R's are linear in the x's, y's and z's:

$$\mathfrak{R}=\begin{pmatrix}R_1\\R_2\\\vdots\\R_n\end{pmatrix}=B\mathfrak{X}, \quad \text{where} \quad \mathfrak{X}=\begin{pmatrix}x_1\\y_1\\\vdots\\z_N\end{pmatrix}, \qquad (1)$$

B being a constant matrix of n rows and $3N$ columns. Now after the operation C, the molecule is indistinguishable from its original self, and we have a new equation obtained by writing (1) down for the new molecule:

$$\overset{c}{R} = B\overset{c}{\mathscr{X}} \tag{2}$$

Here $\overset{c}{R}$'s are the coordinates of the molecule which will be brought into coincidence with R's by the operation C, and

$$\overset{c}{\mathscr{X}} = \begin{pmatrix} \Gamma_c\begin{pmatrix} x_{c^{-1}1} \\ y_{c^{-1}1} \\ z_{c^{-1}1} \end{pmatrix} \\ \vdots \\ \Gamma_c\begin{pmatrix} x_{c^{-1}N} \\ y_{c^{-1}N} \\ z_{c^{-1}N} \end{pmatrix} \end{pmatrix} \tag{3}$$

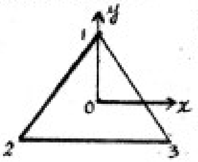
where $c^{-1}i$ is the nucleus which will become after the operation C the nucleus i. Let Z_c be a square matrix of order 3N with the elements ${}_{ix}(Z_c)_{jx} = \delta_{j,c^{-1}i}$, ${}_{ix}(Z_c)_{jy} = 0$ etc. $i,j = 1,2,\cdots N$,

and let T_c stand for $\begin{pmatrix} \Gamma_c & & \\ & \ddots & \\ & & \Gamma_c \end{pmatrix}$

then (2) and (3) give $\overset{c}{R} = BT_cZ_c\mathscr{X}$. (4)

This equation holds for every operation C and is the mathematical expression of the symmetry of the molecule.

EXAMPLE Consider three equivalent nuclei forming an equilateral triangle. Let C be the operation: Rotation counterclockwise through $120°$ about O. Then



$$c^{-1}1 = 3, \quad c^{-1}2 = 1, \quad c^{-1}3 = 2,$$

and

$$Z_c = \begin{pmatrix} 0 & 0 & \begin{smallmatrix}1&&\\&1&\\&&1\end{smallmatrix} \\ \begin{smallmatrix}1&&\\&1&\\&&1\end{smallmatrix} & 0 & 0 \\ 0 & \begin{smallmatrix}1&&\\&1&\\&&1\end{smallmatrix} & 0 \end{pmatrix}, \qquad \Gamma_c = \begin{pmatrix} -\frac{1}{2} & -\frac{\sqrt{3}}{2} & 0 \\ \frac{\sqrt{3}}{2} & -\frac{1}{2} & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Let R_1 be the increment of the distance $\overline{12}$. $R_1 = \frac{1}{2}x_1 + \frac{\sqrt{3}}{2}y_1 - \frac{1}{2}x_2 - \frac{\sqrt{3}}{2}y_2$.

Then $\overset{c}{R}_1$ is that of $\overline{31}$, and $\overset{c}{R}_1 = -\frac{1}{2}x_1 + \frac{\sqrt{3}}{2}y_1 + \frac{1}{2}x_3 - \frac{\sqrt{3}}{2}y_3$.

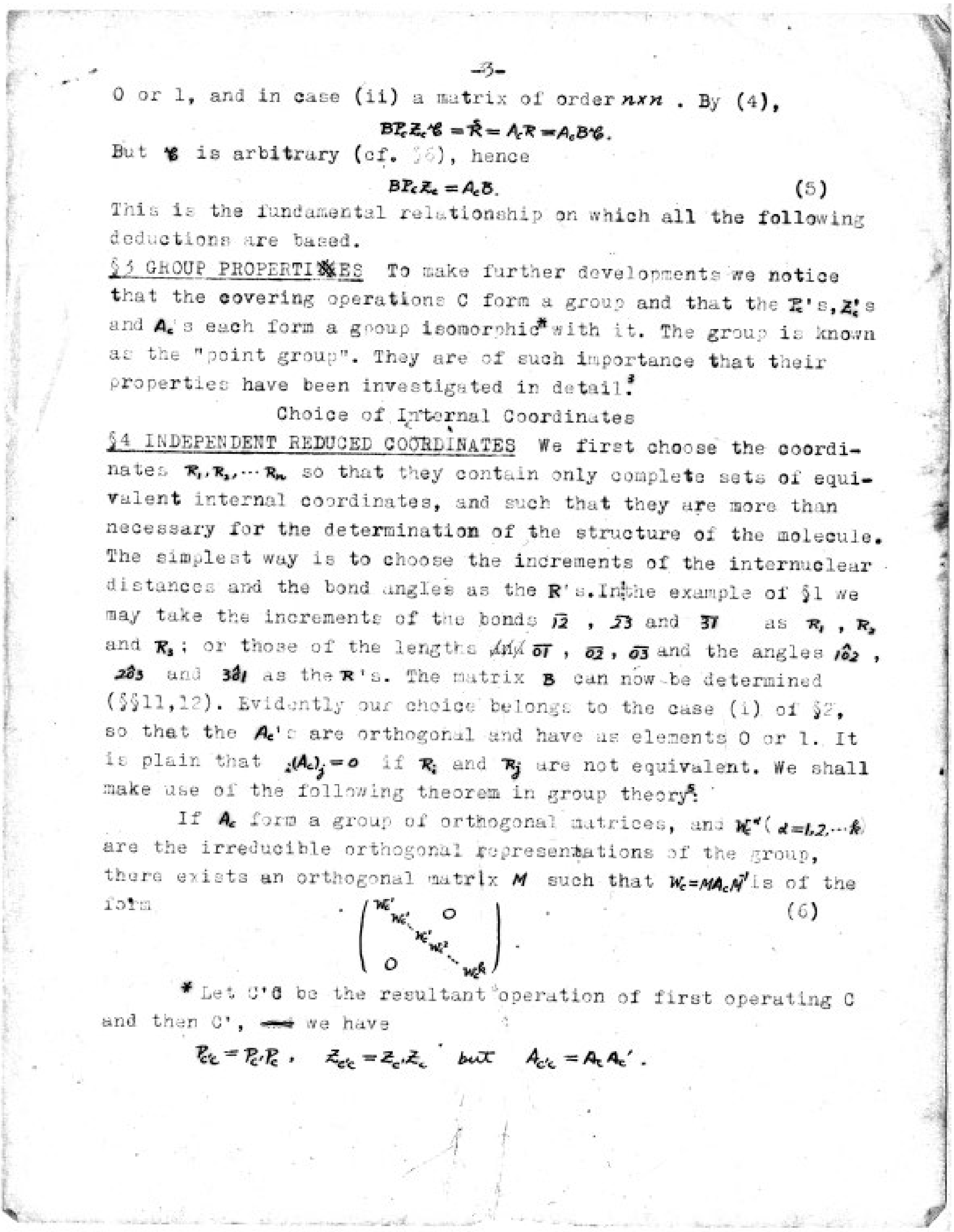
Thus

$$B = \left(\tfrac{1}{2}, \tfrac{\sqrt{3}}{2}, 0, -\tfrac{1}{2}, -\tfrac{\sqrt{3}}{2}, 0, 0, 0, 0\right).$$

(4) becomes the identity

$$-\tfrac{1}{2}x_1 + \tfrac{\sqrt{3}}{2}y_1 + \tfrac{1}{2}x_3 - \tfrac{\sqrt{3}}{2}y_3 = \left(\tfrac{1}{2}\ \tfrac{\sqrt{3}}{2}\ 0\ -\tfrac{1}{2}\ -\tfrac{\sqrt{3}}{2}\ 0\ 0\ 0\ 0\right)\begin{pmatrix} \Gamma_c & 0 & 0 \\ 0 & \Gamma_c & 0 \\ 0 & 0 & \Gamma_c \end{pmatrix} Z_c \begin{pmatrix} x_1 \\ y_1 \\ \vdots \\ z_3 \end{pmatrix}.$$

§2 FUNDAMENTAL RELATIONSHIP In some instances the coordinates $R_1, R_2, \cdots R_n$ are sufficient to determine $\overset{c}{R}_1, \overset{c}{R}_2, \cdots \overset{c}{R}_n$ for all covering operations C. This is the case if (i) the R's contain only complete sets of equivalent coordinates (e.g. in the last example when $\begin{smallmatrix}R_1\\R_2\\R_3\end{smallmatrix}$ = increment of $\begin{smallmatrix}\overline{12}\\\overline{23}\\\overline{31}\end{smallmatrix}$); or if (ii) the R's are all that are necessary to describe the internal structure of the molecule. In both cases we have for small vibrations $\overset{c}{R} = \overset{c}{A}R$. where A_c is in case (i) an orthogonal matrix having as elements



-3-

0 or 1, and in case (ii) a matrix of order $n \times n$. By (4),

$$BP_cZ_c\mathcal{C} = \hat{R} = A_cR = A_cB\mathcal{C}.$$

But $\mathcal{C}$ is arbitrary (cf. §6), hence

$$BP_cZ_c = A_cB. \qquad (5)$$

This is the fundamental relationship on which all the following deductions are based.

§3 GROUP PROPERTIES To make further developments we notice that the covering operations C form a group and that the P's, Z_c's and A_c's each form a group isomorphic* with it. The group is known as the "point group". They are of such importance that their properties have been investigated in detail[3].

Choice of Internal Coordinates

§4 INDEPENDENT REDUCED COORDINATES We first choose the coordinates $R_1, R_2, \cdots R_n$ so that they contain only complete sets of equivalent internal coordinates, and such that they are more than necessary for the determination of the structure of the molecule. The simplest way is to choose the increments of the internuclear distances and the bond angles as the R's. In the example of §1 we may take the increments of the bonds $\overline{12}$, $\overline{23}$ and $\overline{31}$ as R_1 , R_2 and R_3 ; or those of the lengths $\overline{01}$, $\overline{02}$, $\overline{03}$ and the angles $1\hat{0}2$, $2\hat{0}3$ and $3\hat{0}1$ as the R's. The matrix B can now be determined (§§11,12). Evidently our choice belongs to the case (i) of §2, so that the A_c's are orthogonal and have as elements 0 or 1. It is plain that $_i(A_c)_j = 0$ if R_i and R_j are not equivalent. We shall make use of the following theorem in group theory[5]:

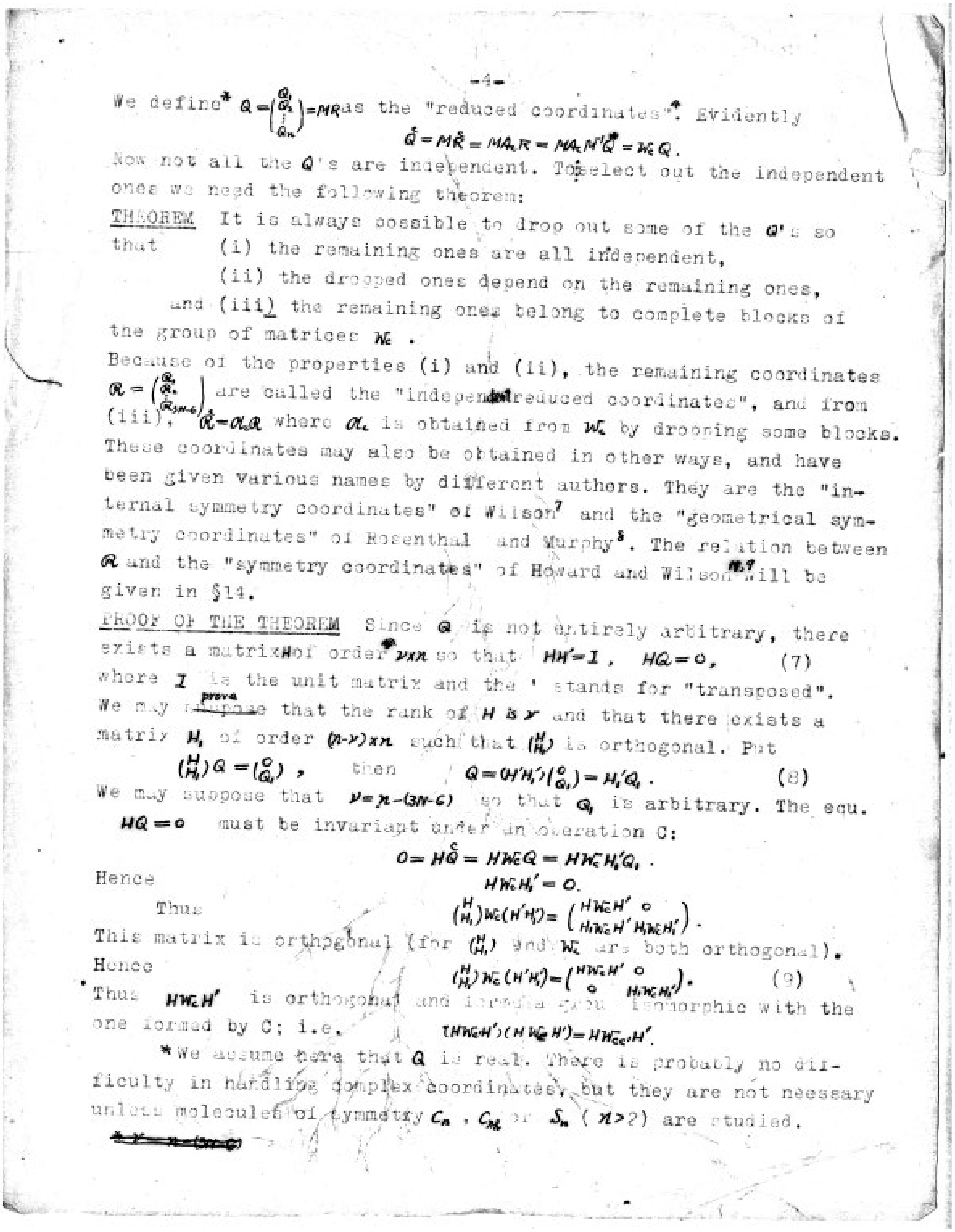
If A_c form a group of orthogonal matrices, and W_c^α $(\alpha = 1, 2, \cdots k)$ are the irreducible orthogonal representations of the group, there exists an orthogonal matrix M such that $W_c = MA_cM^{-1}$ is of the form

$$\begin{pmatrix} W_c^1 & & & & & 0 \\ & W_c^1 & & & & \\ & & \ddots & & & \\ & & & W_c^1 & & \\ & & & & W_c^2 & \\ 0 & & & & & \ddots \; W_c^k \end{pmatrix}. \qquad (6)$$

---

* Let C'C be the resultant operation of first operating C and then C', we have

$$P_{c'c} = P_{c'}P_c , \quad Z_{c'c} = Z_{c'}Z_c \quad \text{but} \quad A_{c'c} = A_cA_{c'} .$$



We define* $Q = \begin{pmatrix} Q_1 \\ Q_2 \\ \vdots \\ Q_n \end{pmatrix} = MR$ as the "reduced coordinates"*. Evidently

$$\overset{c}{Q} = M\overset{c}{R} = MA_cR = MA_cM'Q = W_cQ.$$

Now not all the Q's are independent. To select out the independent ones we need the following theorem:

THEOREM It is always possible to drop out some of the Q's so that (i) the remaining ones are all independent,

(ii) the dropped ones depend on the remaining ones,

and (iii) the remaining ones belong to complete blocks of the group of matrices W_c.

Because of the properties (i) and (ii), the remaining coordinates $\mathcal{R} = \begin{pmatrix} \mathcal{R}_1 \\ \mathcal{R}_2 \\ \vdots \\ \mathcal{R}_{3N-6} \end{pmatrix}$ are called the "independent reduced coordinates", and from (iii), $\overset{c}{\mathcal{R}} = \alpha_c\mathcal{R}$ where α_c is obtained from W_c by dropping some blocks. These coordinates may also be obtained in other ways, and have been given various names by different authors. They are the "internal symmetry coordinates" of Wilson[7] and the "geometrical symmetry coordinates" of Rosenthal and Murphy[8]. The relation between $\mathcal{R}$ and the "symmetry coordinates" of Howard and Wilson[10,9] will be given in §14.

PROOF OF THE THEOREM Since Q is not entirely arbitrary, there exists a matrix H of order* $\nu \times n$ so that $HH' = I$, $HQ = 0$, (7)

where I is the unit matrix and the ' stands for "transposed". We may ~~suppose~~ prove that the rank of H is ν and that there exists a matrix H_1 of order $(n-\nu)\times n$ such that $\binom{H}{H_1}$ is orthogonal. Put

$$\binom{H}{H_1}Q = \binom{0}{Q_1}, \quad \text{then} \quad Q = (H'H_1')\binom{0}{Q_1} = H_1'Q_1. \qquad (8)$$

We may suppose that $\nu = n-(3N-6)$ so that Q_1 is arbitrary. The equ. $HQ = 0$ must be invariant under an operation C:

$$0 = H\overset{c}{Q} = HW_cQ = HW_cH_1'Q_1.$$

Hence

$$HW_cH_1' = 0.$$

Thus

$$\binom{H}{H_1}W_c(H'H_1') = \begin{pmatrix} HW_cH' & 0 \\ H_1W_cH' & H_1W_cH_1' \end{pmatrix}.$$

This matrix is orthogonal (for $\binom{H}{H_1}$ and W_c are both orthogonal). Hence

$$\binom{H}{H_1}W_c(H'H_1') = \begin{pmatrix} HW_cH' & 0 \\ 0 & H_1W_cH_1' \end{pmatrix}. \qquad (9)$$

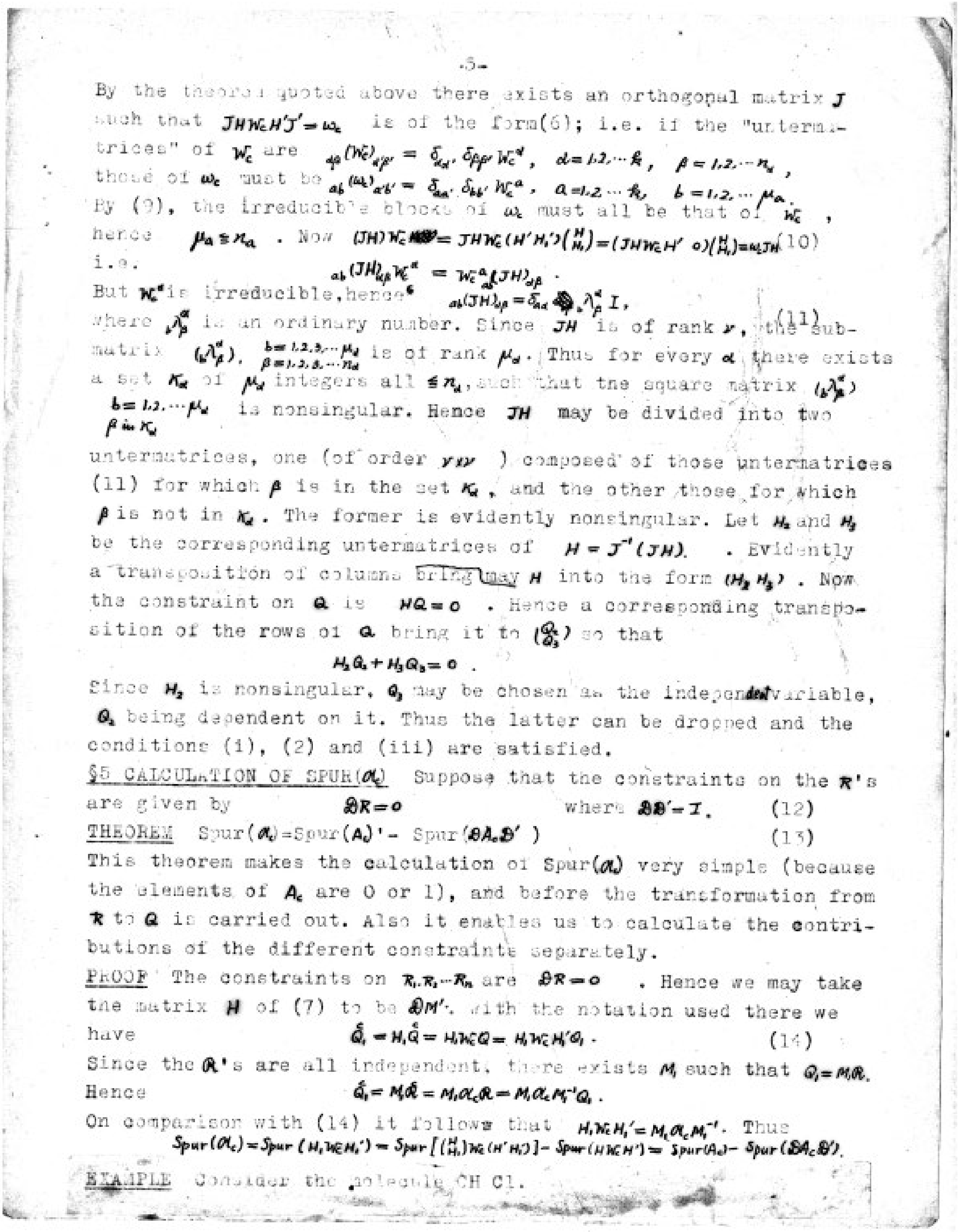
Thus HW_cH' is orthogonal and forms a group isomorphic with the one formed by C; i.e.

$$(HW_cH')(HW_{c'}H') = HW_{cc'}H'.$$

---

* We assume here that Q is real. There is probably no difficulty in handling complex coordinates, but they are not neesary unless molecules of symmetry C_n, C_{nh} or S_n ($n>2$) are studied.

~~* $\nu = n-(3N-6)$~~



-5-

By the theorem quoted above there exists an orthogonal matrix J such that $JHW_cH'J' = \omega_c$ is of the form(6); i.e. if the "untermatrices" of W_c are ${}_{\alpha\beta}(W_c)_{\alpha'\beta'} = \delta_{\alpha\alpha'}\delta_{\beta\beta'}W_c^{\alpha}$, $\alpha = 1,2,\cdots k$, $\beta = 1,2,\cdots n_\alpha$, those of ω_c must be ${}_{ab}(\omega_c)_{a'b'} = \delta_{aa'}\delta_{bb'}W_c^{a}$, $a = 1,2\cdots k$, $b = 1,2,\cdots \mu_a$. By (9), the irreducible blocks of ω_c must all be that of W_c, hence $\mu_a \leq n_a$. Now $(JH)W_c = JHW_c(H'H_1')\binom{H}{H_1} = (JHW_cH' \;\; 0)\binom{H}{H_1} = \omega_c JH$ (10) i.e.

$${}_{ab}(JH)_{\alpha\beta}W_c^{\alpha} = W_c^{a}\,{}_{ab}(JH)_{\alpha\beta}\,.$$

But W_c^{α} is irreducible, hence ${}_{ab}(JH)_{\alpha\beta} = \delta_{a\alpha}\,{}_b\lambda_\beta^{\alpha} I$, (11) where ${}_b\lambda_\beta^{\alpha}$ is an ordinary number. Since JH is of rank ν, the submatrix $({}_b\lambda_\beta^{\alpha})$, $\begin{matrix} b = 1,2,3,\cdots \mu_\alpha \\ \beta = 1,2,3,\cdots n_\alpha \end{matrix}$ is of rank μ_α. Thus for every α there exists a set K_α of μ_α integers all $\leq n_\alpha$, such that the square matrix $({}_b\lambda_\beta^{\alpha})$ $\begin{matrix} b = 1,2,\cdots \mu_\alpha \\ \beta \text{ in } K_\alpha \end{matrix}$ is nonsingular. Hence JH may be divided into two untermatrices, one (of order $\nu\times\nu$) composed of those untermatrices (11) for which β is in the set K_α, and the other those for which β is not in K_α. The former is evidently nonsingular. Let H_2 and H_3 be the corresponding untermatrices of $H = J^{-1}(JH)$. Evidently a transposition of columns ~~bring~~ may H into the form $(H_2\,H_3)$. Now the constraint on Q is $HQ = 0$. Hence a corresponding transposition of the rows of Q bring it to $\binom{Q_2}{Q_3}$ so that

$$H_2Q_2 + H_3Q_3 = 0\,.$$

Since H_2 is nonsingular, Q_3 may be chosen as the independent variable, Q_2 being dependent on it. Thus the latter can be dropped and the conditions (1), (2) and (iii) are satisfied.

§5 CALCULATION OF SPUR($\mathfrak{A}_c$) Suppose that the constraints on the $\mathcal{R}$'s are given by $\mathcal{B}\mathcal{R} = 0$ where $\mathcal{B}\mathcal{B}' = I$. (12)

THEOREM Spur($\mathfrak{A}_c$)=Spur(A_c)'- Spur($\mathcal{B}A_c\mathcal{B}'$) (13)

This theorem makes the calculation of Spur($\mathfrak{A}_c$) very simple (because the elements of A_c are 0 or 1), and before the transformation from $\mathcal{R}$ to Q is carried out. Also it enables us to calculate the contributions of the different constraints separately.

PROOF The constraints on $\mathcal{R}_1,\mathcal{R}_2\cdots\mathcal{R}_n$ are $\mathcal{B}\mathcal{R} = 0$. Hence we may take the matrix H of (7) to be $\mathcal{B}M'$. With the notation used there we have $\overset{c}{Q}_1 = H_1\overset{c}{Q} = H_1W_cQ = H_1W_cH_1'Q_1\,.$ (14)

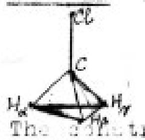
Since the $\mathcal{R}$'s are all independent, there exists M_1 such that $Q_1 = M_1\mathcal{R}$. Hence $\overset{c}{Q}_1 = M_1\overset{c}{\mathcal{R}} = M_1\mathfrak{A}_c\mathcal{R} = M_1\mathfrak{A}_cM_1^{-1}Q_1\,.$

On comparison with (14) it follows that $H_1W_cH_1' = M_1\mathfrak{A}_cM_1^{-1}$. Thus

$$\mathrm{Spur}(\mathfrak{A}_c) = \mathrm{Spur}(H_1W_cH_1') = \mathrm{Spur}\left[\binom{H}{H_1}W_c(H'H_1')\right] - \mathrm{Spur}(HW_cH') = \mathrm{Spur}(A_c) - \mathrm{Spur}(\mathcal{B}A_c\mathcal{B}')\,.$$

EXAMPLE Consider the molecule CH Cl.

EXAMPLE Consider the molecule CH_3Cl. Take



R_1, R_2, R_3, R_4 =increments of the distances $\overline{CCl}, \overline{CH_\alpha}, \overline{CH_\beta}, \overline{CH_\gamma}$.

R_5, R_6, R_7 =increments of the angles $Cl\hat{C}H_\alpha, Cl\hat{C}H_\beta, Cl\hat{C}H_\gamma$,

R_8, R_9, R_{10} =increments of the angles $H_\beta\hat{C}H_\gamma, H_\gamma\hat{C}H_\alpha, H_\alpha\hat{C}H_\beta$.

The constraint~~s~~ is $[R_5+R_6+R_7]f+[R_8+R_9+R_{10}]=0$,
where f is a constant. Thus $\mathfrak{D}=\frac{1}{\sqrt{3f^2+3}}(0\;0\;0\;0\;f\;f\;f\;1\;1\;1)$.
The theorem leads to

Operation C		Spur(A_c)	Spur$(\mathfrak{D}A_c\mathfrak{D}')$*	Spur$(\mathfrak{A}_c)$
No Motion	:C_1	10	1	9
Rotation about $\overline{CCl}$ through	$120°:C_2$	1	1	0
	$240°:C_3$			
Reflection about the plane	$ClCH_\alpha:C_4$	4	1	3
	$ClCH_\beta:C_5$			
	$ClCH_\gamma:C_6$			

The Kinetic and the Potential Energies

§6 CHOICE OF AXES Suppose that $\begin{pmatrix}R_1\\R_2\\\vdots\\R_{3N-6}\end{pmatrix}=\mathbf{R}$ are the increments of $3N-6$ independent internal coordinates. For small vibrations, $\mathbf{R}=B\mathfrak{X}$ where $\mathfrak{X}=\begin{pmatrix}x_1\\\vdots\\x_{3N}\end{pmatrix}$ is defined in §1. Suppose that the equilibrium positions of the nuclei are, in rectangular coordinates, $(X_1Y_1Z_1), (X_2Y_2Z_2), \cdots (X_NY_NZ_N)$. Let $m_1, m_2, \cdots m_N$ be their masses. Write

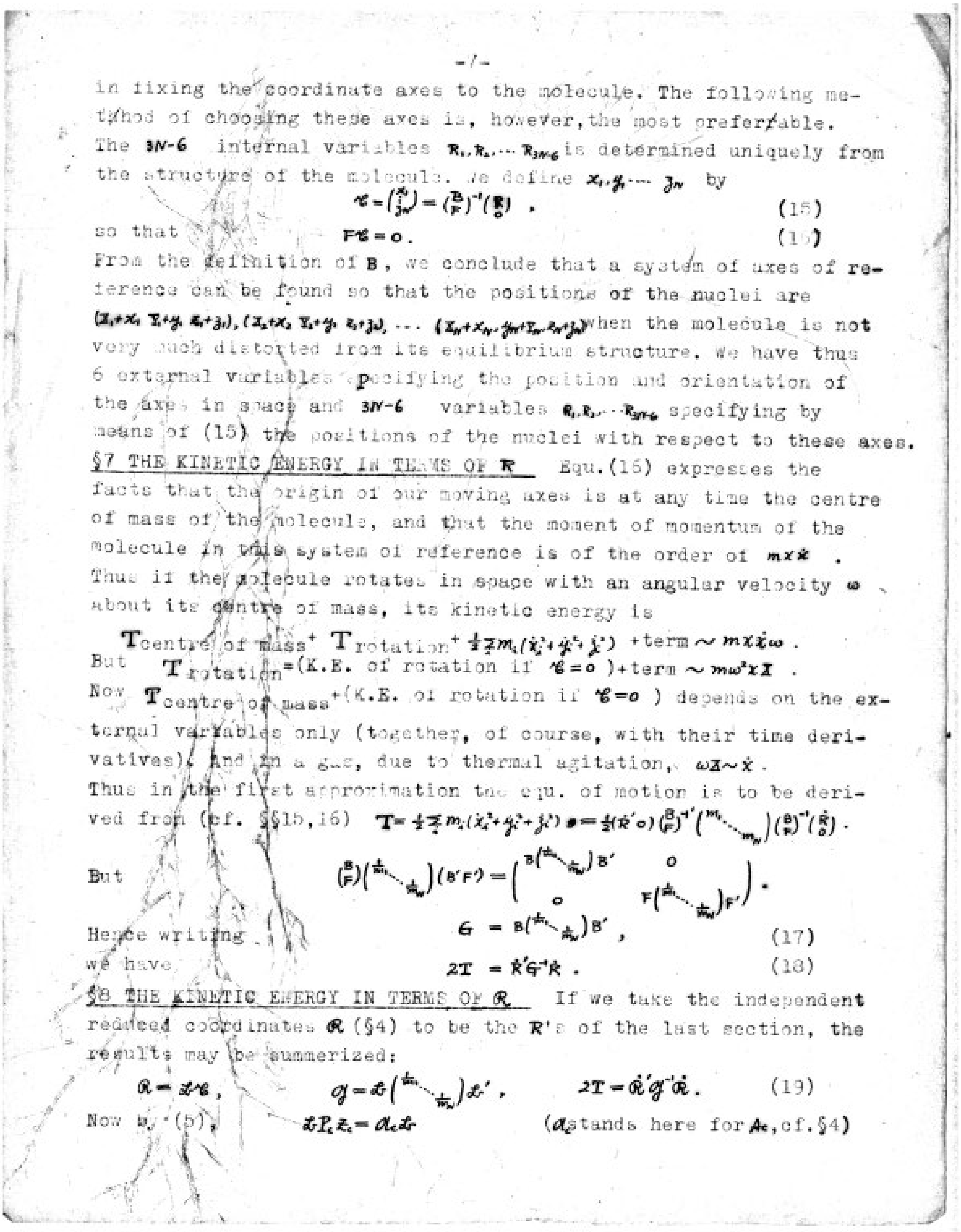
$$F=\begin{pmatrix} m_1 & 0 & 0 & m_2 & 0 & 0 & \cdot & \cdot & \cdot & 0 \\ 0 & m_1 & 0 & 0 & m_2 & 0 & \cdot & \cdot & \cdot & 0 \\ 0 & 0 & m_1 & 0 & 0 & m_2 & \cdot & \cdot & \cdot & m_N \\ 0 & m_1Z_1 & -m_1Y_1 & 0 & m_2Z_2 & -m_2Y_2 & \cdot & \cdot & \cdot & -m_NY_N \\ -m_1Z_1 & 0 & m_1X_1 & -m_2Z_2 & 0 & m_2X_2 & \cdot & \cdot & \cdot & m_NX_N \\ m_1Y_1 & -m_1X_1 & 0 & m_2Y_2 & -m_2X_2 & 0 & \cdot & \cdot & \cdot & 0 \end{pmatrix}$$

Then the first column of $B\begin{pmatrix}\frac{1}{m_1} & & \\ & \ddots & \\ & & \frac{1}{m_N}\end{pmatrix}F'$ is the ~~1/N~~ matrix $B\begin{pmatrix}1\\0\\0\\1\\0\\0\\\vdots\end{pmatrix}$ which is the value of $\mathbf{R}$ when all the nuclei are displaced by 1 unit of lenth along the X-axis. But the $\mathbf{R}$'s are the increments of internal variables, hence $B\begin{pmatrix}1\\0\\0\\1\\\vdots\\0\end{pmatrix}=0$. In the same way we can show that

$$B\begin{pmatrix}\frac{1}{m_1} & & \\ & \ddots & \\ & & \frac{1}{m_N}\end{pmatrix}F'=0.$$

We have thus far described the molecule in a certain system of rectangular coordinates. But when the molecule moves and rotates in space as well as vibrates, there remains an arbitrariness

* From this it is immediately seen that any constraint having the same coefficient for equivalent coordinates contributes 1 to Spur$(\mathfrak{D}A_c\mathfrak{D}')$.



in fixing the coordinate axes to the molecule. The following method of choosing these axes is, however, the most prefer~~r~~able. The $3N-6$ internal variables $\mathbf{R}_1, \mathbf{R}_2, \cdots \mathbf{R}_{3N-6}$ is determined uniquely from the structure of the molecule. We define $x_1, y_1, \cdots z_N$ by

$$\mathcal{X} = \begin{pmatrix} x_1 \\ \vdots \\ z_N \end{pmatrix} = \begin{pmatrix} B \\ F \end{pmatrix}^{-1} \begin{pmatrix} \mathbf{R} \\ 0 \end{pmatrix} . \qquad (15)$$

so that

$$F\mathcal{X} = 0 . \qquad (16)$$

From the definition of B, we conclude that a system of axes of reference can be found so that the positions of the nuclei are $(X_1+x_1,\ Y_1+y_1,\ Z_1+z_1), (X_2+x_2,\ Y_2+y_2,\ Z_2+z_2), \dots (X_N+x_N,\ Y_N+y_N,\ Z_N+z_N)$ when the molecule is not very much distorted from its equilibrium structure. We have thus 6 external variables specifying the position and orientation of the axes in space and $3N-6$ variables $\mathbf{R}_1, \mathbf{R}_2, \cdots \mathbf{R}_{3N-6}$ specifying by means of (15) the positions of the nuclei with respect to these axes.

§7 THE KINETIC ENERGY IN TERMS OF $\mathbf{R}$ Equ. (16) expresses the facts that the origin of our moving axes is at any time the centre of mass of the molecule, and that the moment of momentum of the molecule in this system of reference is of the order of $m x \dot{x}$. Thus if the molecule rotates in space with an angular velocity ω about its centre of mass, its kinetic energy is

$$T_{\text{centre of mass}} + T_{\text{rotation}} + \tfrac{1}{2}\textstyle\sum m_i(\dot{x}_i^2+\dot{y}_i^2+\dot{z}_i^2) + \text{term} \sim m x \dot{x} \omega .$$

But $T_{\text{rotation}} = (\text{K.E. of rotation if } \mathcal{X}=0) + \text{term} \sim m\omega^2 x I$.

Now $T_{\text{centre of mass}} + (\text{K.E. of rotation if } \mathcal{X}=0)$ depends on the external variables only (together, of course, with their time derivatives). And in a gas, due to thermal agitation, $\omega I \sim \dot{x}$. Thus in the first approximation the equ. of motion is to be derived from (cf. §§15,16)

$$T = \tfrac{1}{2}\textstyle\sum m_i(\dot{x}_i^2+\dot{y}_i^2+\dot{z}_i^2) = \tfrac{1}{2}(\dot{\mathbf{R}}'\ 0)\begin{pmatrix} B \\ F \end{pmatrix}^{-1\prime}\begin{pmatrix} m_1 & & \\ & \ddots & \\ & & m_N \end{pmatrix}\begin{pmatrix} B \\ F \end{pmatrix}^{-1}\begin{pmatrix} \dot{\mathbf{R}} \\ 0 \end{pmatrix} .$$

But

$$\begin{pmatrix} B \\ F \end{pmatrix}\begin{pmatrix} \frac{1}{m_1} & & \\ & \ddots & \\ & & \frac{1}{m_N} \end{pmatrix}(B'\ F') = \begin{pmatrix} B\begin{pmatrix} \frac{1}{m_1} & & \\ & \ddots & \\ & & \frac{1}{m_N} \end{pmatrix}B' & 0 \\ 0 & F\begin{pmatrix} \frac{1}{m_1} & & \\ & \ddots & \\ & & \frac{1}{m_N} \end{pmatrix}F' \end{pmatrix} .$$

Hence writing

$$G = B\begin{pmatrix} \frac{1}{m_1} & & \\ & \ddots & \\ & & \frac{1}{m_N} \end{pmatrix}B' , \qquad (17)$$

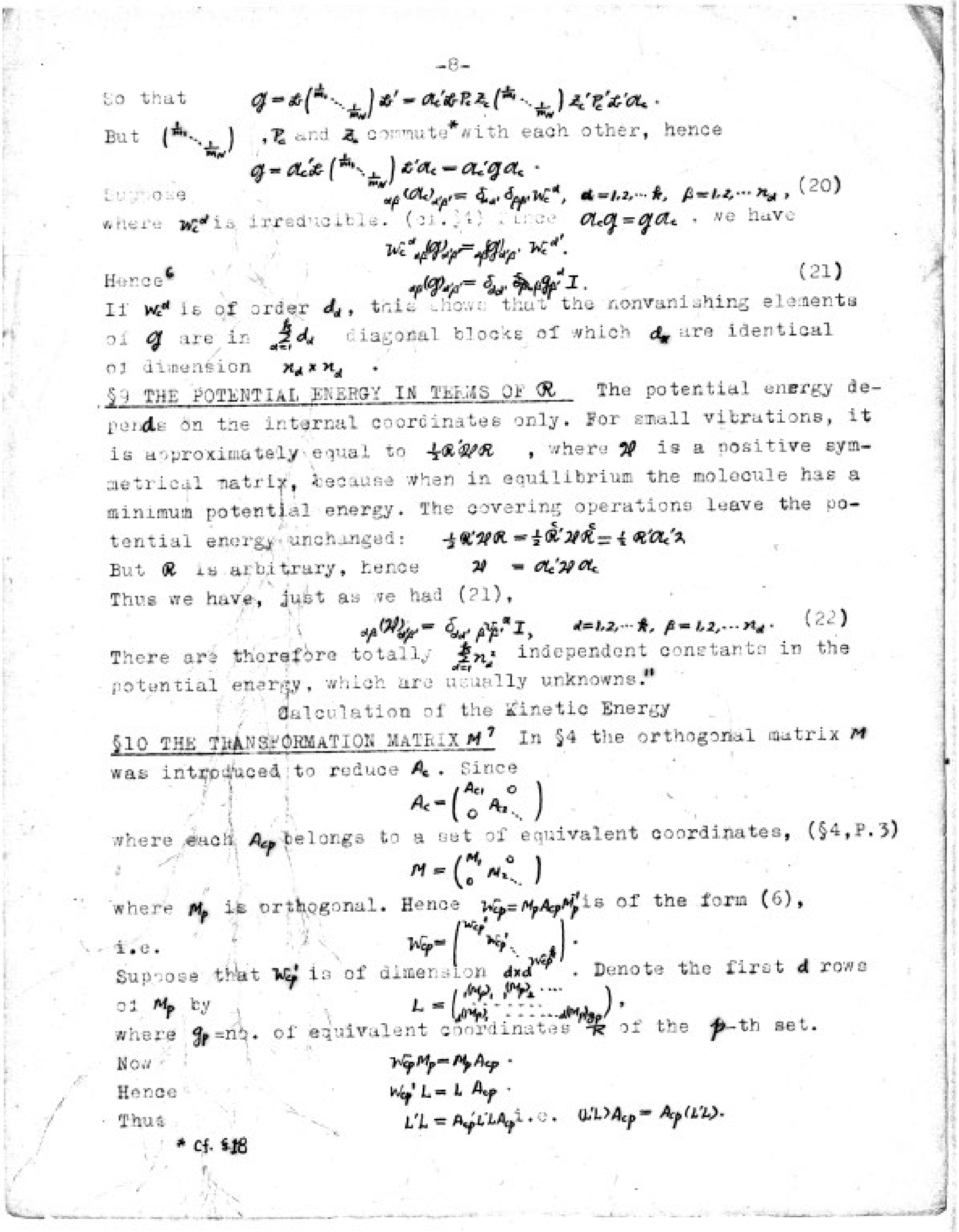
we have

$$2T = \dot{\mathbf{R}}'G^{-1}\dot{\mathbf{R}} . \qquad (18)$$

§8 THE KINETIC ENERGY IN TERMS OF $\mathcal{R}$ If we take the independent reduced coordinates $\mathcal{R}$ (§4) to be the $\mathbf{R}$'s of the last section, the results may be summerized:

$$\mathcal{R} = \mathcal{L}\mathcal{X} , \qquad \mathcal{G} = \mathcal{L}\begin{pmatrix} \frac{1}{m_1} & & \\ & \ddots & \\ & & \frac{1}{m_N} \end{pmatrix}\mathcal{L}' , \qquad 2T = \dot{\mathcal{R}}'\mathcal{G}^{-1}\dot{\mathcal{R}} . \qquad (19)$$

Now by (5), $\mathcal{L}P_c\mathcal{Z}_c = \mathcal{A}_c\mathcal{L}$ ($\mathcal{A}$ stands here for A_c, cf. §4)



So that $\mathcal{G}=\mathcal{X}\begin{pmatrix}\frac{1}{m_1} & & \\ & \ddots & \\ & & \frac{1}{m_N}\end{pmatrix}\mathcal{X}'=\mathcal{A}_c'\mathcal{X}P_cZ_c\begin{pmatrix}\frac{1}{m_1} & & \\ & \ddots & \\ & & \frac{1}{m_N}\end{pmatrix}Z_c'P_c'\mathcal{X}'\mathcal{A}_c$.

But $\begin{pmatrix}\frac{1}{m_1} & & \\ & \ddots & \\ & & \frac{1}{m_N}\end{pmatrix}$, P_c and Z_c commute* with each other, hence

$$\mathcal{G}=\mathcal{A}_c'\mathcal{X}\begin{pmatrix}\frac{1}{m_1} & & \\ & \ddots & \\ & & \frac{1}{m_N}\end{pmatrix}\mathcal{X}'\mathcal{A}_c=\mathcal{A}_c'\mathcal{G}\mathcal{A}_c .$$

Suppose

$${}_{\alpha\beta}(\mathcal{A}_c)_{\alpha'\beta'}=\delta_{\alpha\alpha'}\delta_{\beta\beta'}\mathfrak{M}_c^{\alpha},\quad \alpha=1,2,\cdots k,\quad \beta=1,2,\cdots n_\alpha, \tag{20}$$

where $\mathfrak{M}_c^{\alpha}$ is irreducible. (cf. §4) Since $\mathcal{A}_c\mathcal{G}=\mathcal{G}\mathcal{A}_c$, we have

$$\mathfrak{M}_c^{\alpha}\,{}_{\alpha\beta}(\mathcal{G})_{\alpha'\beta'}={}_{\alpha\beta}(\mathcal{G})_{\alpha'\beta'}\,\mathfrak{M}_c^{\alpha'}.$$

Hence[6]

$${}_{\alpha\beta}(\mathcal{G})_{\alpha'\beta'}=\delta_{\alpha\alpha'}\,{}_{\beta}g_{\beta'}^{\alpha}I. \tag{21}$$

If $\mathfrak{M}_c^{\alpha}$ is of order d_α, this shows that the nonvanishing elements of $\mathcal{G}$ are in $\sum_{\alpha=1}^{k}d_\alpha$ diagonal blocks of which d_α are identical of dimension $n_\alpha\times n_\alpha$.

§9 THE POTENTIAL ENERGY IN TERMS OF $\mathcal{R}$

The potential energy depends on the internal coordinates only. For small vibrations, it is approximately equal to $\frac{1}{2}\mathcal{R}'\mathcal{V}\mathcal{R}$, where $\mathcal{V}$ is a positive symmetrical matrix, because when in equilibrium the molecule has a minimum potential energy. The covering operations leave the potential energy unchanged: $\frac{1}{2}\mathcal{R}'\mathcal{V}\mathcal{R}=\frac{1}{2}\overset{c}{\mathcal{R}}{}'\mathcal{V}\overset{c}{\mathcal{R}}=\frac{1}{2}\mathcal{R}'\mathcal{A}_c'\mathcal{V}$

But $\mathcal{R}$ is arbitrary, hence $\mathcal{V}=\mathcal{A}_c'\mathcal{V}\mathcal{A}_c$

Thus we have, just as we had (21),

$${}_{\alpha\beta}(\mathcal{V})_{\alpha'\beta'}=\delta_{\alpha\alpha'}\,{}_{\beta}\gamma_{\beta'}^{\alpha}I,\qquad \alpha=1,2,\cdots k,\ \beta=1,2,\cdots n_\alpha. \tag{22}$$

There are therefore totally $\sum_{\alpha=1}^{k}n_\alpha^2$ independent constants in the potential energy, which are usually unknowns.[ii]

Calculation of the Kinetic Energy

§10 THE TRANSFORMATION MATRIX M[7]

In §4 the orthogonal matrix M was introduced to reduce A_c. Since

$$A_c=\begin{pmatrix}A_{c1} & & 0\\ & A_{c2} & \\ 0 & & \ddots\end{pmatrix}$$

where each A_{cp} belongs to a set of equivalent coordinates, (§4, P.3)

$$M=\begin{pmatrix}M_1 & & 0\\ & M_2 & \\ 0 & & \ddots\end{pmatrix}$$

where M_p is orthogonal. Hence $\mathfrak{M}_{cp}=M_pA_{cp}M_p'$ is of the form (6), i.e.

$$\mathfrak{M}_{cp}=\begin{pmatrix}\mathfrak{M}_{cp}^{1} & & \\ & \mathfrak{M}_{cp}^{1} & \\ & & \ddots \\ & & & \mathfrak{M}_{cp}^{k}\end{pmatrix}.$$

Suppose that $\mathfrak{M}_{cp}^{1}$ is of dimension $d\times d$. Denote the first d rows of M_p by

$$L=\begin{pmatrix}{}_1(M_p)_1 & {}_1(M_p)_2 & \cdots \\ \vdots & \cdots & \\ {}_d(M_p)_1 & \cdots & {}_d(M_p)_{g_p}\end{pmatrix},$$

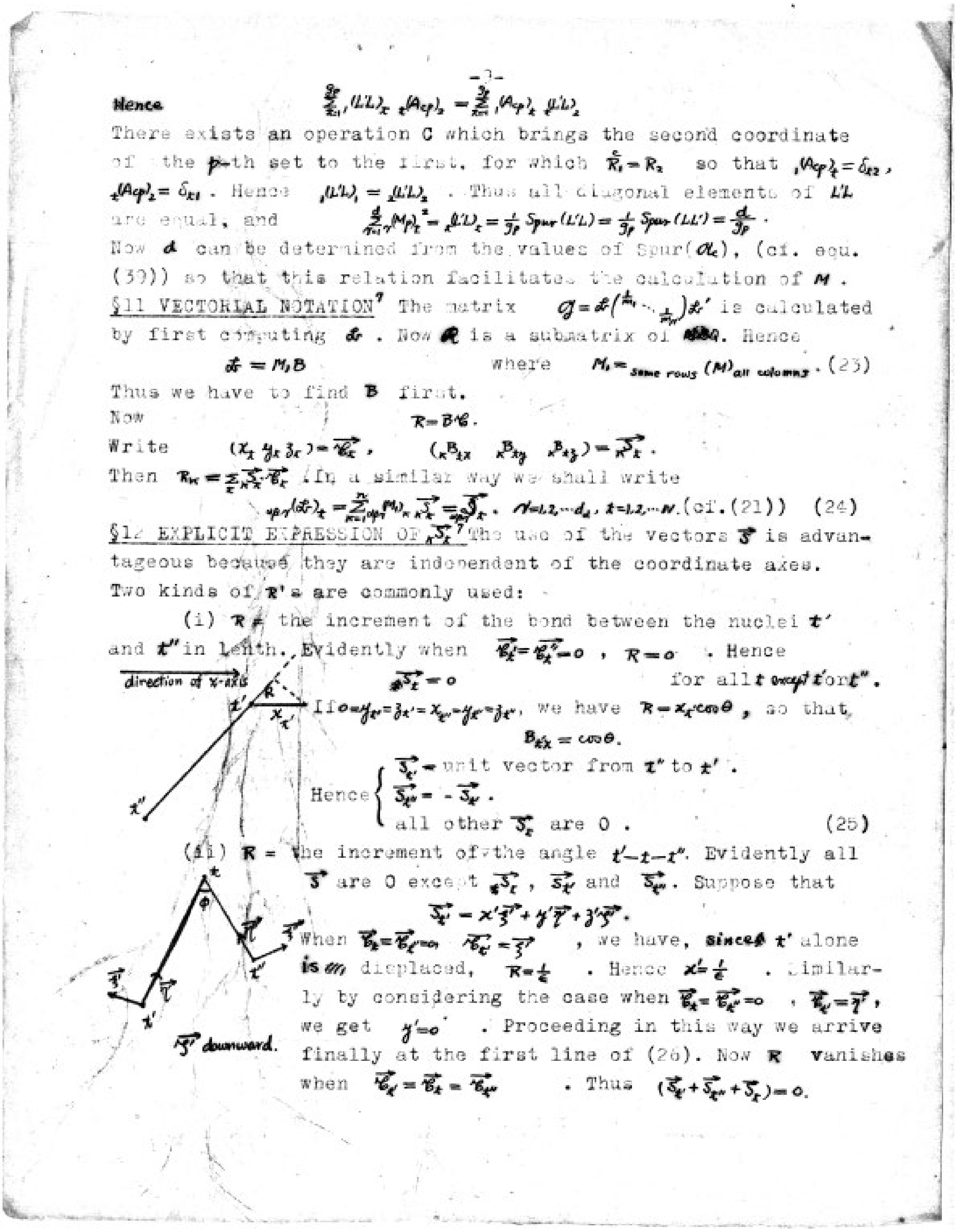
where g_p = no. of equivalent coordinates R of the p-th set.

Now $\mathfrak{M}_{cp}M_p=M_pA_{cp}$.

Hence $\mathfrak{M}_{cp}^{1}L=L\,A_{cp}$.

Thus $L'L=A_{cp}'L'LA_{cp}$ i.e. $(L'L)A_{cp}=A_{cp}(L'L)$.

* Cf. §18



– 7 –

Hence $\sum_{\kappa=1}^{g_p} {}_1(L'L)_\kappa \, {}_\kappa(A_{cp})_2 = \sum_{\kappa=1}^{g_p} {}_1(A_{cp})_\kappa \, {}_\kappa(L'L)_2$

There exists an operation C which brings the second coordinate of the p-th set to the first, for which $\overset{c}{R}_1 = R_2$ so that ${}_1(A_{cp})_\kappa = \delta_{\kappa 2}$, ${}_\kappa(A_{cp})_2 = \delta_{\kappa 1}$. Hence ${}_1(L'L)_1 = {}_2(L'L)_2$. Thus all diagonal elements of $L'L$ are equal, and $\sum_{r=1}^{d} {}_r(M_p)_\kappa^2 = {}_\kappa(L'L)_\kappa = \frac{1}{g_p} \, \mathrm{Spur}(L'L) = \frac{1}{g_p} \, \mathrm{Spur}(LL') = \frac{d}{g_p}$.
Now d can be determined from the values of Spur($\mathfrak{A}_c$), (cf. equ. (39)) so that this relation facilitates the calculation of M.

§11 VECTORIAL NOTATION[7] The matrix $G = \mathfrak{L}\left(\frac{1}{m_1} \cdots \frac{1}{m_N}\right)\mathfrak{L}'$ is calculated by first computing $\mathfrak{L}$. Now $\mathfrak{R}$ is a submatrix of [illegible]. Hence

$$\mathfrak{L} = M_1 B \qquad \text{where} \qquad M_1 = {}_{\text{some rows}}(M)_{\text{all columns}} . \quad (23)$$

Thus we have to find B first.
Now $\mathfrak{R} = B\mathfrak{X}$.
Write $(x_t \; y_t \; z_t) = \vec{r}_t$, $({}_\kappa B_{tx} \; {}_\kappa B_{ty} \; {}_\kappa B_{tz}) = {}_\kappa\vec{S}_t$.
Then $R_\kappa = \sum_t {}_\kappa\vec{S}_t \cdot \vec{r}_t$. In a similar way we shall write

$${}_{cp\,r}(\mathfrak{L})_t = \sum_{\kappa=1}^{n} {}_{cp\,r}(M)_\kappa \, {}_\kappa\vec{S}_t = {}_{cp\,r}\vec{S}_t , \quad r = 1,2,\cdots d_c , \; t = 1,2,\cdots N . \text{ (cf. (21))} \quad (24)$$

§12 EXPLICIT EXPRESSION OF ${}_\kappa\vec{S}_t$[7] The use of the vectors $\vec{S}$ is advantageous because they are independent of the coordinate axes. Two kinds of R's are commonly used:

(i) R = the increment of the bond between the nuclei t' and t'' in length. Evidently when $\vec{r}_{t'} = \vec{r}_{t''} = 0$, $R = 0$. Hence ${}_\kappa\vec{S}_t = 0$ for all t except t' or t''.

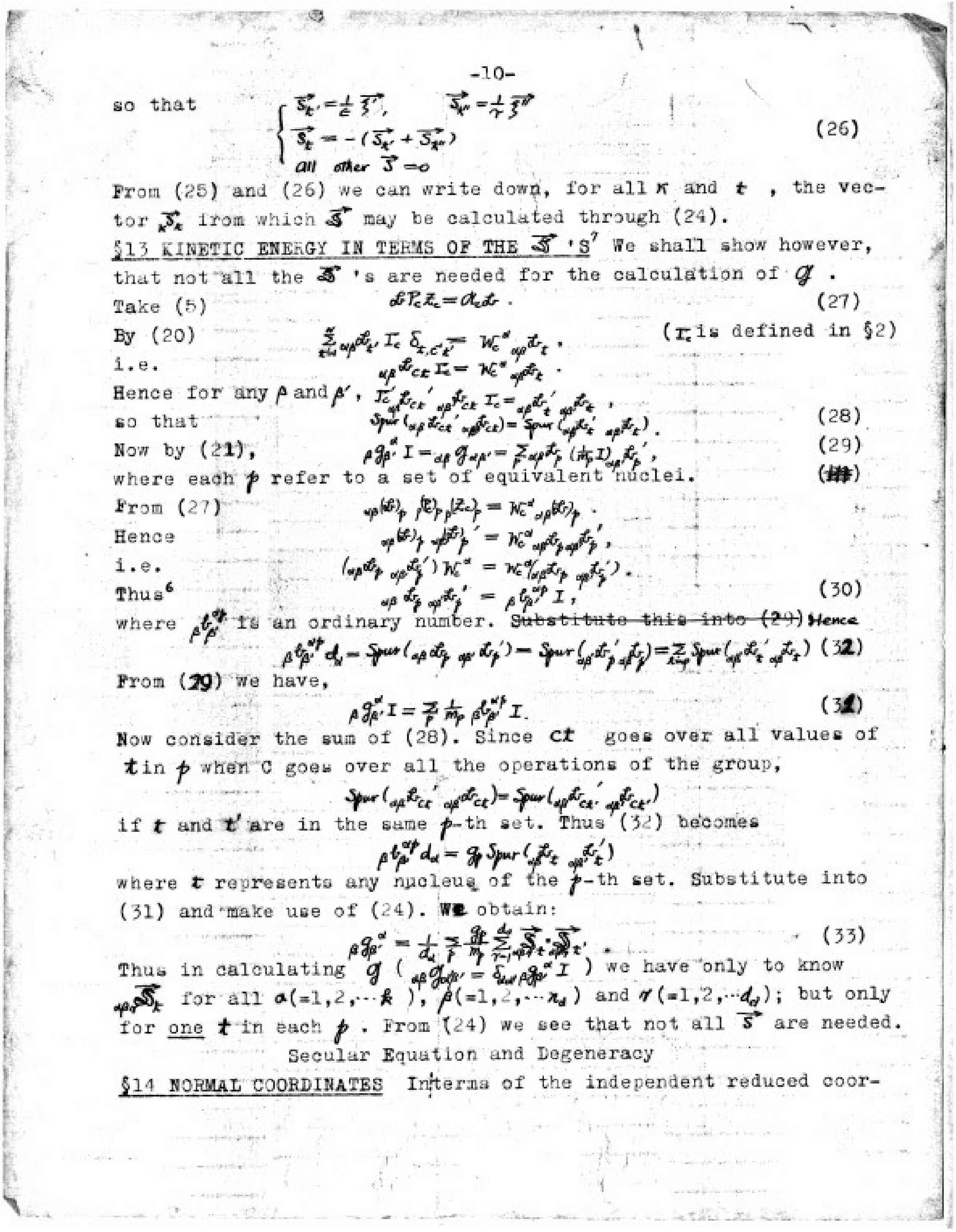
If $0 = y_{t'} = z_{t'} = x_{t''} = y_{t''} = z_{t''}$, we have $R = x_{t'}\cos\theta$, so that
$B_{t'x} = \cos\theta$.

$$\text{Hence} \begin{cases} \vec{S}_{t'} = \text{unit vector from } t'' \text{ to } t' . \\ \vec{S}_{t''} = -\vec{S}_{t'} . \\ \text{all other } \vec{S}_t \text{ are } 0 . \end{cases} \quad (25)$$

(ii) R = the increment of the angle t'–t–t''. Evidently all $\vec{S}$ are 0 except ${}_\kappa\vec{S}_t$, $\vec{S}_{t'}$ and $\vec{S}_{t''}$. Suppose that

$$\vec{S}_{t'} = x'\vec{\xi} + y'\vec{\eta} + z'\vec{\zeta} .$$

When $\vec{r}_t = \vec{r}_{t''} = 0$, $\vec{r}_{t'} = \vec{\xi}$, we have, since t' alone is displaced, $R = \frac{1}{\ell}$. Hence $x' = \frac{1}{\ell}$. Similarly by considering the case when $\vec{r}_t = \vec{r}_{t''} = 0$, $\vec{r}_{t'} = \vec{\eta}$, we get $y' = 0$. Proceeding in this way we arrive finally at the first line of (26). Now R vanishes when $\vec{r}_{t'} = \vec{r}_t = \vec{r}_{t''}$. Thus $(\vec{S}_{t'} + \vec{S}_{t''} + \vec{S}_t) = 0$.



so that

$$\begin{cases} \vec{S}_{t'} = \frac{1}{\epsilon}\vec{\xi}', \qquad \vec{S}_{t''} = \frac{1}{\tau}\vec{\xi}'' \\ \vec{S}_t = -(\vec{S}_{t'} + \vec{S}_{t''}) \\ \text{all other } \vec{\xi} = 0 \end{cases} \tag{26}$$

From (25) and (26) we can write down, for all κ and t, the vector ${}_\kappa\vec{S}_t$ from which $\vec{\mathfrak{S}}$ may be calculated through (24).

§13 KINETIC ENERGY IN TERMS OF THE $\vec{\mathfrak{S}}$ 'S[7] We shall show however, that not all the $\vec{\mathfrak{S}}$'s are needed for the calculation of $\mathfrak{g}$.

Take (5)

$$\mathfrak{t}\,P_c Z_c = \mathfrak{A}_c \mathfrak{t}. \tag{27}$$

By (20)

$$\sum_{t'=1}^{N} {}_{\alpha\beta}\mathfrak{t}_{t'}\, T_c\, \delta_{t,ct'} = W_c^{\alpha}\, {}_{\alpha\beta}\mathfrak{t}_t. \qquad (T_c \text{ is defined in §2})$$

i.e.

$${}_{\alpha\beta}\mathfrak{t}_{ct}\, T_c = W_c^{\alpha}\, {}_{\alpha\beta}\mathfrak{t}_t.$$

Hence for any β and β',

$$T_c'\, {}_{\alpha\beta}\mathfrak{t}_{ct}'\, {}_{\alpha\beta}\mathfrak{t}_{ct}\, T_c = {}_{\alpha\beta}\mathfrak{t}_t'\, {}_{\alpha\beta}\mathfrak{t}_t,$$

so that

$$\mathrm{Spur}({}_{\alpha\beta}\mathfrak{t}_{ct}'\, {}_{\alpha\beta}\mathfrak{t}_{ct}) = \mathrm{Spur}({}_{\alpha\beta}\mathfrak{t}_t'\, {}_{\alpha\beta}\mathfrak{t}_t). \tag{28}$$

Now by (21),

$${}_\beta g_{\beta'}^{\alpha} I = {}_{\alpha\beta}\mathfrak{g}_{\alpha\beta'} = \sum_p {}_{\alpha\beta}\mathfrak{t}_p \left(\frac{1}{m_p} I\right) {}_{\alpha\beta'}\mathfrak{t}_p', \tag{29}$$

where each p refer to a set of equivalent nuclei. ([illegible])

From (27)

$${}_{\alpha\beta}(\mathfrak{t})_p\, {}_p(\mathfrak{T})_p\, {}_p(Z_c)_p = W_c^{\alpha}\, {}_{\alpha\beta}(\mathfrak{t})_p.$$

Hence

$${}_{\alpha\beta}(\mathfrak{t})_p\, {}_{\alpha\beta}(\mathfrak{t})_p' = W_c^{\alpha}\, {}_{\alpha\beta}\mathfrak{t}_p\, {}_{\alpha\beta}\mathfrak{t}_p',$$

i.e.

$$({}_{\alpha\beta}\mathfrak{t}_p\, {}_{\alpha\beta'}\mathfrak{t}_p')\, W_c^{\alpha} = W_c^{\alpha}({}_{\alpha\beta}\mathfrak{t}_p\, {}_{\alpha\beta'}\mathfrak{t}_p').$$

Thus[6]

$${}_{\alpha\beta}\mathfrak{t}_p\, {}_{\alpha\beta'}\mathfrak{t}_p' = {}_\beta t_{\beta'}^{\alpha p} I, \tag{30}$$

where ${}_\beta t_{\beta'}^{\alpha p}$ is an ordinary number. ~~Substitute this into (29)~~ Hence

$${}_\beta t_{\beta'}^{\alpha p} d_\alpha = \mathrm{Spur}({}_{\alpha\beta}\mathfrak{t}_p\, {}_{\alpha\beta'}\mathfrak{t}_p') = \mathrm{Spur}({}_{\alpha\beta'}\mathfrak{t}_p'\, {}_{\alpha\beta}\mathfrak{t}_p) = \sum_{t \in p} \mathrm{Spur}({}_{\alpha\beta}\mathfrak{t}_t'\, {}_{\alpha\beta}\mathfrak{t}_t) \tag{32}$$

From (29) we have,

$${}_\beta g_{\beta'}^{\alpha} I = \sum_p \frac{1}{m_p}\, {}_\beta t_{\beta'}^{\alpha p} I. \tag{31}$$

Now consider the sum of (28). Since ct goes over all values of t in p when C goes over all the operations of the group,

$$\mathrm{Spur}({}_{\alpha\beta}\mathfrak{t}_{ct}'\, {}_{\alpha\beta'}\mathfrak{t}_{ct}) = \mathrm{Spur}({}_{\alpha\beta}\mathfrak{t}_{ct'}'\, {}_{\alpha\beta'}\mathfrak{t}_{ct'})$$

if t and t' are in the same p-th set. Thus (32) becomes

$${}_\beta t_{\beta'}^{\alpha p} d_\alpha = g_p\, \mathrm{Spur}({}_{\alpha\beta}\mathfrak{t}_t\, {}_{\alpha\beta'}\mathfrak{t}_t')$$

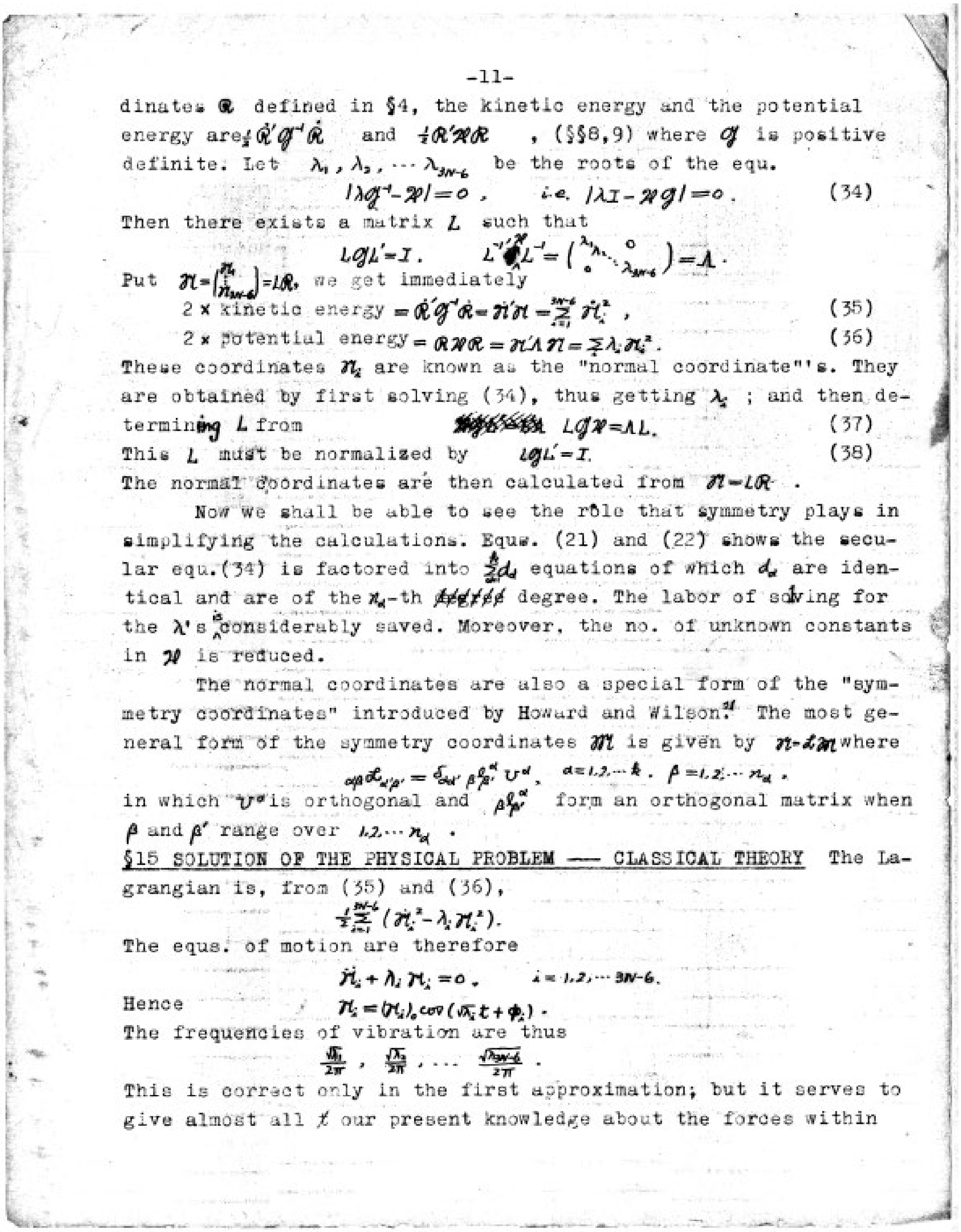
where t represents any nucleus of the p-th set. Substitute into (31) and make use of (24). We obtain:

$${}_\beta g_{\beta'}^{\alpha} = \frac{1}{d_\alpha}\sum_p \frac{g_p}{m_p}\sum_{\tau=1}^{d_\alpha} {}_{\alpha\beta\tau}\vec{\mathfrak{S}}_t \cdot {}_{\alpha\beta'\tau}\vec{\mathfrak{S}}_t. \tag{33}$$

Thus in calculating $\mathfrak{g}$ (${}_{\alpha\beta}\mathfrak{g}_{\alpha'\beta'} = \delta_{\alpha\alpha'}\, {}_\beta g_{\beta'}^{\alpha} I$) we have only to know ${}_{\alpha\beta\tau}\vec{\mathfrak{S}}_t$ for all $\alpha(=1,2,\cdots k)$, $\beta(=1,2,\cdots n_\alpha)$ and $\tau(=1,2,\cdots d_\alpha)$; but only for one t in each p. From (24) we see that not all $\vec{S}$ are needed.

Secular Equation and Degeneracy

§14 NORMAL COORDINATES In terms of the independent reduced coor-



dinates $\mathcal{R}$ defined in §4, the kinetic energy and the potential energy are $\frac{1}{2}\dot{\mathcal{R}}'\mathcal{G}^{-1}\dot{\mathcal{R}}$ and $\frac{1}{2}\mathcal{R}'\mathcal{V}\mathcal{R}$, (§§8,9) where $\mathcal{G}$ is positive definite. Let $\lambda_1, \lambda_2, \cdots \lambda_{3N-6}$ be the roots of the equ.

$$|\lambda\mathcal{G}^{-1}-\mathcal{V}|=0\,, \qquad i.e.\ |\lambda I-\mathcal{V}\mathcal{G}|=0\,. \tag{34}$$

Then there exists a matrix L such that

$$L\mathcal{G}L'=I\,. \qquad L'^{-1}\mathcal{V}L^{-1}=\begin{pmatrix}\lambda_1 & & & 0\\ & \lambda_2 & & \\ & & \ddots & \\ 0 & & & \lambda_{3N-6}\end{pmatrix}=\Lambda\,.$$

Put $\mathcal{N}=\begin{pmatrix}\mathcal{N}_1\\ \vdots\\ \mathcal{N}_{3N-6}\end{pmatrix}=L\mathcal{R}$, we get immediately

$$2\times\text{kinetic energy}=\dot{\mathcal{R}}'\mathcal{G}^{-1}\dot{\mathcal{R}}=\dot{\mathcal{N}}'\dot{\mathcal{N}}=\sum_{i=1}^{3N-6}\dot{\mathcal{N}}_i^2\,, \tag{35}$$

$$2\times\text{potential energy}=\mathcal{R}\mathcal{V}\mathcal{R}=\mathcal{N}'\Lambda\mathcal{N}=\sum_i\lambda_i\mathcal{N}_i^2\,. \tag{36}$$

These coordinates $\mathcal{N}_i$ are known as the "normal coordinate"'s. They are obtained by first solving (34), thus getting λ_i ; and then determining L from

$$L\mathcal{G}\mathcal{V}=\Lambda L\,. \tag{37}$$

This L must be normalized by $\quad L\mathcal{G}L'=I.$ (38)

The normal coordinates are then calculated from $\mathcal{N}=L\mathcal{R}$.

Now we shall be able to see the rôle that symmetry plays in simplifying the calculations. Equs. (21) and (22) shows the secular equ. (34) is factored into $\sum_{\alpha}^{k} d_\alpha$ equations of which d_α are identical and are of the n_α-th degree. The labor of solving for the λ's is considerably saved. Moreover, the no. of unknown constants in $\mathcal{V}$ is reduced.

The normal coordinates are also a special form of the "symmetry coordinates" introduced by Howard and Wilson.[21] The most general form of the symmetry coordinates $\mathfrak{M}$ is given by $\mathcal{N}=\mathcal{L}\mathfrak{M}$ where

$${}_{\alpha\beta}\mathcal{L}_{\alpha'\beta'}=\delta_{\alpha\alpha'}\,{}_{\beta}\ell^{\alpha}_{\beta'}\,U^{\alpha}, \qquad \alpha=1,2,\cdots k\,. \quad \beta=1,2,\cdots n_\alpha\,.$$

in which U^{α} is orthogonal and ${}_{\beta}\ell^{\alpha}_{\beta'}$ form an orthogonal matrix when β and β' range over $1,2,\cdots n_\alpha$.

§15 SOLUTION OF THE PHYSICAL PROBLEM —— CLASSICAL THEORY The Lagrangian is, from (35) and (36),

$$\frac{1}{2}\sum_{i=1}^{3N-6}(\dot{\mathcal{N}}_i^2-\lambda_i\mathcal{N}_i^2).$$

The equs. of motion are therefore

$$\ddot{\mathcal{N}}_i+\lambda_i\mathcal{N}_i=0\,. \qquad i=1,2,\cdots 3N-6.$$

Hence $\quad \mathcal{N}_i=(\mathcal{N}_i)_0\cos(\sqrt{\lambda_i}\,t+\phi_i)\,.$

The frequencies of vibration are thus

$$\frac{\sqrt{\lambda_1}}{2\pi}\,,\ \frac{\sqrt{\lambda_2}}{2\pi}\,,\ \cdots\ \frac{\sqrt{\lambda_{3N-6}}}{2\pi}\,.$$

This is correct only in the first approximation; but it serves to give almost all ~~f~~ our present knowledge about the forces within the molecules.

§16 SOLUTION OF THE PHYSICAL PROBLEM —— QUANTUM THEORY Now

$$T=\frac{1}{2}\sum_i \dot{x}_i^2, \qquad V=\frac{1}{2}\sum_i \lambda_i x_i^2 .$$

The wave equation is most easily obtained from the variational formulation of the problem:

$$\delta\int\left\{\sum \hbar^2 g^{ij}\frac{\partial\psi^*}{\partial x_i}\frac{\partial\psi}{\partial x_j}+(V-E)\psi^*\psi\right\}\sqrt{g}\,dx_1\cdots dx_{3N-6}=0,$$

where g^{ij} is given by $T=g_{ij}\dot{x}^i\dot{x}^j$. Hence

$$\sum_i \frac{1}{2}\left(-\hbar^2\frac{\partial^2}{\partial x_i^2}+\lambda_i x_i^2\right)\psi=E\psi .$$

The electric moment in any direction is, in the first approximation

$$D_0+\sum_{i=1}^{3N-6} D_i x_i$$

where D_i is immediately calculable from (15). Thus the wave function and selection rules, hence the frequencies in the vibrational spectra of the molecule, are the same as those of $3N-6$ independent harmonic oscillators with coordinates $x_1, x_2, \cdots x_{3N-6}$. The frequencies are therefore exactly those obtained classically.

§17 THE DEGREE OF DEGENERACY We have seen that the secular equ. (34) is factorized into d_1 identical equs. of the n_1-th degree, d_2 identical equs. of the n_2-th degree, ... There must therefore be n_1 frequencies each corresponding to d_1 different normal moods of vibration (d_1-fold degenerate). The nos. $d_1, d_2 \ldots$ are accordingly called "the degrees of degeneracy".

It is interesting that they can be determined together with $n_1, n_2, \cdots$ without carrying out the calculation of the normal coordinates. For, by (20)

$$\sum_\alpha n_\alpha \,\mathrm{Spur}(\mathcal{W}_c^\alpha)=\mathrm{Spur}(\mathcal{A}_c).$$

But

$$\sum_c \mathrm{Spur}(\mathcal{W}_c^\alpha)\,\mathrm{Spur}(\mathcal{W}_c^\beta)=\delta_{\alpha\beta}\,h ,$$

where h = order of the group. Hence

$$n_\alpha=\frac{1}{h}\sum_c \mathrm{Spur}(\mathcal{A}_c)\,\mathrm{Spur}(\mathcal{W}_c^\alpha). \qquad (39)$$

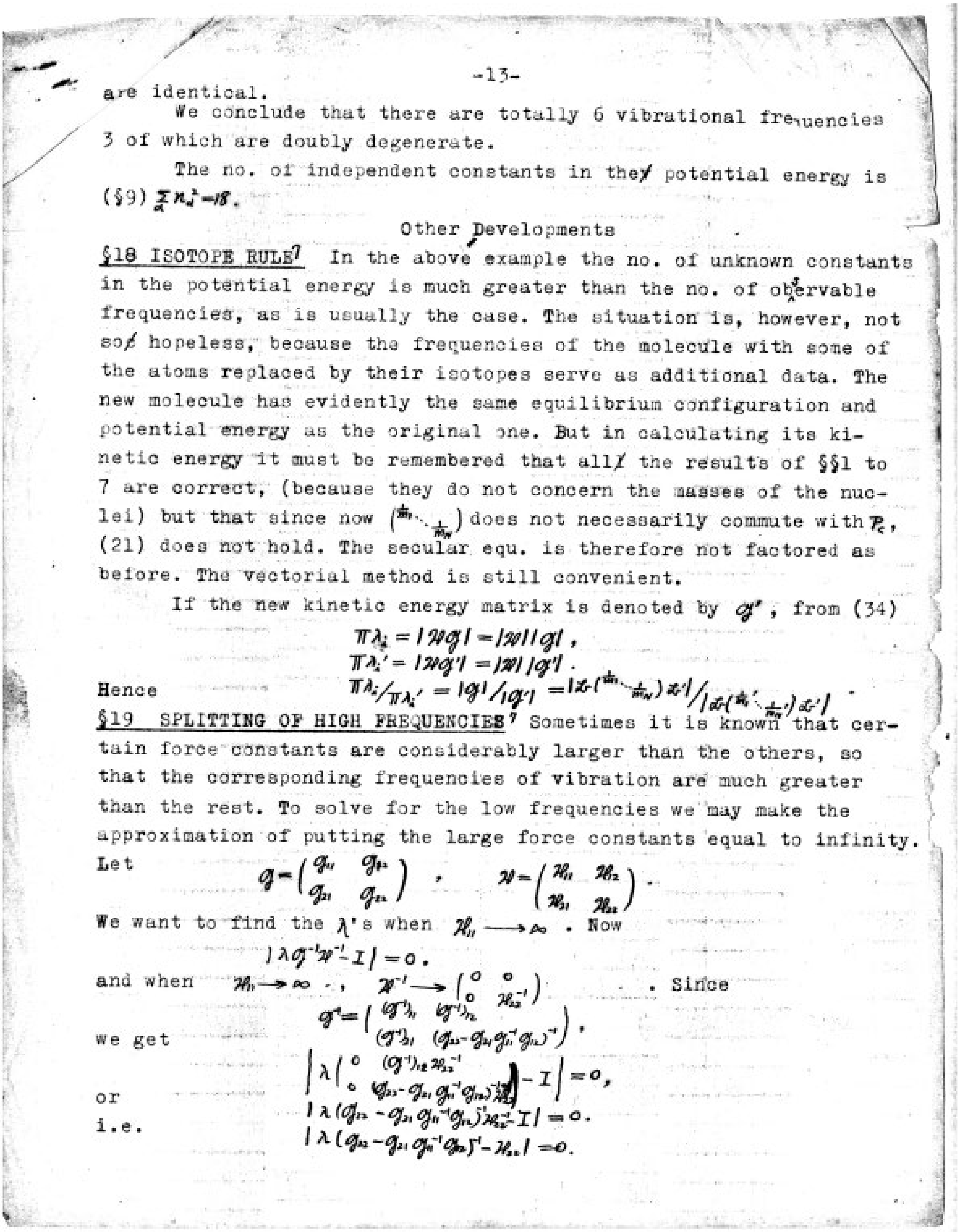
Now both Spur($\mathcal{W}_c^\alpha$) and d_α can be found[3] from a table of characters of the point group, and Spur($\mathcal{A}_c$) can be calculated through the theorem in §5, so that d_α and n_α are easily calculable from (39).

EXAMPLE We consider again the molecule CH_3Cl. (§5) From the left table of characters we get

Representation	d_α	Spur($\mathcal{W}_c$)		
		$c=c_1$	$c=c_2, c_3$	$c=c_4, c_5, c_6$
$\mathcal{W}^1$	1	1	1	1
$\mathcal{W}^2$	1	1	1	-1
$\mathcal{W}^3$	2	2	-1	0

$$n_1=3, \quad n_2=0, \quad n_3=3 .$$

It follows that the secular equ. is factored into 3 cubic equs. 2 of which



are identical.

We conclude that there are totally 6 vibrational frequencies 3 of which are doubly degenerate.

The no. of independent constants in the~~y~~ potential energy is (§9) $\sum_\alpha n_\alpha^2 = 18$.

Other Developments

§18 ISOTOPE RULE[7] In the above example the no. of unknown constants in the potential energy is much greater than the no. of observable frequencies, as is usually the case. The situation is, however, not so hopeless, because the frequencies of the molecule with some of the atoms replaced by their isotopes serve as additional data. The new molecule has evidently the same equilibrium configuration and potential energy as the original one. But in calculating its kinetic energy it must be remembered that all the results of §§1 to 7 are correct, (because they do not concern the masses of the nuclei) but that since now $\left(\frac{1}{m_1} \ddots \frac{1}{m_N}\right)$ does not necessarily commute with P_e, (21) does not hold. The secular equ. is therefore not factored as before. The vectorial method is still convenient.

If the new kinetic energy matrix is denoted by $\mathcal{G}'$, from (34)

$$\prod \lambda_i = |\mathcal{K}\mathcal{G}| = |\mathcal{K}||\mathcal{G}|,$$

$$\prod \lambda_i' = |\mathcal{K}\mathcal{G}'| = |\mathcal{K}||\mathcal{G}'|.$$

Hence

$$\prod \lambda_i \Big/ \prod \lambda_i' = |\mathcal{G}| \big/ |\mathcal{G}'| = \left|\mathcal{L}\left(\frac{1}{m_1} \ddots \frac{1}{m_N}\right)\mathcal{L}'\right| \Big/ \left|\mathcal{L}\left(\frac{1}{m_1'} \ddots \frac{1}{m_N'}\right)\mathcal{L}'\right|.$$

§19 SPLITTING OF HIGH FREQUENCIES[7] Sometimes it is known that certain force constants are considerably larger than the others, so that the corresponding frequencies of vibration are much greater than the rest. To solve for the low frequencies we may make the approximation of putting the large force constants equal to infinity. Let

$$\mathcal{G} = \begin{pmatrix} \mathcal{G}_{11} & \mathcal{G}_{12} \\ \mathcal{G}_{21} & \mathcal{G}_{22} \end{pmatrix}, \qquad \mathcal{K} = \begin{pmatrix} \mathcal{K}_{11} & \mathcal{K}_{12} \\ \mathcal{K}_{21} & \mathcal{K}_{22} \end{pmatrix}.$$

We want to find the λ's when $\mathcal{K}_{11} \longrightarrow \infty$. Now

$$|\lambda \mathcal{G}^{-1}\mathcal{K}^{-1} - I| = 0.$$

and when $\mathcal{K}_{11} \longrightarrow \infty$, $\mathcal{K}^{-1} \longrightarrow \begin{pmatrix} 0 & 0 \\ 0 & \mathcal{K}_{22}^{-1} \end{pmatrix}$. Since

$$\mathcal{G}^{-1} = \begin{pmatrix} (\mathcal{G}^{-1})_{11} & (\mathcal{G}^{-1})_{12} \\ (\mathcal{G}^{-1})_{21} & (\mathcal{G}_{22} - \mathcal{G}_{21}\mathcal{G}_{11}^{-1}\mathcal{G}_{12})^{-1} \end{pmatrix},$$

we get

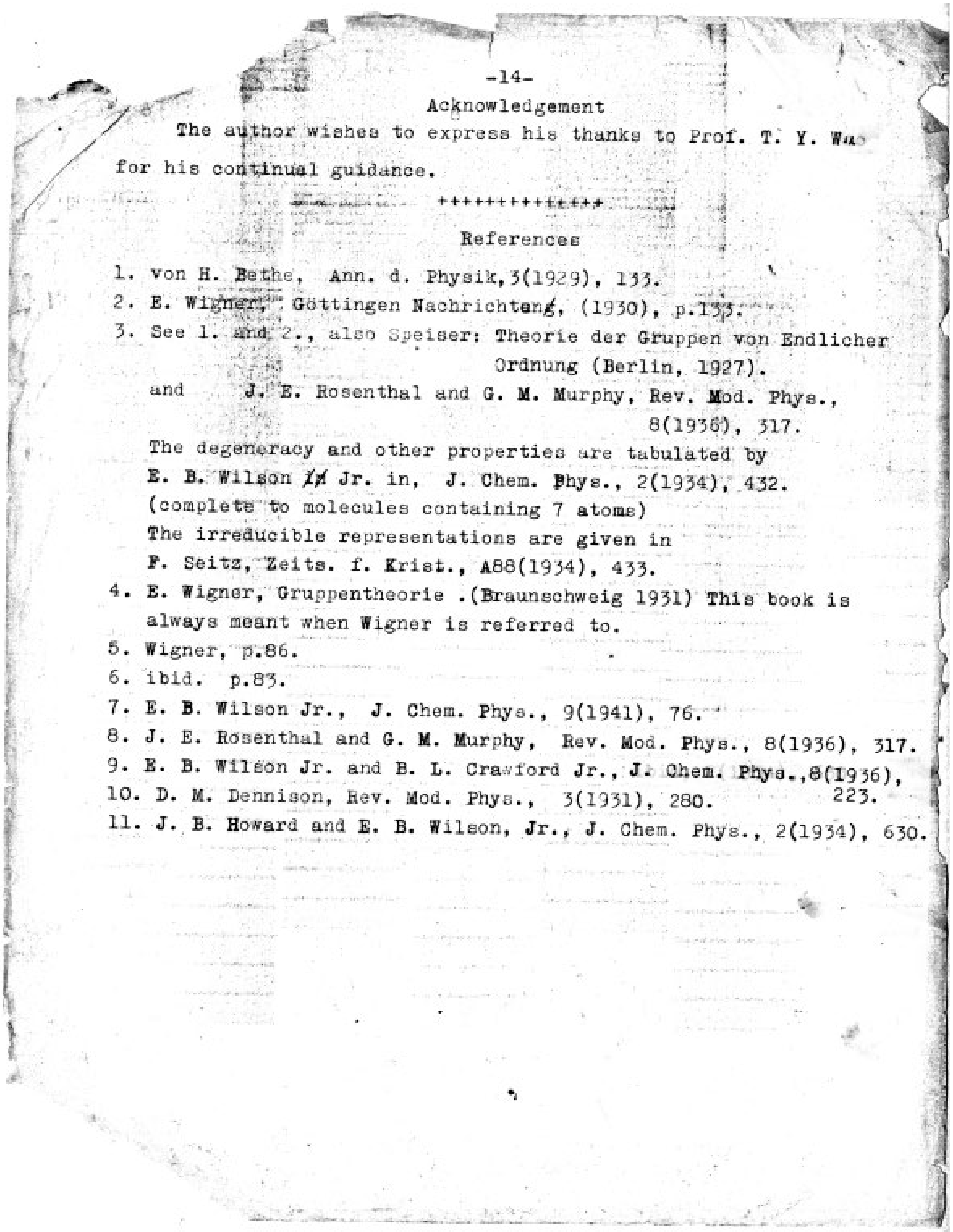
$$\left|\lambda \begin{pmatrix} 0 & (\mathcal{G}^{-1})_{12}\mathcal{K}_{22}^{-1} \\ 0 & (\mathcal{G}_{22} - \mathcal{G}_{21}\mathcal{G}_{11}^{-1}\mathcal{G}_{12})^{-1}\mathcal{K}_{22}^{-1} \end{pmatrix} - I\right| = 0,$$

or

$$|\lambda(\mathcal{G}_{22} - \mathcal{G}_{21}\mathcal{G}_{11}^{-1}\mathcal{G}_{12})^{-1}\mathcal{K}_{22}^{-1} - I| = 0.$$

i.e.

$$|\lambda(\mathcal{G}_{22} - \mathcal{G}_{21}\mathcal{G}_{11}^{-1}\mathcal{G}_{12})^{-1} - \mathcal{K}_{22}| = 0.$$



-14-

Acknowledgement

The author wishes to express his thanks to Prof. T. Y. Wu for his continual guidance.

++++++++++++++

References

1. von H. Bethe, Ann. d. Physik, 3(1929), 133.
2. E. Wigner, Göttingen Nachrichten~~g~~, (1930), p.133.
3. See 1. and 2., also Speiser: Theorie der Gruppen von Endlicher Ordnung (Berlin, 1927).
   and J. E. Rosenthal and G. M. Murphy, Rev. Mod. Phys., 8(1936), 317.
   The degeneracy and other properties are tabulated by E. B. Wilson ~~II~~ Jr. in, J. Chem. Phys., 2(1934), 432.
   (complete to molecules containing 7 atoms)
   The irreducible representations are given in F. Seitz, Zeits. f. Krist., A88(1934), 433.
4. E. Wigner, Gruppentheorie .(Braunschweig 1931) This book is always meant when Wigner is referred to.
5. Wigner, p.86.
6. ibid. p.83.
7. E. B. Wilson Jr., J. Chem. Phys., 9(1941), 76.
8. J. E. Rosenthal and G. M. Murphy, Rev. Mod. Phys., 8(1936), 317.
9. E. B. Wilson Jr. and B. L. Crawford Jr., J. Chem. Phys., 8(1936), 223.
10. D. M. Dennison, Rev. Mod. Phys., 3(1931), 280.
11. J. B. Howard and E. B. Wilson, Jr., J. Chem. Phys., 2(1934), 630.

Investigations in the Statistical Theory of Superlattices

1944 thesis for MSc degree

National Tsing Hua University

INVESTIGATIONS IN THE STATISTICAL THEORY OF SUPERLATTICES

A Dissertation
Submitted to
The Faculty of the Graduate School of Science
in Candidacy for
The Degree of Master of Science

By C. N. Yang (楊振寧)

Kunming, China
June, 1944

CONTENTS

I. The Variation of the Interaction Energy with Change of Lattice Constants and Change of the Degree of Order

II. A Generalization of the Quasi-chemical Method in the Statistical Theory of Superlattices

THE VARIATION OF THE INTERACTION ENERGY WITH CHANGE OF LATTICE CONSTANTS AND CHANGE OF THE DEGREE OF ORDER

By C. N. Yang
National Tsing Hua University,
Kunming, China

ABSTRACT

The change of the lattice constants due to the order-disordering process in a superlattice is investigated by using the condition of minimum free energy in Bethe's theory. It is found that the interaction energy depends on the degree of order when the external pressure is kept constant. The specific heat at constant pressure given by the theory is compared with experiment. Another source of the variation of interaction energy is the change of atomic arrangements. This is also investigated from the view point of Wang's formulation of the free energy in Bethe's approximations.

1. INTRODUCTION

The binary alloy CuAu is face-centred cubic when disordered and tetragonal when ordered. This change of lattice form can be studied thermodynamically if we know the energy and the entropy of the crystal. Some calculations along this line has already been made by Wilson[1] who used Bethe's method to find the energy but Bragg-Williams' method to find the entropy of the crystal. It will be shown in the present paper that Bethe's method can be carried through in the calculations, making it self-consistent. The results are comparable with Gorsky's measurements[2].

The change of lattice constants evidently affects the interaction energy between the atoms, and must consequently produce a change in the configurational energy and the specific heat of the crystal. We shall see that the effect is in the right direction to bring the theory into closer agreement with experiment, because it tends to make the energy increase more rapidly near the critical temperature. An actual calculation of the specific heat at variable lattice constant but constant external pressure for β-brass is given in section 3.

Now the interaction energy can also be influenced by a change of the atomic arrangements. Mott[3] has shown from a study of the electronic distribution in superlattices that the interaction energy decreases as the degree of order decreases. The actual relation between the two is naturally very complicated. A linear dependence (of the average interaction energy upon the degree of order) has been assumed by Lin[4] in attempting to explain the occurrence of the maximum critical temperature of a face-centred alloy at the concentration ratio 1:3. In order to justify the assumption we shall view the problem from a new angle by the introduction of the free energy in Bethe's approximation[5]. In this way it is found that the interaction energy as a function of the degree of order must satisfy certain equations obtained from a set of conditions of consistency. This same set of conditions of consistency makes also possible the calculation of the energy of the crystal without appealing to Bragg-Williams' theory as Lin did.

2. THE VARIATION OF LATTICE CONSTANTS

We shall form the partition function at constant lattice constants l_1 and l_2, and then obtain their equilibrium values from the equations determining the generalized reactions. Let $\frac{1}{2}zNm$ be the number of A-B neighbors in the crystal. If $g(m)$ is the

---

[1] Wilson, Proc. Camb. Phil. Soc. $\underline{34}$, 81(1938).

[2] Gorsky, Zeit. f. Phys. $\underline{50}$, 64(1928).

[3] Mott, Proc. Phys. Soc. $\underline{49}$, 258(1937).

[4] Lin, Chinese J. Phys. $\underline{3}$, 182(1939).

[5] Wang,"Free Energy in the Statistical Theory of Order-Disorder Transformations", Science Report of National Tsing Hua University, series A, 30-th anniversary Memorial Number (1941), printed but failed to appear.

number of arrangements of the atoms for the given value of m, and $W(l_1, l_2, m)$ the configurational energy of the crystal, the configurational partition function is

$$f(m, T, l_1, l_2) = g(m) \exp(-W/kT).$$

The equilibrium value $\bar{m}$ of m is determined from the condition of a maximum of f:

$$\frac{\partial}{\partial \bar{m}} \log f(\bar{m}, T, l_1, l_2) = 0.$$

The generalized reactions are given by

$$\begin{aligned} L_i &= kT \frac{\mathrm{d}}{\mathrm{d}l_i} \log f(\bar{m}, T, l_1, l_2) = kT \frac{\mathrm{d}\bar{m}}{\mathrm{d}l_i} \frac{\partial}{\partial \bar{m}} \log f + kT \frac{\partial}{\partial l_i} \log f = kT \frac{\partial}{\partial l_i} \log f \\ &= -\frac{\partial}{\partial l_i} W(l_1, l_2, \bar{m}). \end{aligned} \tag{1}$$

To study the change of lattice form in CuAu we divide the face-centred lattice into four simple cubic sublattices 1, 2, 3, 4[*]. Let the shortest distance between the sites of 1 and 2, or 3 and 4 be l_1, that between the sites of 1 and 3, 1 and 4, 2 and 3 or 2 and 4 be l_2, so that the former is the distance between neighbouring Au-Au or Cu-Cu atoms and the latter that between neighbouring Au-Cu atoms when the crystal is perfectly ordered. The interaction energies V_{AA}, V_{AB} and V_{BB} are functions of l_1, and l_2.

If the number of sites of each sublattice is $\frac{1}{2}N$, the number of pairs of sites between sublattices 1 and 2 must be $4(\frac{1}{2})N = 2N$. Denote by m_{12} the fraction of A-B pairs among these. Then the number of

$$\begin{aligned} &A\text{-}A \text{ pairs is} \quad \frac{1}{2}\left[4\left(\frac{N}{2}\theta_1 + \frac{N}{2}\theta_2\right) - 2Nm_{12}\right] = N[\theta_1 + \theta_2 - m_{12}], \\ &B\text{-}B \text{ pairs is} \quad \frac{1}{2}\left[4\left(\frac{N}{2}\{1-\theta_1\} + \frac{N}{2}\{1-\theta_2\}\right) - 2Nm_{12}\right] = N[2 - \theta_1 - \theta_2 - m_{12}], \end{aligned}$$

where θ_i is the fraction of sites of sublattice i occupied by A atoms. Thus the energy of interaction between the atoms on sublattices 1 and 2 is

$$N[(\theta_1 + \theta_2 - m_{12})V_{AA}(l_1) + 2m_{12}V_{AB}(l_1) + (2 - \theta_1 - \theta_2 - m_{12})V_{BB}(l_1)].$$

Writing

$$c = \frac{1}{4}(\theta_1 + \theta_2 + \theta_3 + \theta_4) \qquad \text{and} \qquad V = \frac{1}{2}(V_{AA} + V_{BB}) - V_{AB},$$

we get the energy of the whole crystal

$$\begin{aligned} W = {} & N[4cV_{AA}(l_1) + 4(1-c)V_{BB}(l_1) - 2(m_{12} + m_{34})V(l_1)] + \\ & N[8cV_{AA}(l_2) + 8(1-c)V_{BB}(l_2) - 2(m_{13} + m_{14} + m_{23} + m_{24})V(l_2)]. \end{aligned} \tag{2}$$

---

[*] Cf. Fig.27 in Rev. Mod. Phys. 10, 1(1938).

With this value for W, (1) becomes

$$L_1 = -N[4cV'_{AA}(l_1) + 4(1-c)V'_{BB}(l_1) - 2(\bar{m}_{12} + \bar{m}_{34})V'(l_1)], \tag{3}$$

and

$$L_2 = -N[8cV'_{AA}(l_2) + 8(1-c)V'_{BB}(l_2) - 2(\bar{m}_{13} + \bar{m}_{14} + \bar{m}_{23} + \bar{m}_{24})V'(l_2)]. \tag{4}$$

To solve for l_1 and l_2 as functions of T we must first know the $\bar{m}$'s, which are usually very complicated. Wilson[1] discussed the values of l_1 and l_2 only in the cases when the alloy is disordered and when the order is nearly perfect. We shall also confine our attention to these cases.

(i) Disordered. In this case there is no difference between the four sublattices so that all the $\bar{m}_{ij}$'s are equal to $\bar{m}$. (3) and (4) reduce to

$$\begin{aligned} L_1 &= -N[4cV'_{AA}(l_1) + 4(1-c)V'_{BB}(l_1) - 4\bar{m}V'(l_1)], \\ L_2 &= -N[8cV'_{AA}(l_2) + 8(1-c)V'_{BB}(l_2) - 8\bar{m}V'(l_2)]. \end{aligned} \tag{5}$$

If $L_1 = L_2 = 0$, this shows that $l_1 = l_2$, so that the crystal is cubic.

(ii) Order nearly perfect. When $c = \frac{1}{2}$, and the order is nearly perfect,

$$\theta_1 = \theta_2 \cong 1, \quad \theta_3 = \theta_4 \cong 0, \qquad \theta_1 + \theta_3 = 1, \quad \theta_1 - \theta_3 = s.$$

There are only a few B atoms on sublattices 1 and 2. Hence approximately

$$\bar{m}_{12} = (1-\theta_1) + (1-\theta_2) = 2\theta_3 = 1 - s.$$

By the same reason we can obtain the number of A-A pairs of neighbours between the sublattices 1 and 3:

$$N(\theta_1 + \theta_3 - \bar{m}_{13}) = 4\left(\frac{1}{2}N\theta_3\right).$$

Thus

$$\bar{m}_{13} = \theta_1 - \theta_3 = s.$$

We can now write down all the $\bar{m}$'s:

$$\bar{m}_{12} = \bar{m}_{34} = 1 - s, \quad \bar{m}_{13} = \bar{m}_{23} = \bar{m}_{14} = \bar{m}_{24} = s.$$

These equations are correct to the first order of $(1-s)$. Substituting them into (3) and (4) we obtain

$$\begin{aligned} L_1 &= -2N[V'_{AA}(l_1) + V'_{BB}(l_1) - 2(1-s)V'(l)], \\ L_2 &= -4N[V'_{AA}(l_2) + V'_{BB}(l_2) - 2sV'(l_2)]. \end{aligned}$$

These are exactly equations (27) in Wilson's paper, from which an expression of the degree of tetragonality in agreement with Gorsky's measurements[2] can be obtained.

3. THE EFFECT OF THE CHANGE OF LATTICE CONSTANTS ON THE INTERACTION ENERGY

In the alloy CuAu the gold atoms and the copper atoms are in contact when the order is perfect. Since the copper atom is somewhat smaller than the gold atom, the size of the crystal must increase when gold atoms exchange their positions with copper atoms. Thus with increasing disorder the distance between the atoms increases and hence the interaction energies diminish. The disordering process is therefore effected with more ease near the critical temperature than it is at lower temperatures; and we expect the specific heat at constant pressure to possess a steeper and higher maximum at the critical temperature than the specific heat at constant volume.

Now we shall calculate in length the specific heat at constant pressure of the alloy β-brass, which forms the simplest type of superlattice that can be studied statistically. Bethe's method will be used.

The configurational energy of the crystal is, in Easthope's[6] notations:

$$W = -N_{AB}V + \frac{1}{2}Nz[c(V_{AA} - V_{BB}) + V_{BB}]. \tag{6}$$

Substitution of this expression into (1) gives

$$0 = -\bar{m}V'(l) + c[V'_{AA}(l) - V'_{BB}(l)] + V'_{BB}(l),$$

when the pressure is put equal to zero. Now the variation of V is not very large, so that to a sufficient approximation we may assume the linear relations

$$[V'_{AA}(l) - V'_{BB}(l)]/V'(l) = -K_0 + K_1 V, \tag{7}$$

and

$$V'_{BB}(l)/V'(l) = -J_0 + J_1 V. \tag{8}$$

These three last equations give, after eliminating $V'_{AA}(l)$ and $V'_{BB}(l)$:

$$V = \frac{\bar{m} + (cK_0 + J_0)}{cK_1 + J_1}. \tag{9}$$

We have already seen that V increases as $\bar{m}$ increases, hence $cK_1 + J_1$ must be positive. The other constant $cK_0 + J_0$ must also be positive in order that V may be positive with only a relatively small variation.

Eisenschitz[7] has calculated the specific heat at constant pressure by Bragg-Williams' method. He assumed that the interaction energies depend on a parameter u in the following way:

$$\frac{1}{2}(V_{AA} + V_{BB}) = \phi[(1-a) + a(1-u)^2], \qquad V_{AB} = \phi b u^2,$$

---

[6] Easthope, Proc. Camb. Phil. Soc. 33, 502(1937); 34, 68(1938).

[7] Eisenschitz, Proc. Roy. Soc. 68, 546(1938).

where $a = .225$, $b = .203$ and u is of the order of unity. Comparing this with (7) and (8) we see that his assumption is equivalent (approximately) to ours if $\frac{1}{2}K_0 + J_0 = 1.22$ and $\frac{1}{2}K_1 + J_1 = .508 \times 10^{14}\mathrm{erg}^{-1}$. But with these values the specific heat at the critical temperature would be too large. In order to make $(C_p)_{T_c} = 5.1R$ as given by the measurements of Sykes and Wilkinson[8] we assume (cf. eq.(12) below)

$$\frac{1}{2}K_0 + J_0 = 1.79.$$

With this value for $\frac{1}{2}K_0 + J_0$, the relative variation of V can be shown to be within 1.3%.

We can now start from (9) and the equations given by Easthope[6] for the determination of $\bar{m}$ as a function of the temperature and V to obtain the specific heat at constant pressure:

$$C_p = \frac{\mathrm{d}W}{\mathrm{d}T} = \left(\frac{\partial W}{\partial \bar{m}}\right)_l \frac{\mathrm{d}\bar{m}}{\mathrm{d}x}\frac{\mathrm{d}x}{\mathrm{d}T} + \left(\frac{\partial W}{\partial l}\right)_{\bar{m}} \frac{\mathrm{d}l}{\mathrm{d}T} = \left(\frac{\partial W}{\partial \bar{m}}\right)_l \frac{\mathrm{d}\bar{m}}{\partial x}\frac{\mathrm{d}x}{\mathrm{d}T}. \tag{10}$$

But

$$\frac{\mathrm{d}x}{\mathrm{d}T} = \frac{xV/kT^2}{1 + \frac{x}{kT}\frac{\mathrm{d}V}{\mathrm{d}\bar{m}}\frac{\mathrm{d}\bar{m}}{\mathrm{d}x}}. \tag{11}$$

Hence by (6)

$$C_p = \frac{\frac{1}{2}Nzkx(\log x)^2(-\frac{\mathrm{d}\bar{m}}{\mathrm{d}x})}{1 + \frac{x\log x}{\bar{m}+(cK_0+J_0)}(-\frac{\mathrm{d}\bar{m}}{\mathrm{d}x})}. \tag{12}$$

The value of this expression is calculated for the case $c = \frac{1}{2}$, the constant $\frac{1}{2}K_0 + J_0$ being assumed to be 1.79 to make $(C_p)_{T_c} = 5.1R$. The result is plotted in the accompanying figure together with Bethe's curve[9] for C_V and Sykes ad Wilkinson's experimental[8] data.

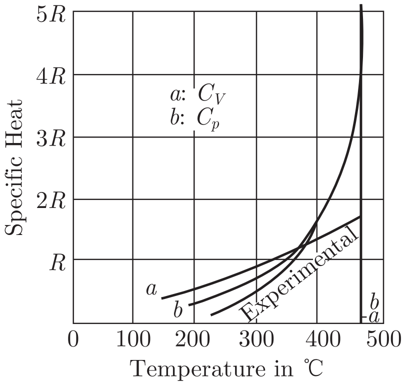


Fig. 1　Configurational specific heat of β-brass.

---

[8] Sykes and Wilkinson, Inst. Metals J. 61, 223(1937).

[9] Nix and Shockley, Rev. Mod. Phys. 10, 1(1938).

4. THE EFFECT OF THE ATOMIC DISTRIBUTION ON THE INTERACTION ENERGY

As has already been mentioned, the interaction energy depends in some very complicated manner upon the degree of order. To study the effect of such a dependence Lin[4] has assumed a linear relationship:

$$V = V_0(1 + \alpha c + \beta m_{AA}) \tag{13}$$

between the interaction energy: V and the fraction of A-A pairs of neighbours: m_{AA}. In this section we shall study the general nature of the variation of V in the light of the theory of the free energy in Bethe's approximation given by Wang[5].

The fundamental equations in Wang's paper are (45), (46) and (39) with ξ_α and ξ_β given by (47), (48), (49) and (50). These equations are still assumed to be valid now V becomes a function of θ_α, θ_β and T. They may be put into the form:

$$\frac{\mathrm{d}}{\mathrm{d}\theta_\alpha}\log Q = -Nr_\alpha \log\xi_\alpha = \frac{\partial}{\partial\theta_\alpha}\log Q_0(\theta_\alpha, \theta_\beta, V/T), \tag{14}$$

$$\frac{\mathrm{d}}{\mathrm{d}\theta_\beta}\log Q = -Nr_\beta \log\xi_\beta = \frac{\partial}{\partial\theta_\beta}\log Q_0(\theta_\alpha, \theta_\beta, V/T), \tag{15}$$

$$E = kT^2\frac{\mathrm{d}}{\mathrm{d}T}\log Q, \tag{16}$$

where Q_0 is the partition function for the case when V is a constant, if we denote by $\frac{\partial}{\partial\theta_\alpha}$, $\frac{\partial}{\partial\theta_\beta}$, $\frac{\partial}{\partial T}$ and $\frac{\partial}{\partial V}$ differentiations when θ_α, θ_β, T and V are regarded as independent of each other, and by $\frac{\mathrm{d}}{\mathrm{d}\theta_\alpha}$ the operator $\frac{\partial}{\partial\theta_\alpha} + \frac{\mathrm{d}V}{\mathrm{d}\theta_\alpha}\frac{\partial}{\partial V}$. (14) and (15) mean that we have assumed with Lin that the equilibrium values of θ_α and θ_β are given by the same equations as in Bethe's approximation. (16) gives the energy of the crystal.

Consistency of (14) and (15) requires

$$\frac{\mathrm{d}}{\mathrm{d}\theta_\alpha}\left(\frac{\partial}{\partial\theta_\beta}\log Q_0\right) = \frac{\mathrm{d}}{\mathrm{d}\theta_\beta}\left(\frac{\partial}{\partial\theta_\alpha}\log Q_0\right),$$

i.e.

$$\frac{\mathrm{d}V}{\mathrm{d}\theta_\alpha}\frac{\partial^2}{\partial\theta_\beta\partial V}\log Q_0 = \frac{\mathrm{d}V}{\mathrm{d}\theta_\beta}\frac{\partial^2}{\partial V\partial\theta_\alpha}\log Q_0. \tag{17}$$

Upon the hypothesis of nearest neighbour interaction the energy in Wang's paper becomes

$$E_0 = kT^2\frac{\partial}{\partial T}\log Q_0 = zNVm_{AA}.$$

But Q_0 depends on T and V through V/T, so that

$$\frac{\partial^2}{\partial\theta_\alpha\partial V}\log Q_0 = -\frac{T}{V}\frac{\partial^2}{\partial T\partial\theta_\alpha}\log Q_0 = -\frac{\partial E_0}{\partial\theta_\alpha}/kTV.$$

Substituting this eq. and a similar one into (17) we obtain

$$\frac{\partial E_0}{\partial \theta_\beta}\frac{\mathrm{d}V}{\mathrm{d}\theta_\alpha} = \frac{\partial E_0}{\partial \theta_\alpha}\frac{\mathrm{d}V}{\mathrm{d}\theta_\beta}, \tag{18}$$

which becomes, if $\frac{\partial E_0}{\partial V}\frac{\mathrm{d}V}{\mathrm{d}\theta_\alpha}\frac{\mathrm{d}V}{\mathrm{d}\theta_\beta}$ is added to both sides,

$$\frac{\mathrm{d}E_0}{\mathrm{d}\theta_\alpha}\frac{\mathrm{d}V}{\mathrm{d}\theta_\beta} = \frac{\mathrm{d}E_0}{\mathrm{d}\theta_\beta}\frac{\mathrm{d}V}{\mathrm{d}\theta_\alpha}.$$

This shows that V and E_0 are connected by a relation independent of θ_α and θ_β, i.e.

$$V = V(E_0, T). \tag{19}$$

A consequence of this result is that the coefficient α in Lin's relation (13) must be zero. This makes, however, the maximum critical temperature for the AB type of superlattice to shift to a value of the concentration different from $\frac{1}{2}$, which contradicts experimental results. One way out of the difficulty is to make some other assumption regarding the dependence of V on the degree of order, such as

$$V = V_0\left[1 + \beta m_{AA} - \frac{\beta(z-1)}{2c(cz-1)}m_{AA}^2\right]. \tag{20}$$

Let us now try to find E in the general case. From (14) and (16) we get

$$\frac{\mathrm{d}E}{\mathrm{d}\theta_\alpha} = kT^2\frac{\mathrm{d}}{\mathrm{d}T}\left(\frac{\partial}{\partial\theta_\alpha}\log Q_0\right) = kT^2\frac{\mathrm{d}(\frac{V}{T})}{\mathrm{d}T}\frac{\partial^2}{\partial\theta_\alpha\partial(\frac{V}{T})}\log Q_0 = -\frac{T^2}{V}\frac{\mathrm{d}(\frac{V}{T})}{\mathrm{d}T}\frac{\partial E_0}{\partial\theta_\alpha}. \tag{21}$$

Similarly

$$\frac{\mathrm{d}E}{\mathrm{d}\theta_\beta} = -\frac{T^2}{V}\frac{\mathrm{d}(\frac{V}{T})}{\mathrm{d}T}\frac{\partial E_0}{\partial\theta_\beta}.$$

Hence

$$\frac{\mathrm{d}E}{\mathrm{d}\theta_\alpha}\frac{\partial E_0}{\partial\theta_\beta} = \frac{\partial E_0}{\partial\theta_\alpha}\frac{\mathrm{d}E}{\mathrm{d}\theta_\beta}.$$

Just as (18) leads to (19), this last equation leads to

$$E = E(E_0, T).$$

Substituting this into (21) we obtain

$$\frac{\partial E}{\partial E_0}\left(\frac{\partial E_0}{\partial\theta_\alpha} + \frac{\partial E_0}{\partial V}\frac{\partial V}{\partial\theta_\alpha}\right) = \frac{\partial E_0}{\partial\theta_\alpha}\left(1 - \frac{T}{V}\frac{\mathrm{d}V}{\mathrm{d}T}\right).$$

But (19) gives

$$\frac{\mathrm{d}V}{\mathrm{d}\theta_\alpha} = \frac{\dfrac{\partial E}{\partial\theta_\alpha}\dfrac{\partial V}{\partial E_0}}{1 - \dfrac{\partial V}{\partial E_0}\dfrac{\partial E_0}{\partial V}} \quad \text{and} \quad \frac{\mathrm{d}V}{\mathrm{d}T} = \frac{\dfrac{\partial V}{\partial T} + \dfrac{\partial V}{\partial E_0}\dfrac{\partial E_0}{\partial T}}{1 - \dfrac{\partial V}{\partial E_0}\dfrac{\partial E_0}{\partial V}}.$$

Hence

$$\frac{\partial E}{\partial E_0}=\left(1-\frac{\partial V}{\partial E_0}\frac{\partial E_0}{\partial V}\right)\left(1-\frac{T}{V}\frac{\mathrm{d}V}{\mathrm{d}T}\right)=1-\frac{\partial V}{\partial E_0}\frac{\partial E_0}{\partial V}-\frac{T}{V}\frac{\mathrm{d}V}{\mathrm{d}T}-\frac{T}{V}\frac{\partial V}{\partial E_0}\frac{\partial E_0}{\partial T}.$$

Now E_0/V is a function of θ_α, θ_β and V/T, so that

$$\frac{\partial E_0}{\partial V}=-\frac{T}{V}\frac{\partial E_0}{\partial T}+\frac{\partial E_0}{\partial V}.$$

Hence

$$\frac{\partial E}{\partial E_0}=1-\frac{E_0}{V}\frac{\partial V}{\partial E_0}-\frac{T}{V}\frac{\partial V}{\partial T}. \tag{22}$$

If V depends on m_{AA} only, and not on T, we have

$$\frac{\partial E}{\partial E_0}=1-\frac{E_0}{V}\frac{\mathrm{d}V}{\mathrm{d}E_0}=V\frac{\mathrm{d}(E_0/V)}{\mathrm{d}E_0}.$$

The boundary condition is given by the case when there is no A atoms, i.e. when $\theta_\alpha+\theta_\beta=0$. In this case $m_{AA}=0$, $E_0=E=0$. Hence

$$E=\int_0^{E_0}\left[V\frac{\mathrm{d}(E_0/V)}{\mathrm{d}E_0}\right]\mathrm{d}E_0=zN\int_0^{m_{AA}}V\mathrm{d}m_{AA}. \tag{23}$$

If the assumption (20) is made, the energy of the crystal is

$$E=zNV_0\left[m_{AA}+\frac{\beta}{2}m_{AA}^2-\frac{\beta(z-1)}{6c(cz-1)}m_{AA}^3\right].$$

When β is not large, this differs very little in numerical value from Bethe's original expression. The specific heat is

$$C_V=zNV\frac{\mathrm{d}m_{AA}}{\mathrm{d}T}=zNV\frac{\mathrm{d}m_{AA}}{\mathrm{d}x}\frac{\frac{xV}{kT^2}}{1+\frac{x}{kT}\frac{\mathrm{d}V}{\mathrm{d}x}}=zNkx(\log x)^2\frac{\mathrm{d}m_{AA}}{\mathrm{d}x}\frac{1}{1+\frac{x}{kT}\frac{\mathrm{d}V}{\mathrm{d}m_{AA}}\frac{\mathrm{d}m_{AA}}{\mathrm{d}x}}.$$

The author is very much indebted to Prof. J. S. Wang for suggesting this problem and for helpful discussions.

A GENERALIZATION OF THE QUASI-CHEMICAL METHOD IN THE STATISTICAL THEORY OF SUPERLATTICES

By C. N. Yang*
National Tsing Hua University,
Kunming, China

ABSTRACT

The quasi-chemical method introduced by Fowler and Guggenheim for the equilibrium distribution of pairs of sites in a superlattice is generalized. It is shown that by considering groups containing large numbers of sites the method may be used to obtain successive approximations of the free energy of the crystal. To analyze the fundamental assumption underlying the method more closely, the hypothesis of the non-interference of local configurations is discussed. The free energy of the crystal is obtained without integration as a closed algebraic expression with the aid of a Legendre transformation. Applications of the results are then made to different approximations for simple and body-centred cubic crystals and for the face-centred cubic crystal Cu_3Au. In each case the free energy is obtained and discussed.

---

*Research Fellow of the China Foundation for the Promotion of Education and Culture.

1. INTRODUCTION

It was shown by Fowler and Guggenheim[1] that the quasi-chemical method, originally devised for the theory of regular solutions, applies equally well to the theory of superlattices with long-distance order. The method is, as they have emphasized, definitely one stage further towards an exact theory than Bragg-Williams' method[2]. When compared with Bethe's[3] or Kirkwood's[4] method it also distinguishes itself in mathematical simplicity. But to be a method that may lead to a consistent scheme of successive approximations, it must be applicable to the n-th approximation in the theory of superlattices. This does not, however, seem possible in the original form of the method given by Fowler and Guggenheim. It is the purpose of the present paper to formulate the quasi-chemical method in a new way which is applicable to high order approximations in the theory of superlattices.

The free energy expression in Bethe's and in the quasi-chemical methods involves an integral. Its evaluation is very complicated and has been carried out[1,5] so far only in Bethe's approximation for simple and body-centered cubic crystals. In the new formulation of the quasi-chemical method, however, it will be shown that a Legendre transformation helps much in avoiding the mathematical difficulties. (It might be noticed that a similar Legendre transformation is used to essentially the same effect in Fowler's formulation of general statistical mechanics. Cf. Fowler, Statistical Mechanics, second edition, p.188.) The free energy is obtained directly as a closed algebraic expression. Its values are given for Bethe's first and second (modified) approximations and for the face-centered alloy Cu_3Au in sections 7 and 8.

To make sure that the quasi-chemical method may actually be used to obtain a series of successively better approximations, we must investigate the free energy in high order approximations and compare it with the partition function of the crystal. This is done in section 5 together with a comparison of the quasi-chemical and Bethe's methods.

Except in the last section we are only concerned with binary alloys with atomic ratio 1:1 forming a (quadratic), simple cubic or body-centered cubic lattice. The generalization of the method to the investigation of alloys with other atomic ratios and forming other types of lattices is easy. In fact, the superior power of the quasi-

---

[1] Fowler and Guggenheim, Proc. Roy. Soc. A174, 189 (1940).

[2] Bragg and Williams, Proc. Roy. Soc. A145, 699 (1934); 151, 540 (1935); 152, 231 (1935).

[3] Bethe, Proc. Roy. Soc. A150, 552 (1935).

[4] Kirkwood, J. Chem. Phys. 6, 70 (1938).

[5] Chang, Proc. Camb. Phil. Soc. 35, 265 (1939); Kirkwood, J. Chem. Phys. 8, 623 (1940); Wang, "Free Energy in the Statistical Theory of Order–Disorder Transformation", Science Report of National Tsing Hua University, Series A, 30-th Anniversary Memorial Number (1941), printed but failed to appear.

chemical method appears to be even more fully revealed when a face-centered lattice is treated. This problem is taken at the end of the paper where an approximate free energy expression for Cu_3Au is obtained and its critical phenomena discussed.

2. REFORMULATION OF THE QUASI-CHEMICAL METHOD

Consider a crystal AB. Let $2N$ be the total number of atoms and z the number of nearest neighbours of each. At low temperatures we can distinguish between the so-called α-sites and β-sites for A and B atoms respectively. Denote by Nr the number of A atoms on α-sites. Let

$$1 - r = w, \qquad r - w = s.$$

The partition function of the crystal is

$$\sum_r p(r, T), \tag{1}$$

where $p(r, T)$ is equal to $\sum \exp(-E/kT)$ over all possible configurations of the crystal with the given value of r. The average energy over all these configurations is

$$\overline{E}(r, T) = kT^2 \frac{\partial}{\partial T} \log p(r, T). \tag{2}$$

But evidently,

$$p(r, \infty) = g(r) = \left[\frac{N!}{(Nr)!(N\omega)!}\right]^2. \tag{3}$$

Hence

$$\log p(r, T) = \log g(r) + \int_\infty^T \frac{1}{kT^2} \overline{E}(r, T) \mathrm{d}T. \tag{4}$$

The problem therefore reduces to one of finding $\overline{E}(r, T)$. Since a direct solution is very difficult we shall try to find an approximate solution by the quasi-chemical method, which is presented below in a form easily generalizable.

There are in the crystal zN nearest pairs of sites α-β. Among these let there be $[q_\alpha, q_\beta]$ with their α-sites occupied by q_α $(= 0, 1)$ wrong (B) atoms and their β-sites by q_β $(= 0, 1)$ wrong (A) atoms. For given r the following relations hold:

$$\begin{aligned} [0,1] + [0,0] + [1,1] + [1,0] &= zN, \\ [1,1] + [1,0] &= zNw, \\ [0,1] \qquad\quad + [1,1] \qquad\quad &= zNw. \end{aligned} \tag{5}$$

Upon the approximation of neighbour interaction the energy of the crystal may be written as

$$E(r, T) = [0,1]V_{AA} + [0,0]V_{AB} + [1,1]V_{AB} + [1,0]V_{BB}, \tag{6}$$

where the V's are the interaction energies between a pair of nearest neighbours.

We may give (5) and (6) a different interpretation by imagining [0,1], [0,0], [1,1], [1,0] and V_{AA}, V_{AB}, V_{AB}, V_{BB} to be respectively the numbers and the molecular internal energies of the four different kinds of molecules XZ, X, XYZ, XY of a gaseous assembly. The interpretation of (6) is that the assembly has the same internal (non-kinetic) energy as the crystal at the given value of r. (5) would mean that there are altogether zN X atoms, zNw Y atoms and zNw Z atoms in the assembly.

The quasi-chemical method consists in taking the averages $\overline{[0,1]}, \overline{[0,0]}, \overline{[1,1]}, \overline{[1,0]}$ of the assembly at any temperature as approximately representing the corresponding averages of the crystal at the same temperature. Whether this approximation is good can only be judged for the present from the results it leads to.

A detailed treatment of the problem of a gaseous assembly has been given by Fowler[6]. We are only interested in our assembly of four different kinds of molecules, for which the results may be summarized as:

$$\begin{aligned} \overline{[0,1]} &= \xi\nu \mathrm{e}^{-V_{AA}/kT}, \qquad & \overline{[0,0]} &= \xi \mathrm{e}^{-V_{AB}/kT}, \\ \overline{[1,1]} &= \xi\mu\nu \mathrm{e}^{-V_{AB}/kT}, \qquad & \overline{[1,0]} &= \xi\mu \mathrm{e}^{-V_{BB}/kT}, \end{aligned} \tag{7}$$

where ξ, μ and ν are to be determined from (6). From (7) we get

$$\frac{\overline{[0,0]}\,\overline{[1,1]}}{\overline{[0,1]}\,\overline{[1,0]}} = x^{-2}, \tag{8}$$

where

$$x = \exp\left[-\frac{1}{2}(V_{AA} + V_{BB} - 2V_{AB})/kT\right]. \tag{9}$$

(8) and (6) together form the starting point of Fowler and Guggenheim's work[1]. The subsequent calculations of $\overline{E}(r,T)$, $p(r,T)$ and the free energy of the crystal are straightforward and will not be repeated here. We shall see later how the free energy can be written down directly without actual integration.

3. GENERALIZATION TO GROUPS OF FOUR SITES

So far we have fixed our attention on the pairs of nearest neighbours in the crystal and have used the quasi-chemical method to obtain the average numbers of the four different kinds of pairs. Now we shall generalize the whole procedure: we shall study all the groups of sites of an arbitrarily chosen form in the crystal, and classifying these groups according to the way they are occupied by atoms we shall obtain the average number of groups in each class by chemical analogy.

---

[6] Fowler, Statistical Mechanics, second edition, pp. 162–163.

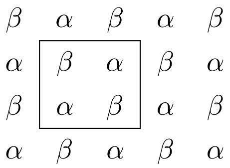


Fig. 1

To make this clear let us consider in detail groups of four sites forming squares (as shown) in a quadratic lattice. We classify these groups into $2^4 = 16$ classes denoted by (0,0,0,0), (0,0,0,1),$\cdots$, (1,1,1,1) respectively, so that all groups in the class (q_1, q_2, q_3, q_4) have q_1 wrong atoms in their upper α-sites, q_2 wrong atoms in their lower α-sites, q_3 wrong atoms in their upper β-sites and q_4 wrong atoms in their lower β-sites. The total number of these groups is N. Hence

$$\sum_{q_i=0}^{1} [q_1, q_2, q_3, q_4] = N, \tag{10}$$

where $[q_1, q_2, q_3, q_4]$ is an abbreviation for the number of groups in the class (q_1, q_2, q_3, q_4). Now the number of all those groups in the crystal with a wrong (B) atom on the upper α-site is just the number of B atoms on the α-sites. Hence

$$\sum_{q} q_i [q_1, q_2, q_3, q_4] = Nw, \qquad i = 1, 2, 3, 4. \tag{11}$$

Let $\chi(q_1, q_2, q_3, q_4)$ be the energy of each group in the class (q_1, q_2, q_3, q_4). It is easy to show that the total energy of the crystal is

$$\overline{E}(r, T) = \sum_{q} \overline{[q_1, q_2, q_3, q_4]}\, \chi(q_1, q_2, q_3, q_4). \tag{12}$$

We may give (10), (11) and (12) an interpretation similar to the one given in section 2 for equations (5) and (6). The same quasi-chemical method used there to obtain (7) leads now to the following averages (approximate) at a given value of r:

$$\overline{[q_1, q_2, q_3, q_4]} = \xi {\mu_1}^{q_1} {\mu_2}^{q_2} {\mu_3}^{q_3} {\mu_4}^{q_4} \mathrm{e}^{-\chi(q_1,q_2,q_3,q_4)/kT}. \tag{13}$$

In this expression the parameters ξ, μ_1, μ_2, μ_3 and μ_4 are to be determined from (10) and (11), which may be written in the form

$$\xi \frac{\partial \phi}{\partial \xi} = N, \qquad \mu_i \frac{\partial \phi}{\partial \mu_i} = Nw \quad (i = 1, 2, 3, 4), \tag{14}$$

if we put

$$\phi(\xi, \mu_1, \mu_2, \mu_3, \mu_4) = \sum_{q} \xi {\mu_1}^{q_1} {\mu_2}^{q_2} {\mu_3}^{q_3} {\mu_4}^{q_4} \mathrm{e}^{-\chi(q_1,q_2,q_3,q_4)/kT}. \tag{15}$$

Or again, in the form

$$\frac{\partial\Psi}{\partial\log\xi}=\frac{\partial\Psi}{\partial\log\mu_1}=\frac{\partial\Psi}{\partial\log\mu_2}=\frac{\partial\Psi}{\partial\log\mu_3}=\frac{\partial\Psi}{\partial\log\mu_4}=0, \tag{16}$$

if we put

$$\Psi=-N\log\xi-\sum_i Nw\log\mu_i+\phi. \tag{17}$$

It can be shown[7] that ξ and μ_i are uniquely determined by (16) at given r and T. Their values at $T=\infty$ are

$$(\xi)_{T=\infty}=Nr^4, \qquad (\mu_i)_{T=\infty}=\frac{w}{r}, \quad i=1,2,3,4, \tag{18}$$

as can be verified by substitution into (14).

To calculate the free energy it is necessary first to evaluate the integral in (4). We shall show that this can be done without first solving (14) for ξ and μ_i. For, by (12) and (13) the integrand may be written

$$\frac{1}{kT^2}\overline{E}(r,T)=\frac{1}{kT^2}\sum_q\overline{[q_1,q_2,q_3,q_4]}\,\chi(q_1,q_2,q_3,q_4)=\frac{\partial\phi}{\partial T}. \tag{19}$$

In the partial differentiation in $\dfrac{\partial\phi}{\partial T}$, ξ and μ_i are treated as independent variables. If, however, we regard them as functions (defined by (14)) of r and T, (16) and (17) lead to the following result:

$$\frac{1}{kT^2}\overline{E}(r,T)=\frac{\partial\phi}{\partial T}=\frac{2\Psi(r,T)}{\partial T}. \tag{20}$$

Mathematically the change of the independent variables from T, ξ and μ_i to T and r is equivalent to the Legendre transformation

$$\xi,\ \mu_1,\ \mu_2,\ \mu_3,\ \mu_4\rightarrow N,\ Nw,\ Nw,\ Nw,\ Nw$$

defined by (14). Substituting (20) into (4) we get

$$\log p(r,T)=\log g(r)+\Psi(r,T)-\Psi(r,\infty), \tag{21}$$

so that the free energy may be written down:

$$F(r,T)=-kT\log p(r,T)=[\log g(r)+\Psi(r,T)-\Psi(r,\infty)](-kT). \tag{22}$$

The equilibrium value $\overline{r}$ of r is obtained by minimizing F:

$$0=\frac{\partial F(\overline{r},T)}{\partial\overline{r}}=-kT\left[\frac{\mathrm{d}\log g(\overline{r})}{\mathrm{d}\overline{r}}+\frac{\partial\Psi(\overline{r},T)}{\partial\overline{r}}-\frac{\partial\Psi(\overline{r},\infty)}{\partial\overline{r}}\right]. \tag{23}$$

---

[7] The proof follows easily (if we put e^{ψ} to be the function Φ) from Lemma 2.42 of Fowler's Statistical Mechanics, second edition.

But by (16) and (17)

$$\frac{\partial \Psi(r,T)}{\partial r} = \sum_i N \log \mu_i, \tag{24}$$

and by (3)

$$\frac{\mathrm{d} \log g(r)}{\mathrm{d} r} = 2N \log \frac{w}{r},$$

so that by (18)

$$\sum_i \log \mu_i = -2 \log \frac{\overline{w}}{\overline{r}} + \left[\sum_i \log \mu_i\right]_{T=\infty} = 2 \log \left(\frac{\overline{w}}{\overline{r}}\right),$$

i.e.

$$\prod_i \mu_i = \left(\frac{\overline{w}}{\overline{r}}\right)^2. \tag{25}$$

It will be shown in the next section that we may put $V_{AA} = V_{BB}$, $V_{AB} = 0$, without altering the specific heat of the crystal if $V = \frac{1}{2}(V_{AA} + V_{BB}) - V_{AB}$ is left unchanged. When this is done, ϕ will be symmetrical with respect to μ_1, μ_2, μ_3 and μ_4, and we conclude that all the μ's are equal from the facts that (i) equation (14) has only one set of solution[7], and (ii) if the conclusion is true (14) becomes, with all μ_i put equal to μ,

$$\xi \frac{\partial \phi}{\partial \xi} = N, \qquad \mu \frac{\partial \phi}{\partial \mu} = 4Nw, \tag{26}$$

which $\underline{\text{does}}$ have[7] a set of solution in ξ and μ. Now ϕ is given by

$$\phi = \xi[1 + 4\mu x^2 + (4\mu^2 x^2 + 2\mu^2 x^4) + 4\mu^3 x^2 + \mu^4], \tag{27}$$

where x is defined by (9). On eliminating ξ from (26) we obtain

$$(1+s)\mu^4 + (2+4s)x^2\mu^3 + 2sx^2(x^2+2)\mu^2 + (4s-2)x^2\mu + (s-1) = 0. \tag{28}$$

The free energy is given by (21) and (18):

$$-\frac{F(r,T)}{2NkT} = r \log r + w \log w - 2w \log \mu + \frac{1}{2} \log(1 + 4\mu x^2 + 4\mu^2 x^2 + 2\mu^2 x^4 + 4\mu^3 x^2 + \mu^4), \tag{29}$$

and the condition of equilibrium by (25):

$$\mu = \sqrt{\frac{1-\overline{s}}{1+\overline{s}}}. \tag{30}$$

To obtain the critical temperature, we expend (28) in powers of s and find after identifying coefficients

$$\log \mu = -\frac{1+6x^2+x^4}{2+2x^2} s + \kappa s^3 + \cdots,$$

which is the only real solution for $\log\mu$. Next we expand (30):

$$\log\mu = -s - \frac{1}{3}s^3 - \cdots .$$

At the critical value x_c of x, these last two equations have a multiple solution at $s = 0$. Hence

$$-\frac{1 + 6{x_c}^2 + {x_c}^4}{2 + 2{x_c}^2} = -1,$$

i.e.

$$x_c = (\sqrt{5} - 2)^{1/2} = .4858\,.$$

4. GENERAL FORM OF THE QUASI-CHEMICAL METHOD

Let us now take a group of any size and form. Let it have a α-sites b β-sites and γ pairs of nearest neighbours. The procedures to obtain an approximate expression for the free energy of the crystal follow exactly the same line as in the special case considered in the last section. Equations (13), (14) and (18) are essentially unchanged:

$$\overline{[q_1, q_2, \ldots]} = \xi\mu_1^{q_1}\mu_2^{q_2}\cdots \mathrm{e}^{-\chi/kT}, \tag{31}$$

$$\phi = \sum_q \xi\mu_1^{q_1}\mu_2^{q_2}\cdots \mathrm{e}^{-\chi/kT}, \quad \xi\frac{\partial\phi}{\partial\xi} = N, \quad \mu_i\frac{\partial\phi}{\partial\mu_i} = Nw, \tag{32}$$

and

$$(\mu_i)_{T=\infty} = \frac{w}{r}. \tag{33}$$

But (12) should be corrected by a factor $\frac{\gamma}{z}$ to account for the duplications in calculating E from the sum of the energies of all the groups in the crystal:

$$\overline{E} = \frac{z}{\gamma}\sum_q \overline{[q_1, q_2, \ldots]}\,\chi(q_1, q_2, \ldots) = \frac{z}{\gamma}kT^2\frac{\partial\phi}{\partial T}. \tag{34}$$

Hence (22) becomes[*]

$$F(r,T) = -kT\left[\log g(r) + \frac{z}{\gamma}\Psi(r,T) - \frac{z}{\gamma}\Psi(r,\infty)\right], \tag{35}$$

or more explicitly, by (3), (17), and (33):

$$F(r,T) = -\frac{zNkT}{\gamma}\left[\log N + \left(a + b - \frac{2\gamma}{z}\right)(r\log r + w\log w) - \log\xi - w\sum_i \log\mu_i\right]. \tag{36}$$

---

[*] Care must be taken when the theory is extended to the case when the atomic ratio is not 1:1. The function Ψ in (35) must then be replaced by $\frac{1}{2}\Psi + \frac{1}{2}\Psi'$ where Ψ' is the function Ψ for the case when the group of interest has the same form as the original one but with α and β sites interchanged.

The derivative is

$$\frac{\partial}{\partial r}F(r,T) = -\frac{zNkT}{\gamma}\log\left[\left(\prod_i \mu_i\right)\left(\frac{r}{w}\right)^{a+b-\frac{2\gamma}{z}}\right], \tag{37}$$

so that the condition of equilibrium is

$$\prod_i \mu_i = \left(\frac{\overline{w}}{\overline{r}}\right)^{a+b-\frac{2\gamma}{z}}. \tag{38}$$

In actual calculations the following points may prove helpful:

(i) The free energy is changed by a constant if V_{AA} and V_{BB} are both replaced by $1/2(V_{AA}+V_{BB})-V_{AB}$, and V_{AB} by 0. To prove this let z_i be the number of sites in the group neighbouring to the site i. Let χ be changed into χ' by the replacement. It is evident that

$$\chi' - \chi = -\gamma V_{AB} + \frac{V_{AA}-V_{BB}}{2}\ (\text{no. of } B\text{-}B \text{ pairs} - \text{no. of } A\text{-}A \text{ pairs}),$$

and that

$$\sum_{\alpha\text{-sites}} q_i z_i - \sum_{\beta\text{-sites}} q_i z_i = \text{no. of } B\text{-}B \text{ pairs} - \text{no. of } A\text{-}A \text{ pairs}.$$

Hence

$$\overline{[q_1, q_2, \ldots]} = \xi \mu_1^{q_1} \mu_2^{q_2} \cdots \mathrm{e}^{-\chi/kT} = \xi' \mu_1'^{q_1} \mu_2'^{q_2} \cdots \mathrm{e}^{-\chi'/kT},$$

if we put

$$\xi' = \xi \mathrm{e}^{-\gamma V_{AB}/kT}, \quad \mu_i' = \mu_i \mathrm{e}^{\pm z_i (V_{AA}-V_{BB})/2kT},$$

where the $+$ sign or the $-$ sign is to be taken according as the site i is an α or a β site. We can now calculate the new free energy and verify the above statement.

(ii) Sites that are symmetrically situated in the group have equal μ's irrespective of their nature if $V_{AA} = V_{BB}$, $V_{AB} = 0$. This has already been shown in the last section. Since the most troublesome part of the calculations is the elimination of the parameters, much might be gained by choosing a group with a large number of sites symmetrically situated.

(iii) The free energy is a function of s^2, so that (38) is always satisfied at $\overline{w} = \overline{r} = 1/2$ (i.e. long distance order = 0). The proof is simple when we have already made $V_{AA} = V_{BB}$, $V_{AB} = 0$, so that an interchange of A and B atoms does not alter the energy. Thus

$$\chi(q_1, q_2, \ldots) = \chi(1-q_1, 1-q_2, \cdots).$$

Putting

$$\xi' = \xi \mu_1 \mu_2 \cdots$$

and

$$\mu_i' = \frac{1}{\mu_i}, \tag{39}$$

we get

$$\xi {\mu_1}^{q_1} {\mu_2}^{q_2} \cdots \mathrm{e}^{-\chi/kT} = \xi' {\mu_1'}^{1-q_1} {\mu_2'}^{1-q_2} \cdots \mathrm{e}^{-\chi/kT}.$$

Thus if (32) is satisfied

$$\begin{aligned}\sum_q (1-q_i)\xi' {\mu_1'}^{1-q_1} {\mu_2'}^{1-q_2} \cdots \mathrm{e}^{-\chi/kT} &= \sum_q \xi {\mu_1}^{q_1} {\mu_2}^{q_2} \cdots \mathrm{e}^{-\chi/kT} - \sum_q q_i \xi {\mu_1}^{q_1} {\mu_2}^{q_2} \cdots \mathrm{e}^{-\chi/kT} \\ &= Nr,\end{aligned}$$

i.e. ξ', μ_1', μ_2', ...would be the solution of (32) with r substituted for w. Hence by (32) and (17)

$$\Psi(1-r,T) = N - N\log\xi' - \sum_i Nr\log\mu_i' = \Psi(r,T)$$

showing that

$$F(1-r,T) = F(r,T). \tag{40}$$

(iv) <u>The parameter for a corner site is always given by</u>

$$\epsilon = \frac{1}{1+s}\left(\sqrt{x^2s^2+(1-s^2)} - sx\right) \tag{41}$$

<u>irrespective of the size of the group, if $V_{AA} = V_{BB}$, $V_{AB} = 0$.</u> By a corner site we mean a site that has only one nearest neighbour in the group. Let ϵ be the selective variable (parameter) of a corner site, and μ_1 that of its only neighbour in the group. If the corner site is dropped, a new group is obtained. We distinguish all quantities referring to this new group by a prime, and obtain at once

$$\xi'\frac{\partial\phi'}{\partial\xi'} = N, \qquad \mu_i'\frac{\partial\phi'}{\partial\mu_i'} = Nw, \quad i = 1,2,\cdots. \tag{42}$$

The sites of the primed group are numbered in the same way as in the unprimed group. Introducing the variable χ defined in (9) we may write

$$\begin{aligned}\phi &= \sum_{p,q} \xi\epsilon^p \mu_1^{q_1}\mu_2^{q_2}\cdots \mathrm{e}^{-\chi/kT} \\ &= \sum_{q_2,\cdots} \xi(1+\epsilon x)\mu_2^{q_2}\mu_3^{q_3}\cdots \mathrm{e}^{-\chi'/kT} + \mu_1\sum_{q_2,\cdots}\xi(\epsilon+x)\mu_2^{q_2}\cdots \mathrm{e}^{-\chi'/kT}.\end{aligned} \tag{43}$$

Let these two terms be denoted by ϕ_0 and ϕ_1 respectively. Since

$$\phi = N, \quad \mu_1\frac{\partial\phi}{\partial\mu_1} = Nw,$$

we have

$$\phi_0 = Nr, \quad \phi_1 = Nw. \tag{44}$$

Now

$$\epsilon\frac{\partial\phi_0}{\partial\epsilon}=\frac{\epsilon x}{1+\epsilon x}\phi_0,\quad \epsilon\frac{\partial\phi_1}{\partial\epsilon}=\frac{\epsilon}{\epsilon+x}\phi_1.$$

Hence $\epsilon\frac{\partial\phi}{\partial\epsilon}=Nw$ leads to

$$\frac{\epsilon x}{1+\epsilon x}Nr+\frac{\epsilon}{\epsilon+x}Nw=Nw, \tag{45}$$

or

$$\frac{w}{r}=\frac{\epsilon(\epsilon+x)}{1+\epsilon x}, \tag{46}$$

the solution of which is (41). Thus the two parameters μ and ν in the approximation discussed in section 2 are equal to ϵ.

(v) The "contribution" to the free energy from a corner atom is such that, in the notations of (iv).

$$F(r,T)=\frac{\gamma-1}{\gamma}F'(r,T)+\frac{1}{\gamma}F_0(r,T), \tag{47}$$

where $F_0(r,T)$ is the free energy when $\gamma=1$, i.e. the free energy in the approximation discussed in section 2. This is proved as follows.

If we put

$$\xi=\xi''\frac{1}{1+\epsilon x},\quad \mu_1=\mu_1''\frac{1+\epsilon x}{\epsilon+x},\quad \mu_i=\mu_i'',\ \ i\geq 2, \tag{48}$$

it is evident from (43) that ϕ would become a function of ξ'', μ_1'', μ_2'', $\cdots$ satisfying the relations

$$\xi''\frac{\partial\phi}{\partial\xi''}\left(=\xi\frac{\partial\phi}{\partial\xi}\right)=N,\quad \mu_i''\frac{\partial\phi}{\partial\mu_i''}\left(=\mu_i\frac{\partial\phi}{\partial\mu_i}\right)=Nw,\quad i=1,2,\cdots. \tag{49}$$

It is also evident that ϕ is the same function of ξ'', μ_1'', μ_2'', $\cdots$ as ϕ' is of ξ', μ_1', μ_2', $\cdots$. Now (42) has only one[7] set of solution in ξ' and μ_i'. Hence from (49) we infer that $\xi'=\xi''$, $\mu_i'=\mu_i''$. Thus

$$\xi=\xi'\frac{1}{1+\epsilon x},\quad \mu_1=\mu_1'\frac{1+\epsilon x}{\epsilon+x},\quad \mu_i=\mu_i',\ \ i\geq 2. \tag{50}$$

(41) and (50) give the parameters μ_i in terms of μ_i'. Inserting them into (36) we obtain

$$\begin{aligned}F(r,T)=&-NkT\frac{z}{\gamma}\left[\log N+\left(a+b-\frac{2\gamma}{z}\right)(r\log r+w\log w)-\log\xi'-w\sum_i\log\mu_i'+\right.\\&\left.r\log(1+\epsilon x)+w\log(\epsilon+x)-w\log\epsilon\right]\\=&\frac{\gamma-1}{\gamma}F'(r,T)-NkT\frac{z}{\gamma}\left[\left(1-\frac{2}{z}\right)(r\log r+w\log w)+\right.\\&\left.r\log(1+\epsilon x)+w\log\frac{\epsilon+x}{\epsilon}\right].\end{aligned} \tag{51}$$

If the original (unprimed) group is a pair of nearest neighbours, we have $\gamma = 1$, and (51) reduces to the expression for the free energy in the approximation discussed in section 2:

$$F_0(r,T) = -zNkT\left[\left(1-\frac{2}{z}\right)(r\log r + w\log w) + r\log(1+\epsilon x) + w\log\frac{\epsilon+x}{\epsilon}\right]. \quad (52)$$

Inserting this back into (51) we get (47).

5. COMPARISON WITH BETHE'S METHOD

The so-called local grand partition function[1,3,8] in Bethe's approximations with long-distance order is identical in form with our function ϕ when all the "interior sites" in the group have the same parameter μ. For the case of equal concentrations for the two kinds of atoms, which is the case so far considered, this parameter has been put equal to unity by Bethe. Since the different terms of the local grand partition function stand for the probabilities of occurrence of the corresponding local groups in the crystal, it is clear that Bethe's method with long-distance order is essentially equivalent to our method plus the assumption that the free energy (35) has a minimum when

$$(\mu)_{\text{interior sites}} = 1. \quad (53)$$

But as we have shown that (38) gives the condition of a minimum of the free energy, the complete* identification of Bethe's and the quasi-chemical methods in any approximation reduces to the mathematical proof of the equivalence of (38) and (53). While this presents no difficulty at all for Bethe's first approximation (section 7), a general proof is by no means easy. We can only satisfy ourselves with the assertion that the two methods are equivalent for large groups, i.e. groups for which

$$a + b - \frac{2\gamma}{z} \ll \gamma.$$

This follows from the fact that if (53) is true

$$\left[\left(\prod \mu_i\right)\left(\frac{r}{w}\right)^{a+b-\frac{2\gamma}{z}}\right]^{\frac{1}{\gamma}} \cong \left[\prod (\mu)_{\text{interior sites}}\right]^{\frac{1}{\gamma}} = 1,$$

so that by (37)

$$\frac{\partial}{\partial r}F(r,T) = 0.$$

---

[8] Easthope, Proc. Camb. Phil. Soc. <u>33</u>, 502 (1937).

[*] "complete" as far as the probabilities of occurrence of the local configurations are concerned. The energy calculations are different in the two methods.

To see how the equilibrium free energy $F(\bar{r},T)$ varies with T in high order approximations, we substitute (38) into (36) and make use of (32):

$$-\frac{F}{zNkT}=\frac{1}{\gamma}\log\left(\sum_q \mu_1^{q_1}\mu_2^{q_2}\cdots \mathrm{e}^{-\chi/kT}\right)+\frac{1}{\gamma}\left(a+b-\frac{2\gamma}{z}\right)\log\bar{r}.$$

The last term is very small for large groups, so that by (53)

$$-\frac{F}{zNkT}=\frac{1}{\gamma}\log\left(\sum \mathrm{e}^{-\chi/kT}\right).$$

6. THE NON-CINTERFERENCE OF LOCAL CONFIGURATIONS

Let us return to the fundamental assumption of the quasi-chemical method, i.e. to (31) which gives the average numbers of the different local configurations (so far called groups) in the crystal. This equation expresses the exact distribution law of an assembly of molecules (cf. the example in section 2) which has an energy γ/z times as large as the crystal. Distinguishing all quantities referring to the assembly of molecules by a subscript m, we get

$$F(r,T)+kT\log g(r)=\frac{z}{\gamma}[F_m(r,T)+kT\log g_m(r)],$$

which is obtained from (4). But if H is the number of arrangements in the crystal lattice having the given values of $[q_1,q_2,\ldots]$,

$$F(r,T)=-kT\log\overline{H}+\overline{E}. \tag{54}$$

Thus

$$\log\frac{\overline{H}}{g(r)}=\frac{z}{\gamma}\log\frac{\overline{H}_m}{g_m(r)}.$$

But[*]

$$H_m=\frac{N!}{\prod\limits_q [q_1,q_2,\ldots]!}, \tag{55}$$

hence dropping the bar we get

$$H=h(r)\left\{\frac{N!}{\prod\limits_q [q_1,q_2,\ldots]!}\right\}^{z/\gamma}, \tag{56}$$

where

$$h(r)=\frac{g(r)}{\{g_m(r)\}^{z/\gamma}}. \tag{57}$$

---

[*] Fowler, Statistical Mechanics, second edition, sections 2.6 and 5.11.

Equation (56) has been referred to in Fowler and Guggenheim's paper[1] as the mathematical expression of the "hypothesis of the non-interference of local configurations", because when $\gamma/z = 1$, the number of arrangements in the crystal consistent with the distribution law $[q_1, q_2, \ldots]$ for the groups of sites is, except for the factor $h(r)$, equal to

$$H_m = \frac{N!}{\prod\limits_q [q_1, q_2, \ldots]!},$$

which is the number of arrangements in the crystal for the given values of $[q_1, q_2, \ldots]$ if the N groups in the crystal are imagined to be separated and are filled independently with atoms. The term "non-interference" comes from the fact that actually the N groups are not separated but are interlocked and cannot be filled independently with atoms, i.e. they "interfere" with each other.

To find the value of $g_m(r)$ we notice that by definition $g_m = \sum H_m$. But $\sum H_m$ is the number of arrangements in the N separated groups considered above if they are to be so filled with atoms that Nw of them have wrong atoms on the sites $i, i = 1, 2, \ldots$. Among the N sites i of the N groups $\frac{N!}{(Nr)!(Nw)!}$ different arrangements are possible. Hence[*]

$$g_m = \sum H_m = \sum \frac{N!}{\prod\limits_q [q_1, q_2, \ldots]!} = \left[\frac{N!}{(Nr)!(Nw)!}\right]^{a+b}. \tag{58}$$

Thus

$$h(r) = \left[\frac{N!}{(Nr)!(Nw)!}\right]^{2-(a+b)\frac{z}{\gamma}}. \tag{59}$$

The free energy of the crystal may be obtained from (54), (56) and (59):

$$\begin{aligned} F(r,T) = \overline{E} - \frac{zNkT}{\gamma}\Bigg\{ \left(a + b - \frac{2\gamma}{z}\right)(r\log r + w\log w) + \log N - \\ \frac{1}{N}\sum_q \overline{[q_1, q_2, \ldots]}\log\overline{[q_1, q_2, \ldots]} \Bigg\}, \end{aligned} \tag{60}$$

which has been obtained above by integration.

---

[*] It might be mentioned in passing that for the special case considered in section 2. (58) gives directly the value of the sum $\sum_x$ in equation (8.5) of Fowler and Guggenheim's paper if their r and q are equal. The generalization to the case r/q is however easy. The result is

$$\sum_x \frac{[zN]!}{[zN(r-x)]![zNx]![zN(1-r-q+x)]![zN(q-x)]!} = \frac{[zN]!}{[zNr]![zN(1-r)]!}\,\frac{[zN]!}{[zNq]![zN(1-q)]!} \tag{60}$$

which is exact. The value of $\log\sum_x$ given by (60) reduces to the approximate expression that Fowler and Guggenheim obtained by identifying $\sum_x$ with its maximum term when N is large.

7. SPECIAL CONSIDERATIONS CONCERNING BETHE'S FIRST AND SECOND APPROXIMATIONS

(i) First Approximation. If an α-site together with its z nearest neighbours are taken as our group of interest, all the sites except the central one are corner sites. Hence their selective variables are all equal to the value of ϵ given in (41). By successive applications of (47) we see that the free energy is exactly $F_0(r, T)$, a fact which has already been pointed out by Fowler and Guggenheim[1]. The selective variable of the central site is given by successive applications of (50)

$$\lambda = \frac{w}{r}\left(\frac{1+\epsilon x}{\epsilon + x}\right)^z. \tag{61}$$

The factor w/r is the selective variable for the central site when it alone forms the group. The equilibrium condition (38) becomes

$$\lambda\epsilon^z = \left(\frac{\overline{w}}{\overline{r}}\right)^{z-1}.$$

But by (61) and (45),

$$\lambda = \left(\frac{w}{r}\right)^{1-z}\epsilon^z.$$

Hence at equilibrium

$$\lambda = 1. \tag{62}$$

Thus the approximation is completely equivalent to Bethe's first approximation, as already mentioned in section 5.

(ii) Second Approximation. Now consider the group of sites occurring in Bethe's second approximation[3]. According to section 4, (iv), the selective variables for the corner sites in the second shell are all equal to ϵ, which is given by (41). But in Bethe's original calculations, the selective variables for the corner sites and the medium sites are made equal, and are found to be different from ϵ. Thus if we use his original method, equation (32) can not be satisfied. (In other words, the probabilities of occurrence or wrong atoms in the corner and the medium sites would be unequal.)

For simplicity we shall drop the corner sites and take as our group of interest the central site, the first shell sites and the medium sites; with selective variables μ, ν and λ respectively. (The contribution by the corner sites can be included in the free energy by simple addition as shown in section 4 (v).) With the notations n, and g_{nm} of Bethe[3] we find

$$\phi = \xi \sum_n (x^n + \mu x^{z-n}) P_n(x, \nu, \lambda), \tag{63}$$

where

$$P_n(x,\nu,\lambda)=\nu^n\sum_m g_{nm}[(1+\lambda)x]^m(x^2+\lambda)^{(\frac{z}{2}-1)n-\frac{m}{2}}(1+\lambda x^2)^{(\frac{z}{2}-1)(z-n)-\frac{m}{2}}.$$

After eliminating ξ and μ, (32) becomes

$$zw=\frac{2\sum\limits_{l,n}x^{l-n}\left(rP_n\lambda\frac{\partial}{\partial\lambda}P_l+wP_l\lambda\frac{\partial}{\partial\lambda}P_n\right)}{(z-2)\Big(\sum\limits_n x^nP_n\Big)\Big(\sum\limits_n x^{-n}P_n\Big)}=\frac{\sum\limits_{l,n}x^{l-n}\left(rP_n\nu\frac{\partial}{\partial\nu}P_l+wP_l\nu\frac{\partial}{\partial\nu}P_n\right)}{\Big(\sum\limits_n x^nP_n\Big)\Big(\sum\limits_n x^{-n}P_n\Big)}. \quad (64)$$

The free energy is obtained from (36):

$$F(r,T)=-\frac{NkT}{z-1}\Bigg[\frac{1}{2}(z^2-4z+4)(r\log r+w\log w)+r\log\sum_n x^nP_n+$$

$$w\log\sum_n x^{z-n}P_n-zw\log\nu-wz\left(\frac{z}{2}-1\right)\log\lambda\Bigg]. \quad (65)$$

8. APPLICATION TO THE CRYSTAL Cu_3Au

For the face-centred crystal Cu_3Au, we may of course follow Peierls[9] and take as our group a central site together with its twelve first shell neighbours. The free energy expression would then contain seven selective variables[*], four of which can be eliminated. The resultant expression is very cumbersome and numerical calculations would be laborious. We therefore make a simpler approximation: the group is taken to be four nearest neighbours forming a tetrahedron. A little geometrical consideration assures us that all such tetrahedrons contain an α-site (for gold atoms) and three β-sites (for copper atoms), an interesting conclusion showing that the tetrahedron might be

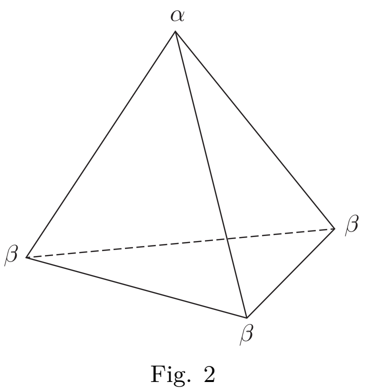


Fig. 2

---

[9] Peierls, Proc. Rey. Soc. London A154, 207(1936).

[*] For α-centred groups, three, and for β-centred, four, parameters are necessary. Both these two kinds of groups must be considered because otherwise the energy of the crystal cannot be easily obtained from the energy of the groups in the crystal.

regarded as a sort of “molecular” structure in a face-centred lattice with atomic ratio 1:3. Our approximation may thus be reasonably expected to reveal the more important features of order–disorder transformation in such alloys.

Let μ and ν be the parameters (for wrong atoms) of the β-sites and the α-sites respectively. Let there be altogether $4N$ atoms. It is easy to see that there are $8N$ groups in the crystal. When Nw atoms on the α-sites are wrong, the equations determining the parameters are

$$8N = \phi = \xi[x^3 + 3x^2\mu + 3x^3\mu^2 + x^6\mu^3 + \nu(x^6 + 3x^3\mu + 3x^2\mu^2 + x^3\mu^3)], \tag{66a}$$

$$8Nw = \nu\frac{\partial\phi}{\partial\nu} = \xi\nu(x^6 + 3x^3\mu + 3x^2\mu^2 + x^3\mu^3), \tag{66b}$$

and

$$8N\left(\frac{w}{3}\right) + 8N\left(\frac{w}{3}\right) + 8N\left(\frac{w}{3}\right) = \mu\frac{\partial\phi}{\partial\mu} = 3x^2\mu\xi[1 + 2x\mu + x^4\mu^2 + \nu(x + 2\mu + x\mu^2)], \tag{66c}$$

where x is defined by (9). The energy of the crystal is (cf. (34)),

$$\overline{E} = \frac{1}{2}kT^2\frac{\partial\phi}{\partial T} + \text{constant}; \tag{67}$$

so that the free energy becomes (cf. (35))

$$F(w,T) = -kT\left[\log g(w) + \frac{1}{2}(\phi - 8N\log\xi - 8Nw\log\nu - 8Nw\log\mu)_{T=\infty}^{T}\right].$$

But

$$\log g(w) = -N\left\{(1-w)\log(1-w) + w\log w + w\log\frac{w}{3} + (3-w)\log[(3-w)/3]\right\},$$

and at $T = \infty$,

$$\nu = \frac{w}{1-w}, \quad \mu = \frac{w}{3-w}, \quad \xi = 8N(1-w)\left(1-\frac{w}{3}\right)^3.$$

Hence

$$-\frac{F(w,T)}{NkT} = -9\log 3 + 4\log 8N + 6w\log w + 3(1-w)\log(1-w) + \\ 3(3-w)\log(3-w) - 4\log\xi - 4w\log\mu - 4w\log\nu. \tag{68}$$

Since ξ and ν can be very easily solved from (66), numerical calculations are quite simple. The equilibrium value of w is given by (cf. (37) and (38))

$$0 = -3\log\frac{(1-\overline{w})(3-\overline{w})}{\overline{w}^2} + 4\log\mu\nu. \tag{69}$$

This is always satisfied at $\overline{w} = \frac{3}{4}$*. Actual calculation shows that the absolute minimum

---

*This is not evident from (69) directly. But if we divide the whole crystal into four sublattices which are all simple cubic and introduce a w for each sublattice so that Nw is the number of A atoms on the i-th sublattice ($i = 1, 2, 3, 4$), it is obvious that the free energy is symmetrical in the w's. From this we infer that (69) is satisfied at $\overline{w} = \frac{3}{4}$.

of the free energy is or is not at $\overline{w} = \frac{3}{4}$ according as $x \geq .2965$ or $x < .2965$. The value of the free energy is plotted in Fig.3. From the form of the graph it is seen that the crystal has a critical temperature at which the long-distance order and (hence) the energy are discontinuous. The critical temperature T_C and the latent heat Q are found to be

$$T_C = .8228\frac{1}{k}\left[\frac{1}{2}(V_{AA}+V_{BB}) - V_{AB}\right], \quad Q = .8824N\left[\frac{1}{2}(V_{AA}+V_{BB}) - V_{AB}\right].$$

In terms of the total energy change from $T = 0$ to $T = \infty$:

$$E_0 = 3N\left[\frac{1}{2}(V_{AA}+V_{BB}) - V_{AB}\right],$$

these quantities become

$$\begin{aligned} T_C &= 1.097E_0/R, \quad &&(T_C = 2.19E_0/R \text{ in Bragg-Williams' approximation and} \\ & &&T_C \cong 1.3E_0/R \text{ in Peierls' approximation.)} \\ Q &= .2941E_0, \quad &&(Q = .218E_0 \text{ in Bragg-Williams' approximation and} \\ & &&Q \cong .36E_0 \text{ in Peierls' approximation.)} \end{aligned}$$

where R stands for $4Nk$.

It will be noticed that due to the lack of a free energy expression Peierls[9] did not give the exact values of these quantities.

In conclusion, the author wishes to express his thanks to Prof. J. S. Wang for valuable criticism and advice.

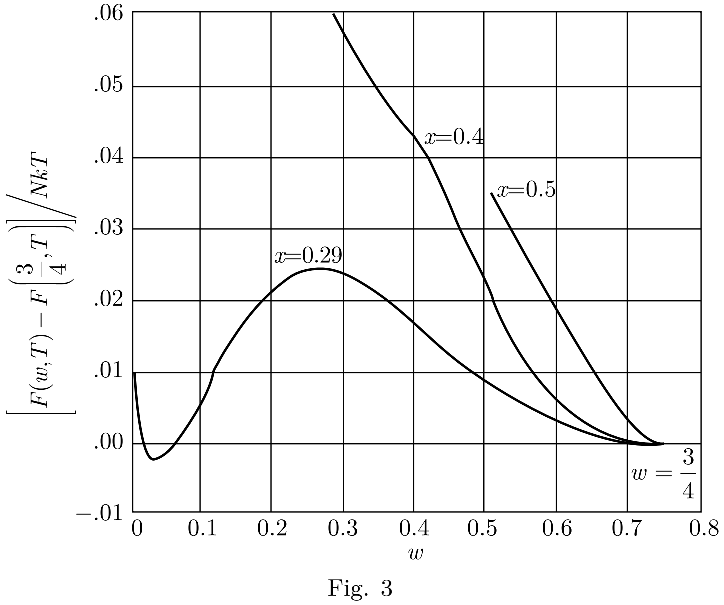
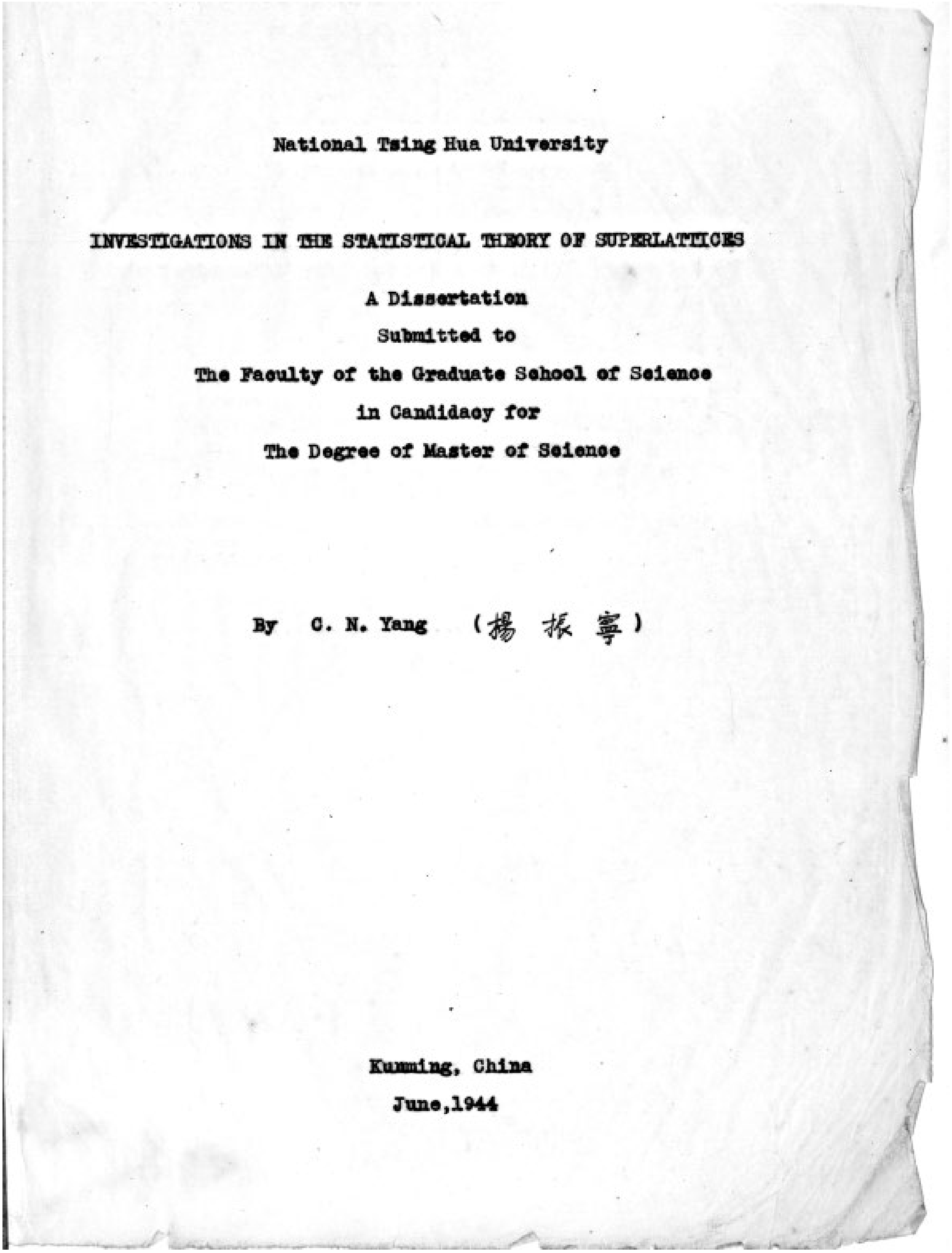


Fig. 3



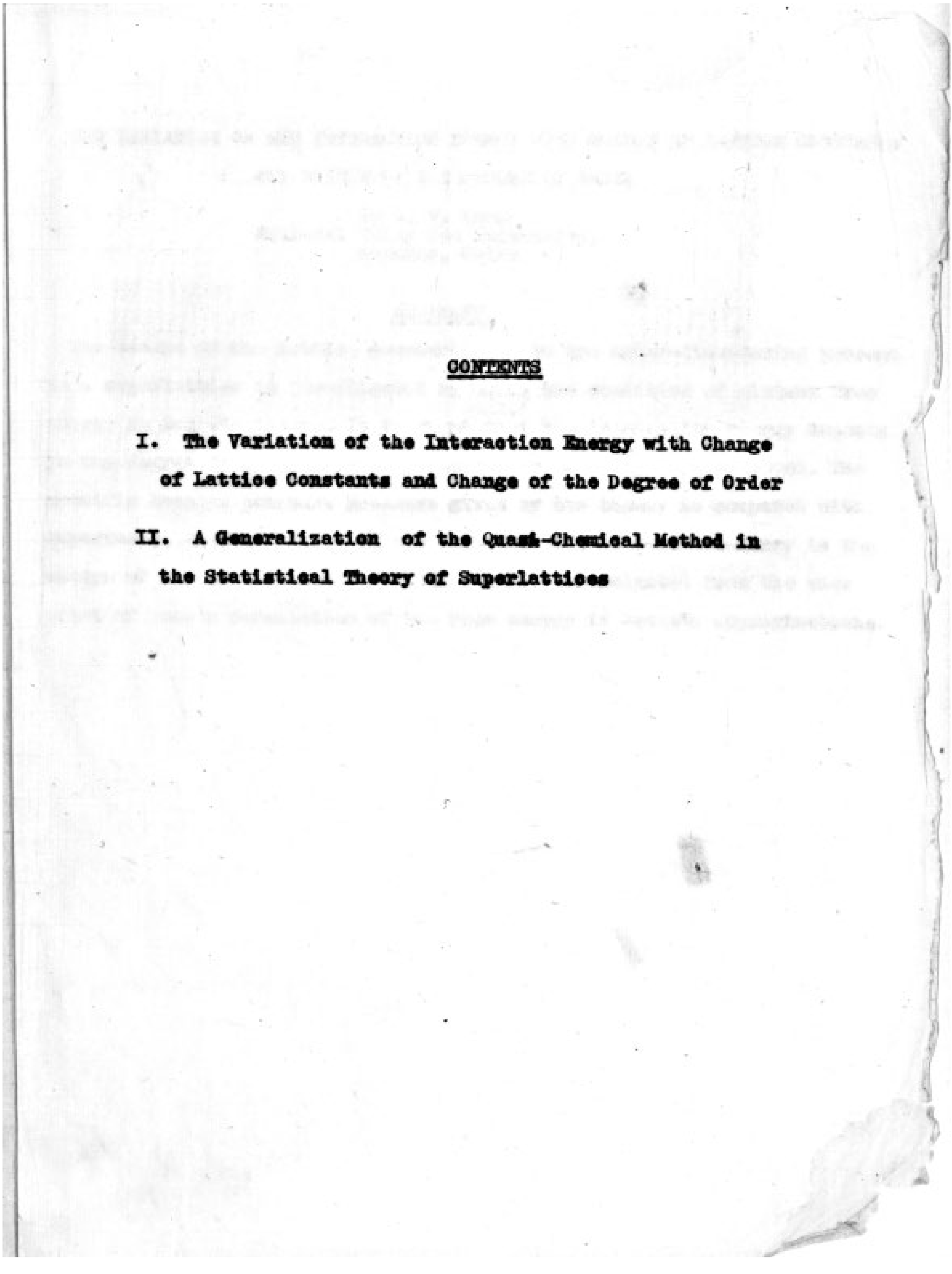
National Tsing Hua University

INVESTIGATIONS IN THE STATISTICAL THEORY OF SUPERLATTICES

A Dissertation
Submitted to
The Faculty of the Graduate School of Science
in Candidacy for
The Degree of Master of Science

By C. N. Yang (楊振寧)

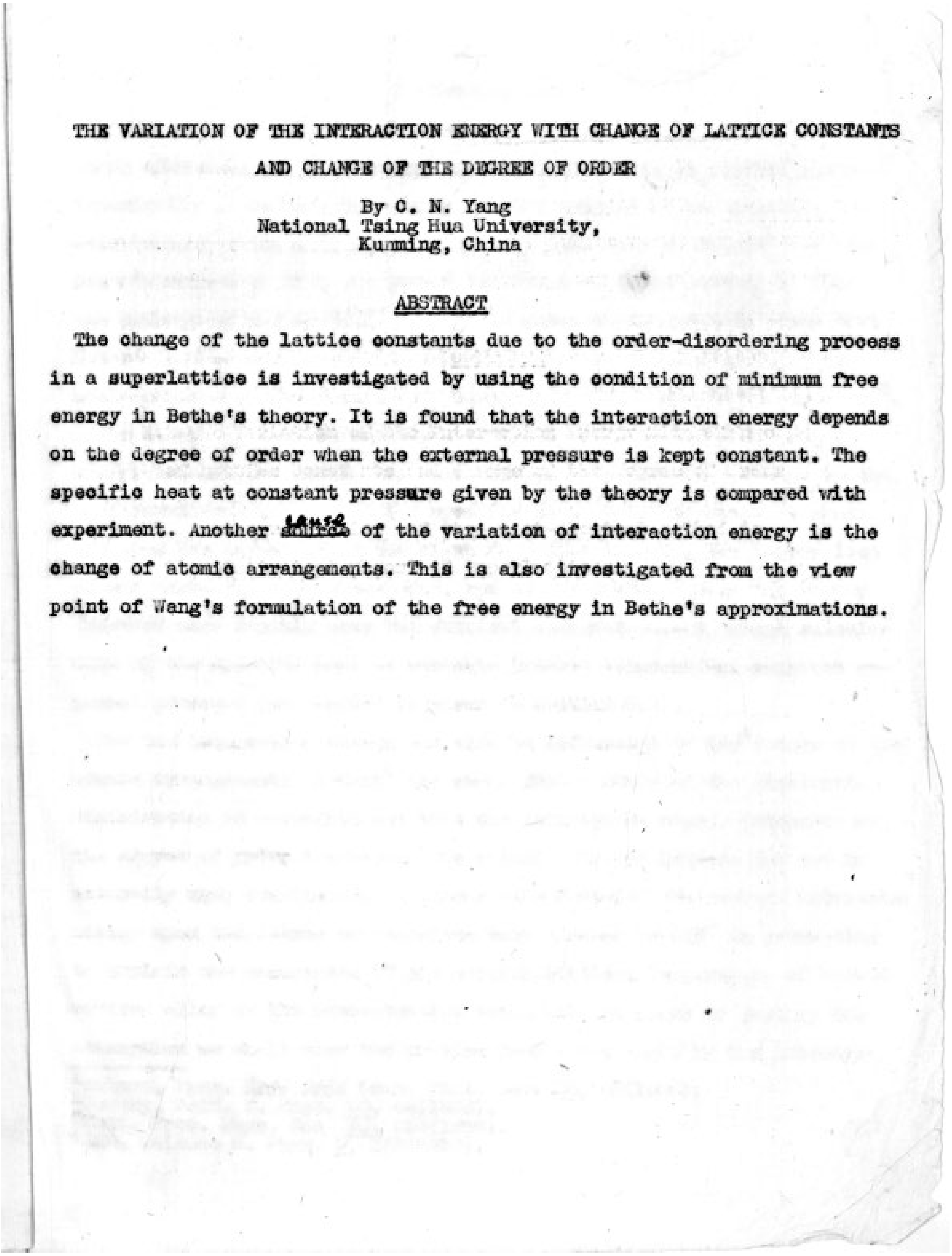
Kunming, China
June,1944



CONTENTS

I. The Variation of the Interaction Energy with Change of Lattice Constants and Change of the Degree of Order

II. A Generalization of the Quasi-Chemical Method in the Statistical Theory of Superlattices

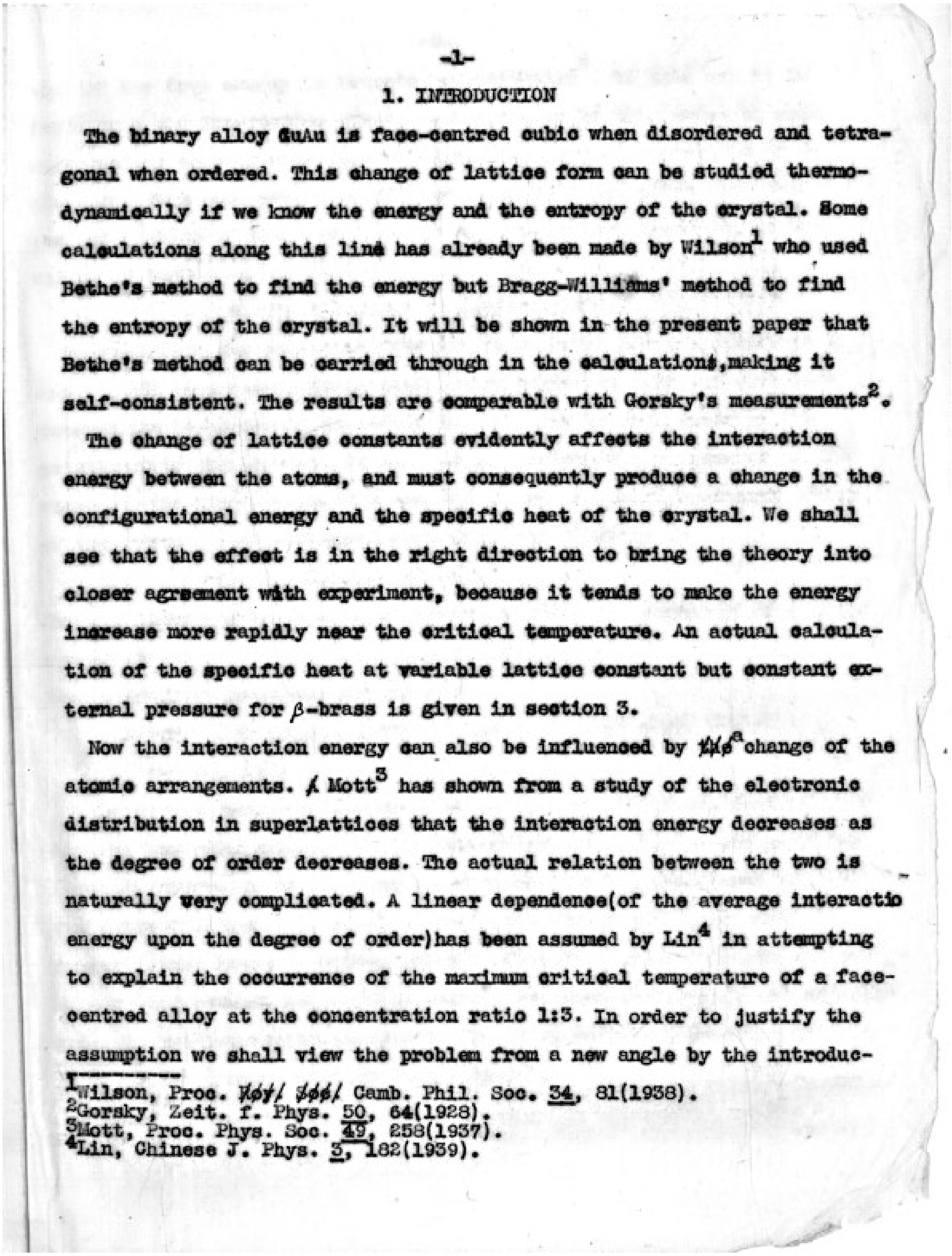


THE VARIATION OF THE INTERACTION ENERGY WITH CHANGE OF LATTICE CONSTANTS AND CHANGE OF THE DEGREE OF ORDER

By C. N. Yang
National Tsing Hua University,
Kunming, China

ABSTRACT

The change of the lattice constants due to the order-disordering process in a superlattice is investigated by using the condition of minimum free energy in Bethe's theory. It is found that the interaction energy depends on the degree of order when the external pressure is kept constant. The specific heat at constant pressure given by the theory is compared with experiment. Another ~~source~~ cause of the variation of interaction energy is the change of atomic arrangements. This is also investigated from the view point of Wang's formulation of the free energy in Bethe's approximations.



1. INTRODUCTION

The binary alloy CuAu is face-centred cubic when disordered and tetragonal when ordered. This change of lattice form can be studied thermodynamically if we know the energy and the entropy of the crystal. Some calculations along this line has already been made by Wilson[1] who used Bethe's method to find the energy but Bragg-Williams' method to find the entropy of the crystal. It will be shown in the present paper that Bethe's method can be carried through in the calculation~~s~~,making it self-consistent. The results are comparable with Gorsky's measurements[2].

The change of lattice constants evidently affects the interaction energy between the atoms, and must consequently produce a change in the configurational energy and the specific heat of the crystal. We shall see that the effect is in the right direction to bring the theory into closer agreement with experiment, because it tends to make the energy increase more rapidly near the critical temperature. An actual calculation of the specific heat at variable lattice constant but constant external pressure for β-brass is given in section 3.

Now the interaction energy can also be influenced by ~~the~~ a change of the atomic arrangements. ~~A~~ Mott[3] has shown from a study of the electronic distribution in superlattices that the interaction energy decreases as the degree of order decreases. The actual relation between the two is naturally very complicated. A linear dependence(of the average interactio energy upon the degree of order)has been assumed by Lin[4] in attempting to explain the occurrence of the maximum critical temperature of a face-centred alloy at the concentration ratio 1:3. In order to justify the assumption we shall view the problem from a new angle by the introduc-

---

[1] Wilson, Proc. ~~Roy.~~ ~~Soc.~~ Camb. Phil. Soc. 34, 81(1938).

[2] Gorsky, Zeit. f. Phys. 50, 64(1928).

[3] Mott, Proc. Phys. Soc. 49, 258(1937).

[4] Lin, Chinese J. Phys. 3, 182(1939).

tion of the free energy in Bethe's approximation[5]. In this way it is found that the interaction energy as a function of the degree of order must satisfy certain equations obtained from a set of conditions of consistency. This same set of conditions of consistency makes also possible the calculation of the energy of the crystal without appealing to Bragg-Williams' theory as Lin did.

2. THE VARIATION OF LATTICE CONSTANTS

We shall form the partition function at constant lattice constants l_1 and l_2, and then obtain their equilibrium values ~~by~~ from the equations determining the generalized reactions. Let $\frac{1}{2}zNm$ be the number of A-B neighbours in the crystal. If $g(m)$ is the number of arrangements of the atoms for the given value of m, and $W(l_1, l_2, m)$ the configurational energy of the crystal, the configurational partition function is

$$f(m,T,l_1,l_2)=g(m)\exp(-W/kT).$$

The equilibrium value $\bar{m}$ of m is determined from the condition of a maximum of f:

$$\frac{\partial}{\partial \bar{m}}\log f(\bar{m},T,l_1,l_2)=0.$$

The generalized reactions are given by

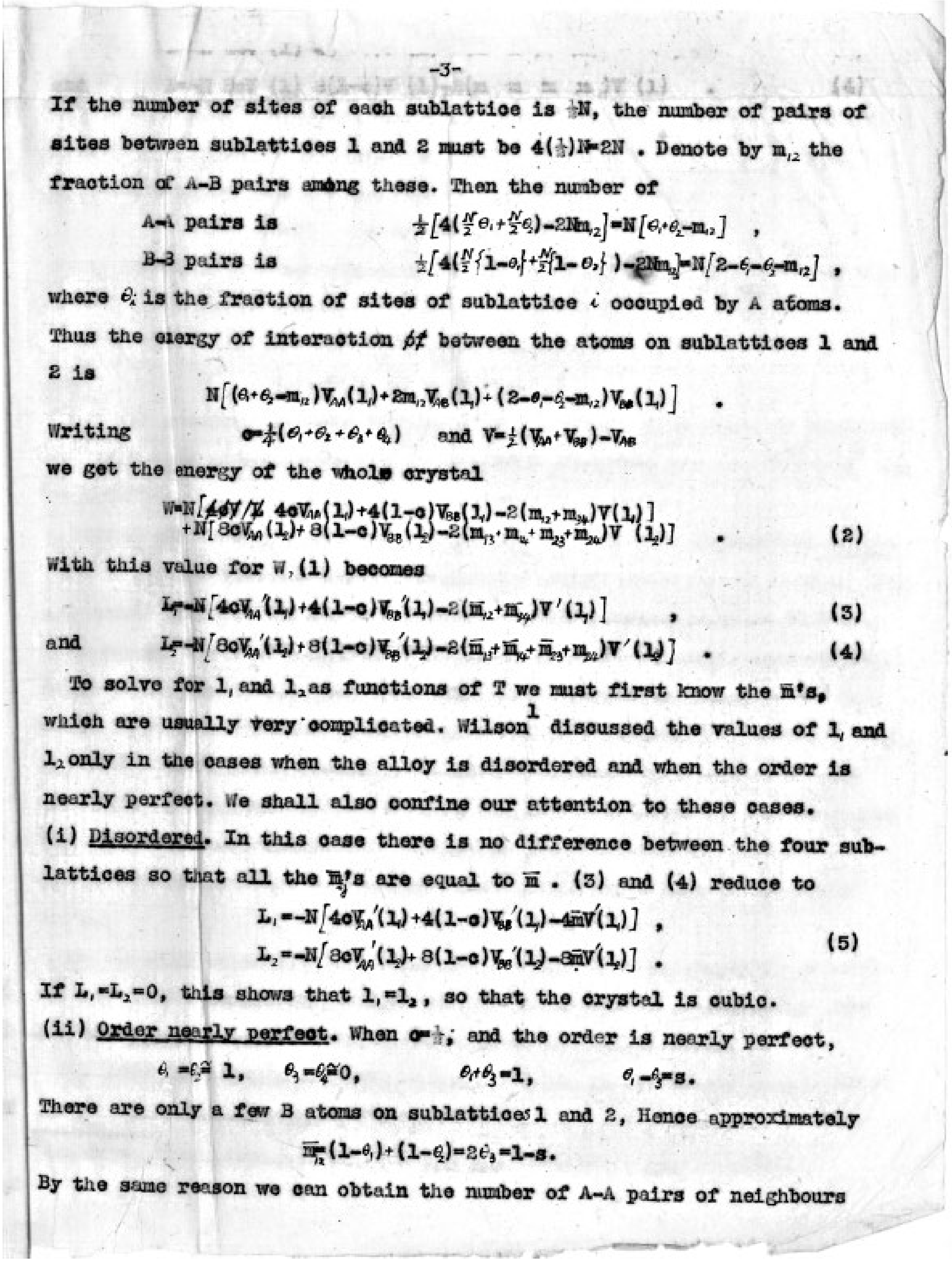
$$L_i=kT\frac{d}{dl_i}\log f(\bar{m},T,l_1,l_2)=kT\frac{d\bar{m}}{dl_i}\frac{\partial}{\partial\bar{m}}\log f+kT\frac{\partial}{\partial l_i}\log f=kT\frac{\partial}{\partial l_i}\log f$$
$$=-\frac{\partial}{\partial l_i}W(l_1,l_2,\bar{m}) \quad . \tag{1}$$

To study the change of lattice form in CuAu we divide the face-centred lattice into four simple cubic sublattices 1,2,3,4*. Let the shortest distance between the sites of 1 and 2, or 3 and 4 be l_1, that between the sites of 1 and 3, 1 and 4, 2 and 3 or 2 and 4 be l_2, so that the former is the distance between neighbouring Au-Au or Cu-Cu atoms and the latter that between neighbouring Au-Cu atoms when the crystal is perfectly ordered. The interaction energies V_{AA}, V_{AB} and V_{BB} are functions of l_1 and l_2.

---

*Cf. Fig. 27 in Rev. Mod. Phys. 10, 1(1938).

[5] Wang, "Free Energy in the Statistical Theory of Order-Disorder Transformations", Science Report of National Tsing Hua University, series A, 30-th anniversary Memorial Number (1941), printed but failed to appear.



-3-

If the number of sites of each sublattice is $\frac{1}{4}N$, the number of pairs of sites between sublattices 1 and 2 must be $4(\frac{1}{2})N=2N$. Denote by m_{12} the fraction of A-B pairs among these. Then the number of

A-A pairs is $\frac{1}{2}[4(\frac{N}{2}\theta_1+\frac{N}{2}\theta_2)-2Nm_{12}]=N[\theta_1+\theta_2-m_{12}]$,

B-B pairs is $\frac{1}{2}[4(\frac{N}{2}\{1-\theta_1\}+\frac{N}{2}\{1-\theta_2\})-2Nm_{12}]=N[2-\theta_1-\theta_2-m_{12}]$,

where θ_i is the fraction of sites of sublattice i occupied by A atoms. Thus the energy of interaction ~~of~~ between the atoms on sublattices 1 and 2 is

$$N[(\theta_1+\theta_2-m_{12})V_{AA}(l_1)+2m_{12}V_{AB}(l_1)+(2-\theta_1-\theta_2-m_{12})V_{BB}(l_1)] .$$

Writing $c=\frac{1}{4}(\theta_1+\theta_2+\theta_3+\theta_4)$ and $V=\frac{1}{2}(V_{AA}+V_{BB})-V_{AB}$

we get the energy of the whole crystal

$$\begin{aligned} W=N[\cancel{4cV/V}\ 4cV_{AA}(l_1)+4(1-c)V_{BB}(l_1)-2(m_{12}+m_{34})V(l_1)] \\ +N[8cV_{AA}(l_2)+8(1-c)V_{BB}(l_2)-2(m_{13}+m_{14}+m_{23}+m_{24})V(l_2)] \end{aligned} \quad . \qquad (2)$$

With this value for W, (1) becomes

$$L_1=-N[4cV_{AA}'(l_1)+4(1-c)V_{BB}'(l_1)-2(\bar{m}_{12}+\bar{m}_{34})V'(l_1)] , \qquad (3)$$

and

$$L_2=-N[8cV_{AA}'(l_2)+8(1-c)V_{BB}'(l_2)-2(\bar{m}_{13}+\bar{m}_{14}+\bar{m}_{23}+m_{24})V'(l_2)] . \qquad (4)$$

To solve for l_1 and l_2 as functions of T we must first know the $\bar{m}$'s, which are usually very complicated. Wilson[1] discussed the values of l_1 and l_2 only in the cases when the alloy is disordered and when the order is nearly perfect. We shall also confine our attention to these cases.

(i) Disordered. In this case there is no difference between the four sublattices so that all the $\bar{m}_{ij}$'s are equal to $\bar{m}$. (3) and (4) reduce to

$$\begin{aligned} L_1&=-N[4cV_{AA}'(l_1)+4(1-c)V_{BB}'(l_1)-4\bar{m}V'(l_1)] , \\ L_2&=-N[8cV_{AA}'(l_2)+8(1-c)V_{BB}'(l_2)-8\bar{m}V'(l_2)] . \end{aligned} \qquad (5)$$

If $L_1=L_2=0$, this shows that $l_1=l_2$, so that the crystal is cubic.

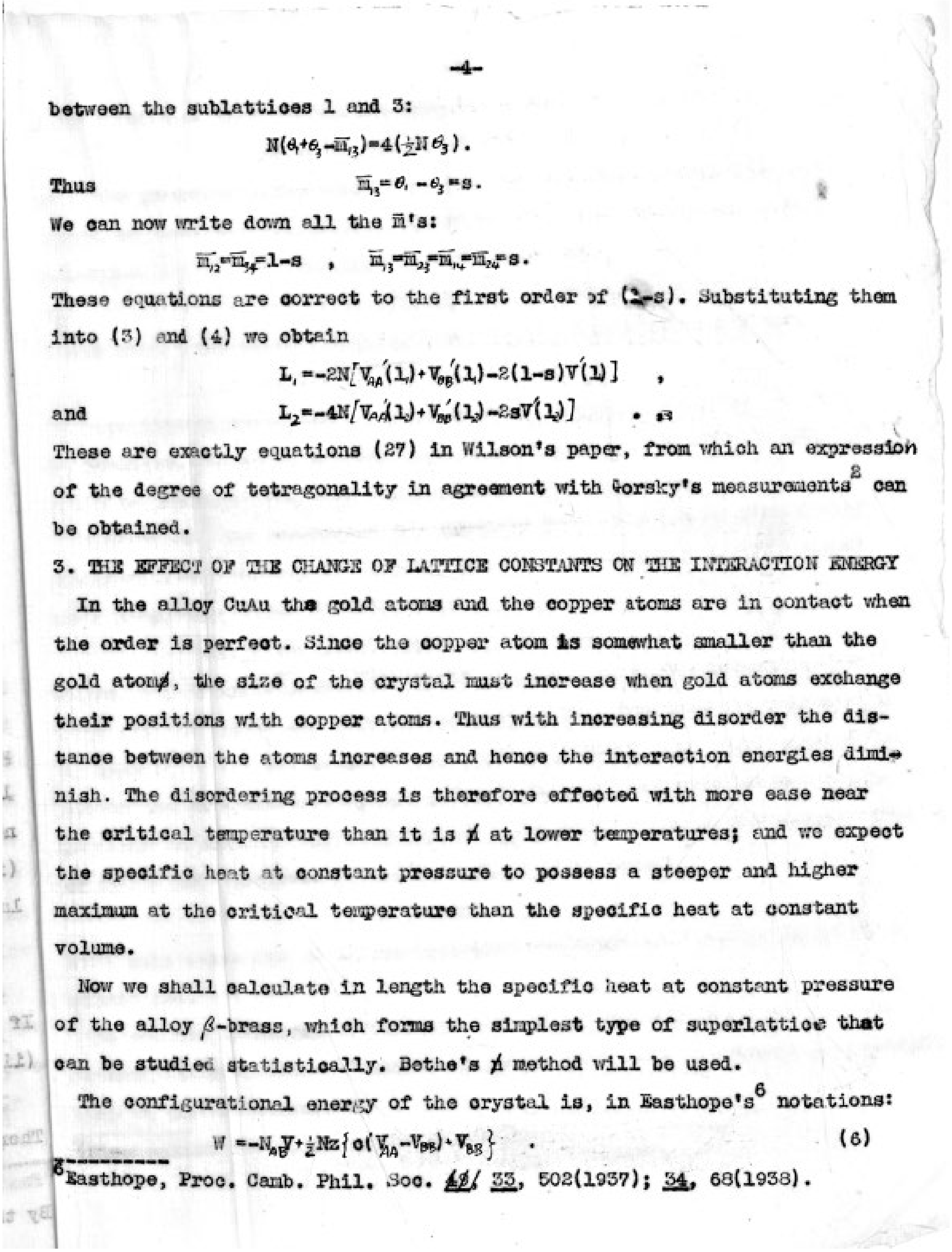
(ii) Order nearly perfect. When $c=\frac{1}{2}$; and the order is nearly perfect,

$$\theta_1=\theta_2\cong 1, \quad \theta_3=\theta_4\cong 0, \quad \theta_1+\theta_3=1, \quad \theta_1-\theta_3=s.$$

There are only a few B atoms on sublattices 1 and 2, Hence approximately

$$\bar{m}_{12}=(1-\theta_1)+(1-\theta_2)=2\theta_3=1-s.$$

By the same reason we can obtain the number of A-A pairs of neighbours



between the sublattices 1 and 3:

$$N(\theta_1+\theta_3-\bar{m}_{13})=4(\tfrac{1}{2}N\theta_3).$$

Thus $$\bar{m}_{13}=\theta_1-\theta_3=s.$$

We can now write down all the $\bar{m}$'s:

$$\bar{m}_{12}=\bar{m}_{34}=1-s \quad , \quad \bar{m}_{13}=\bar{m}_{23}=\bar{m}_{14}=\bar{m}_{24}=s.$$

These equations are correct to the first order of (1-s). Substituting them into (3) and (4) we obtain

$$L_1=-2N[V'_{AA}(1_1)+V'_{BB}(1_1)-2(1-s)V'(1_1)] \quad ,$$

and $$L_2=-4N[V'_{AA}(1_2)+V'_{BB}(1_2)-2sV'(1_2)] \quad .$$

These are exactly equations (27) in Wilson's paper, from which an expression of the degree of tetragonality in agreement with Gorsky's measurements[2] can be obtained.

3. THE EFFECT OF THE CHANGE OF LATTICE CONSTANTS ON THE INTERACTION ENERGY

In the alloy CuAu the gold atoms and the copper atoms are in contact when the order is perfect. Since the copper atom is somewhat smaller than the gold atom~~s~~, the size of the crystal must increase when gold atoms exchange their positions with copper atoms. Thus with increasing disorder the distance between the atoms increases and hence the interaction energies diminish. The disordering process is therefore effected with more ease near the critical temperature than it is ~~i~~ at lower temperatures; and we expect the specific heat at constant pressure to possess a steeper and higher maximum at the critical temperature than the specific heat at constant volume.

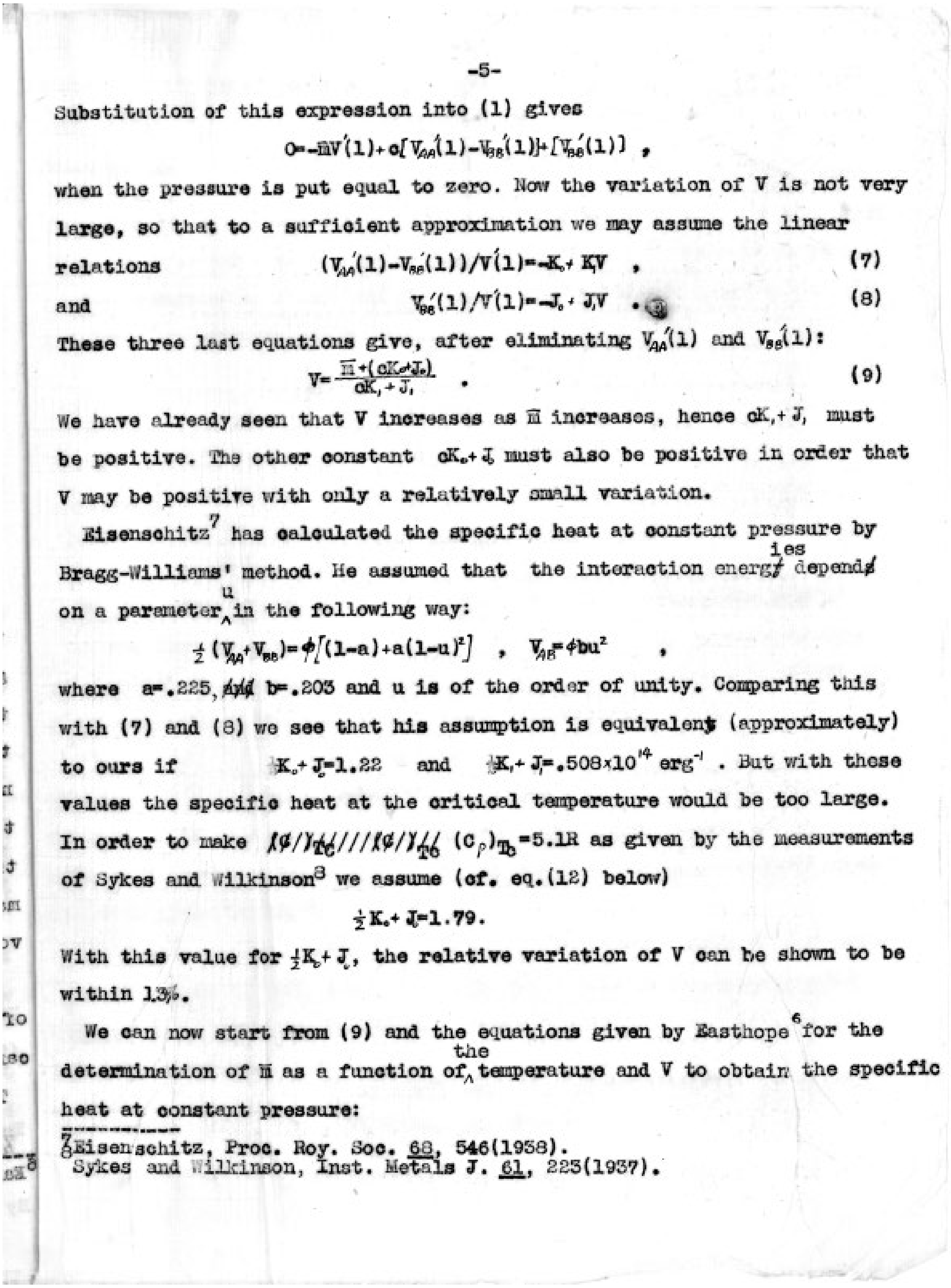
Now we shall calculate in length the specific heat at constant pressure of the alloy β-brass, which forms the simplest type of superlattice that can be studied statistically. Bethe's ~~m~~ method will be used.

The configurational energy of the crystal is, in Easthope's[6] notations:

$$W=-N_{AB}V+\tfrac{1}{2}Nz\{c(V_{AA}-V_{BB})+V_{BB}\} \quad . \qquad (6)$$

---

[6] Easthope, Proc. Camb. Phil. Soc. ~~40~~/ 33, 502(1937); 34, 68(1938).



-5-

Substitution of this expression into (1) gives

$$0=-\overline{\Pi}V'(1)+c[V'_{AA}(1)-V'_{BB}(1)]+[V'_{BB}(1)] ,$$

when the pressure is put equal to zero. Now the variation of V is not very large, so that to a sufficient approximation we may assume the linear relations

$$(V'_{AA}(1)-V'_{BB}(1))/V'(1)=-K_0+K_1V , \qquad (7)$$

and

$$V'_{BB}(1)/V'(1)=-J_0+J_1V . \qquad (8)$$

These three last equations give, after eliminating $V'_{AA}(1)$ and $V'_{BB}(1)$:

$$V=\frac{\overline{\Pi}+(cK_0+J_0)}{cK_1+J_1} . \qquad (9)$$

We have already seen that V increases as $\overline{\Pi}$ increases, hence cK_1+J_1 must be positive. The other constant cK_0+J_0 must also be positive in order that V may be positive with only a relatively small variation.

Eisenschitz[7] has calculated the specific heat at constant pressure by Bragg-Williams' method. He assumed that the interaction energ~~y~~ies depend~~s~~ on a parameter u in the following way:

$$\frac{1}{2}(V_{AA}+V_{BB})=\phi[(1-a)+a(1-u)^2] , \quad V_{AB}=\phi bu^2 ,$$

where $a=.225$, ~~and~~ $b=.203$ and u is of the order of unity. Comparing this with (7) and (8) we see that his assumption is equivalent (approximately) to ours if $\frac{1}{2}K_0+J_0=1.22$ and $\frac{1}{2}K_1+J_1=.508\times10^{14}\ \mathrm{erg}^{-1}$. But with these values the specific heat at the critical temperature would be too large. In order to make ~~(c/)Tc////(c/)Tc~~ $(c_p)_{T_c}=5.1R$ as given by the measurements of Sykes and Wilkinson[8] we assume (cf. eq.(12) below)

$$\frac{1}{2}K_0+J_0=1.79.$$

With this value for $\frac{1}{2}K_0+J_0$, the relative variation of V can be shown to be within 13%.

We can now start from (9) and the equations given by Easthope[6] for the determination of $\overline{\Pi}$ as a function of the temperature and V to obtain the specific heat at constant pressure:

---

[7] Eisenschitz, Proc. Roy. Soc. 68, 546(1938).

[8] Sykes and Wilkinson, Inst. Metals J. 61, 223(1937).

$$C_p=\frac{dW}{dT}=\left(\frac{\partial W}{\partial \bar{m}}\right)_{\ell}\frac{d\bar{m}}{dx}\frac{dx}{dT}+\left(\frac{\partial W}{\partial \ell}\right)_{\bar{m}}\frac{d\ell}{dT}=\left(\frac{\partial W}{\partial \bar{m}}\right)_{\ell}\frac{d\bar{m}}{dx}\frac{dx}{dT} \quad . \tag{10}$$

But

$$\frac{dx}{dT}=(xV/kT^2)/\left(1+\frac{x}{kT}\frac{dV}{d\bar{m}}\frac{d\bar{m}}{dx}\right) \quad . \tag{11}$$

Hence by (6)

$$C_p=\frac{\frac{1}{2}Nzkx(\log x)^2\left(-\frac{d\bar{m}}{dx}\right)}{1+\frac{x\log x}{\bar{m}+(cK_0+J_0)}\left(-\frac{d\bar{m}}{dx}\right)} \quad . \tag{12}$$

The value of this expression is calculated for the case $c=\frac{1}{2}$, the constant $\frac{1}{2}K_0+J_0$ being assumed to be 1.79 to make $(C_p)_{T_c}=5.1R$. The result is plotted in the accompanying figure together with Bethe's curve[9] for C_v and Sykes and Wilkinson's experimental[8] data.

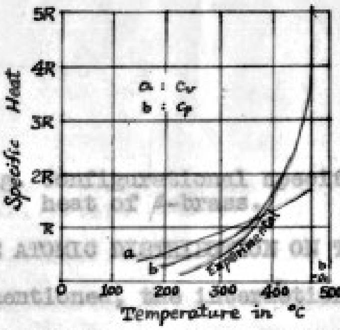


Fig. Configurational specific heat of β-brass.

4. THE EFFECT OF THE ATOMIC DISTRIBUTION ON THE INTERACTION ENERGY

As has already been mentioned the interaction energy depends in some very complicated manner upon the degree of order. To study the effect of such a dependence Lin[4] has assumed a linear relationship:

$$V=V_0(1+\alpha c+\beta m_{AA}) \tag{13}$$

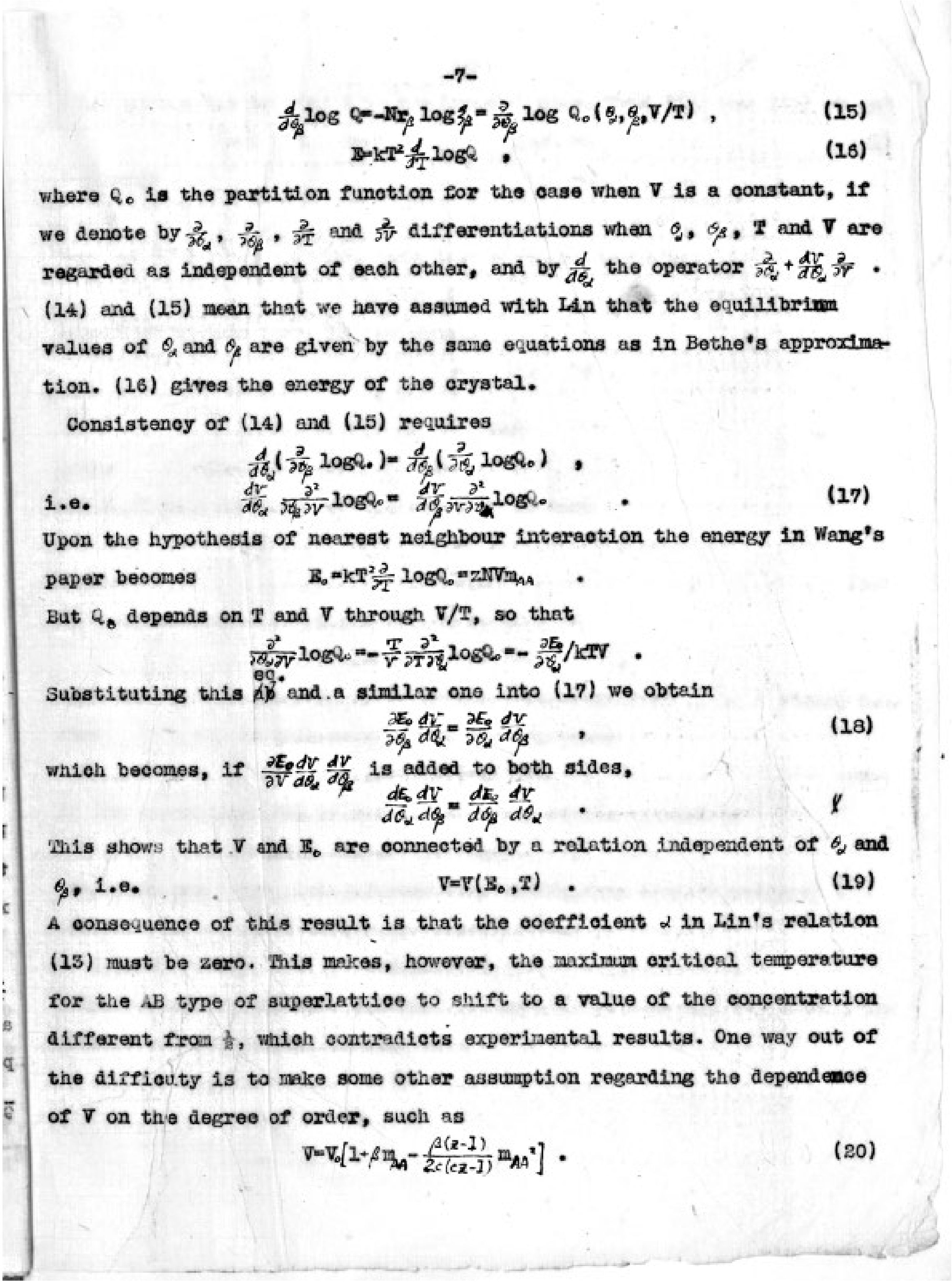
between the interaction energy V and the fraction of A-A pairs of neighbours m_{AA}. In this section we shall study the general nature of the variation of V in the light of the theory of the free energy in Bethe's approximation given by Wang[5].

The fundamental equations in Wang's paper are (45), (46) and (39) with ξ_α and ξ_β given by (47), (48), (49) and (50). These equations are still assumed to be valid now V becomes a function of θ_α, θ_β and T. They may be put into the form:

$$\frac{d}{d\theta_\alpha}\log Q=-Nr_\alpha\log\xi_\alpha=\frac{\partial}{\partial\theta_\alpha}\log Q_0(\theta_\alpha,\theta_\beta,V/T) \quad , \tag{14}$$

---

[9] Nix and Shockley, Rev. Mod. Phys. 10, 1(1938).



$$\frac{d}{d\theta_\beta}\log Q=-Nr_\beta\log\xi_\beta=\frac{\partial}{\partial\theta_\beta}\log Q_0(\theta_\alpha,\theta_\beta,V/T)\;, \tag{15}$$

$$E=kT^2\frac{d}{\partial T}\log Q\;. \tag{16}$$

where Q_0 is the partition function for the case when V is a constant, if we denote by $\frac{\partial}{\partial\theta_\alpha}$, $\frac{\partial}{\partial\theta_\beta}$, $\frac{\partial}{\partial T}$ and $\frac{\partial}{\partial V}$ differentiations when θ_α, θ_β, T and V are regarded as independent of each other, and by $\frac{d}{d\theta_\alpha}$ the operator $\frac{\partial}{\partial\theta_\alpha}+\frac{dV}{d\theta_\alpha}\frac{\partial}{\partial V}$. (14) and (15) mean that we have assumed with Lin that the equilibrimm values of θ_α and θ_β are given by the same equations as in Bethe's approximation. (16) gives the energy of the crystal.

Consistency of (14) and (15) requires

$$\frac{d}{d\theta_\alpha}\left(\frac{\partial}{\partial\theta_\beta}\log Q_0\right)=\frac{d}{d\theta_\beta}\left(\frac{\partial}{\partial\theta_\alpha}\log Q_0\right)\;,$$

i.e.

$$\frac{dV}{d\theta_\alpha}\frac{\partial^2}{\partial\theta_\beta\partial V}\log Q_0=\frac{dV}{d\theta_\beta}\frac{\partial^2}{\partial V\partial\theta_\alpha}\log Q_0\;. \tag{17}$$

Upon the hypothesis of nearest neighbour interaction the energy in Wang's paper becomes

$$E_0=kT^2\frac{\partial}{\partial T}\log Q_0=zNVm_{AA}\;.$$

But Q_0 depends on T and V through V/T, so that

$$\frac{\partial^2}{\partial\theta_\alpha\partial V}\log Q_0=-\frac{T}{V}\frac{\partial^2}{\partial T\partial\theta_\alpha}\log Q_0=-\frac{\partial E_0}{\partial\theta_\alpha}/kTV\;.$$

Substituting this eq. and a similar one into (17) we obtain

$$\frac{\partial E_0}{\partial\theta_\beta}\frac{dV}{d\theta_\alpha}=\frac{\partial E_0}{\partial\theta_\alpha}\frac{dV}{d\theta_\beta}\;, \tag{18}$$

which becomes, if $\frac{\partial E_0}{\partial V}\frac{dV}{d\theta_\alpha}\frac{dV}{d\theta_\beta}$ is added to both sides,

$$\frac{dE_0}{d\theta_\alpha}\frac{dV}{d\theta_\beta}=\frac{dE_0}{d\theta_\beta}\frac{dV}{d\theta_\alpha}\;.$$

This shows that V and E_0 are connected by a relation independent of θ_α and θ_β, i.e.

$$V=V(E_0,T)\;. \tag{19}$$

A consequence of this result is that the coefficient α in Lin's relation (13) must be zero. This makes, however, the maximum critical temperature for the AB type of superlattice to shift to a value of the concentration different from $\frac{1}{2}$, which contradicts experimental results. One way out of the difficuty is to make some other assumption regarding the dependence of V on the degree of order, such as

$$V=V_0\left[1+\beta m_{AA}-\frac{\beta(z-1)}{2c(cz-1)}m_{AA}^2\right]\;. \tag{20}$$

Let us now try to find E in the general case. From (14) and (16) we get

$$\frac{dE}{d\theta_\alpha}=kT^2\frac{d}{dT}\left(\frac{\partial}{\partial\theta_\alpha}\log\Theta_0\right)=kT^2\frac{d\left(\frac{V}{T}\right)}{dT}\frac{\partial^2}{\partial\theta_\alpha\partial\left(\frac{V}{T}\right)}\log\Theta_0=-\frac{T^2}{V}\frac{d\left(\frac{V}{T}\right)}{dT}\frac{\partial E_0}{\partial\theta_\alpha}. \qquad (21)$$

Similarly $\frac{dE}{d\theta_\beta}=-\frac{T^2}{V}\frac{d\left(\frac{V}{T}\right)}{dT}\frac{\partial E_0}{\partial\theta_\beta}$.

Hence $$\frac{dE}{d\theta_\alpha}\frac{\partial E_0}{\partial\theta_\beta}=\frac{\partial E_0}{\partial\theta_\alpha}\frac{dE}{d\theta_\beta}\quad .$$

Just as (18) leads to (19), this last equation leads to

$$E=E(E_0,T)\quad .$$

Substituting this into (21) we obtain

$$\frac{\partial E}{\partial E_0}\left[\frac{\partial E_0}{\partial\theta_\alpha}+\frac{\partial E_0}{\partial V}\frac{\partial V}{\partial\theta_\alpha}\right]=\frac{\partial E_0}{\partial\theta_\alpha}\left(1-\frac{T}{V}\frac{dV}{dT}\right)\quad .$$

But (19) gives $$\frac{dV}{d\theta_\alpha}=\frac{\frac{\partial E}{\partial\theta_\alpha}\frac{\partial V}{\partial E_0}}{1-\frac{\partial V}{\partial E_0}\frac{\partial E_0}{\partial V}}\qquad\text{and}\qquad\frac{dV}{dT}=\frac{\frac{\partial V}{\partial T}+\frac{\partial V}{\partial E_0}\frac{\partial E_0}{\partial T}}{1-\frac{\partial V}{\partial E_0}\frac{\partial E_0}{\partial V}}\quad .$$

Hence $$\frac{\partial E}{\partial E_0}=\left(1-\frac{\partial V}{\partial E_0}\frac{\partial E_0}{\partial V}\right)\left(1-\frac{T}{V}\frac{dV}{dT}\right)=1-\frac{\partial V}{\partial E_0}\frac{\partial E_0}{\partial V}-\frac{T}{V}\frac{dV}{dT}-\frac{T}{V}\frac{\partial V}{\partial E_0}\frac{\partial E_0}{\partial T}\quad .$$

Now E_0/V is a function of θ_α, θ_β and V/T, so that

$$\frac{\partial E_0}{\partial V}=-\frac{T}{V}\frac{\partial E_0}{\partial T}+\frac{E_0}{V}\quad .$$

Hence $$\frac{\partial E}{\partial E_0}=1-\frac{E_0}{V}\frac{\partial V}{\partial E_0}-\frac{T}{V}\frac{\partial V}{\partial T}\quad . \qquad (22)$$

If V depends on m_{AA} only, and not on T, we have

$$\frac{\partial E}{\partial E_0}=1-\frac{E_0}{V}\frac{dV}{dE_0}=V\frac{d(E_0/V)}{dE_0}\quad .$$

The boundary condition is given by the case when there is no A atoms, i.e. when $\theta_\alpha+\theta_\beta=0$. In this case $m_{AA}=0$, $E_0=E=0$. Hence

$$E=\int_0^{E_0}\left(V\frac{d(E_0/V)}{dE_0}\right)dE_0=\cancel{E\,d\log(E_0/V)}=zN\int_0^{m_{AA}}V\,dm_{AA}\quad . \qquad (23)$$

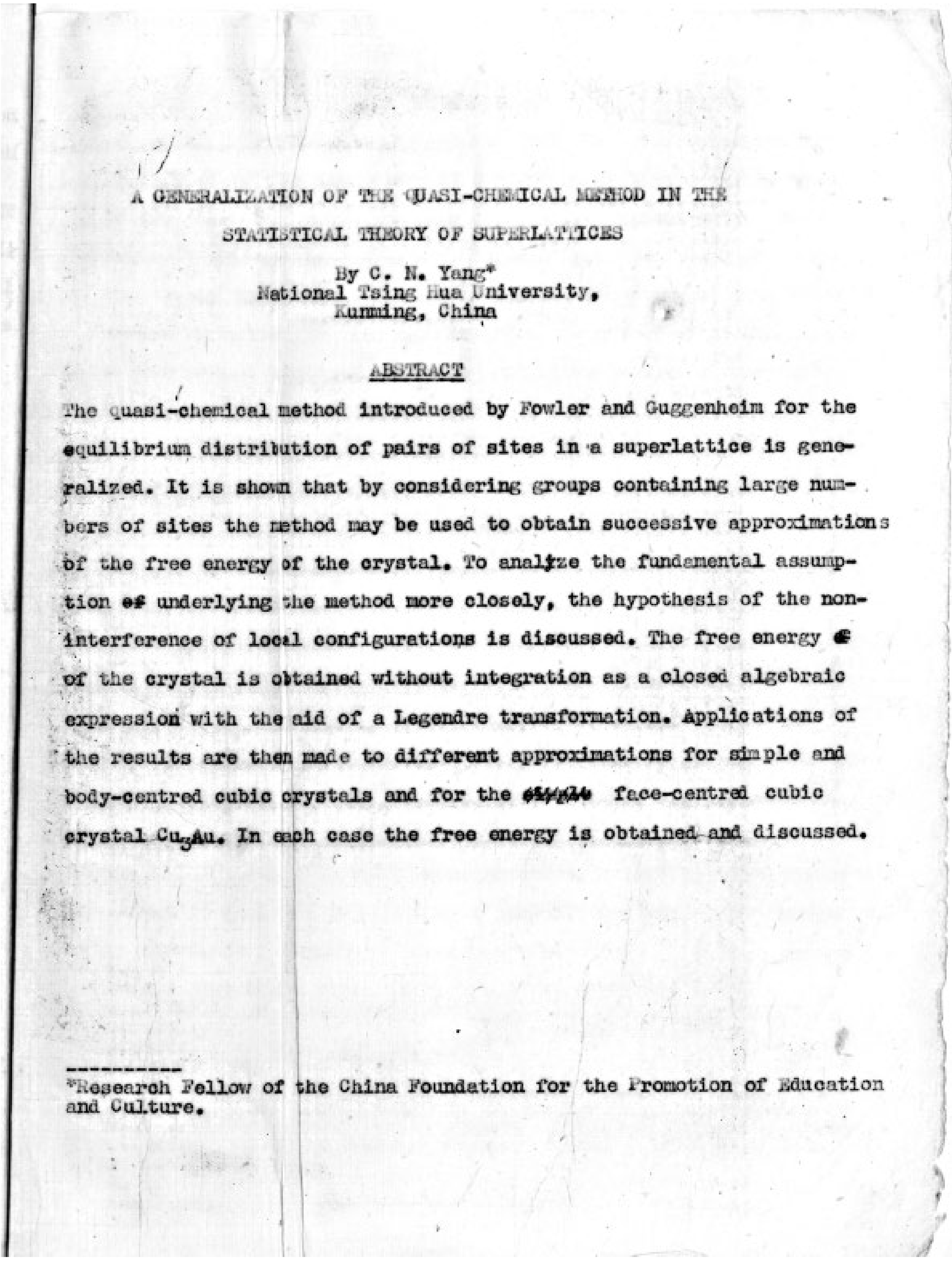
If the assumption (20) is made, the energy of the crystal is

$$E=zNV_0\left[m_{AA}+\frac{\beta}{2}m_{AA}^2-\frac{\beta(z-1)}{6c(cz-1)}m_{AA}^3\right]\quad .$$

When β is not large, this differs very little in numerical value from Bethe's original expression. The specific heat is

$$C_V=zNV\frac{dm_{AA}}{dT}=zNV\frac{dm_{AA}}{dx}\frac{\frac{xV}{kT^2}}{1+\frac{x}{kT}\frac{dV}{dx}}=zNkx(\log x)^2\frac{dm_{AA}}{dx}\frac{1}{1+\frac{x}{kT}\frac{dV}{dm_{AA}}\frac{dm_{AA}}{dx}}\quad .$$

The author is very much indebted to Prof. J. S. Wang for suggesting this problem and for helpful discussions.



A GENERALIZATION OF THE QUASI-CHEMICAL METHOD IN THE STATISTICAL THEORY OF SUPERLATTICES

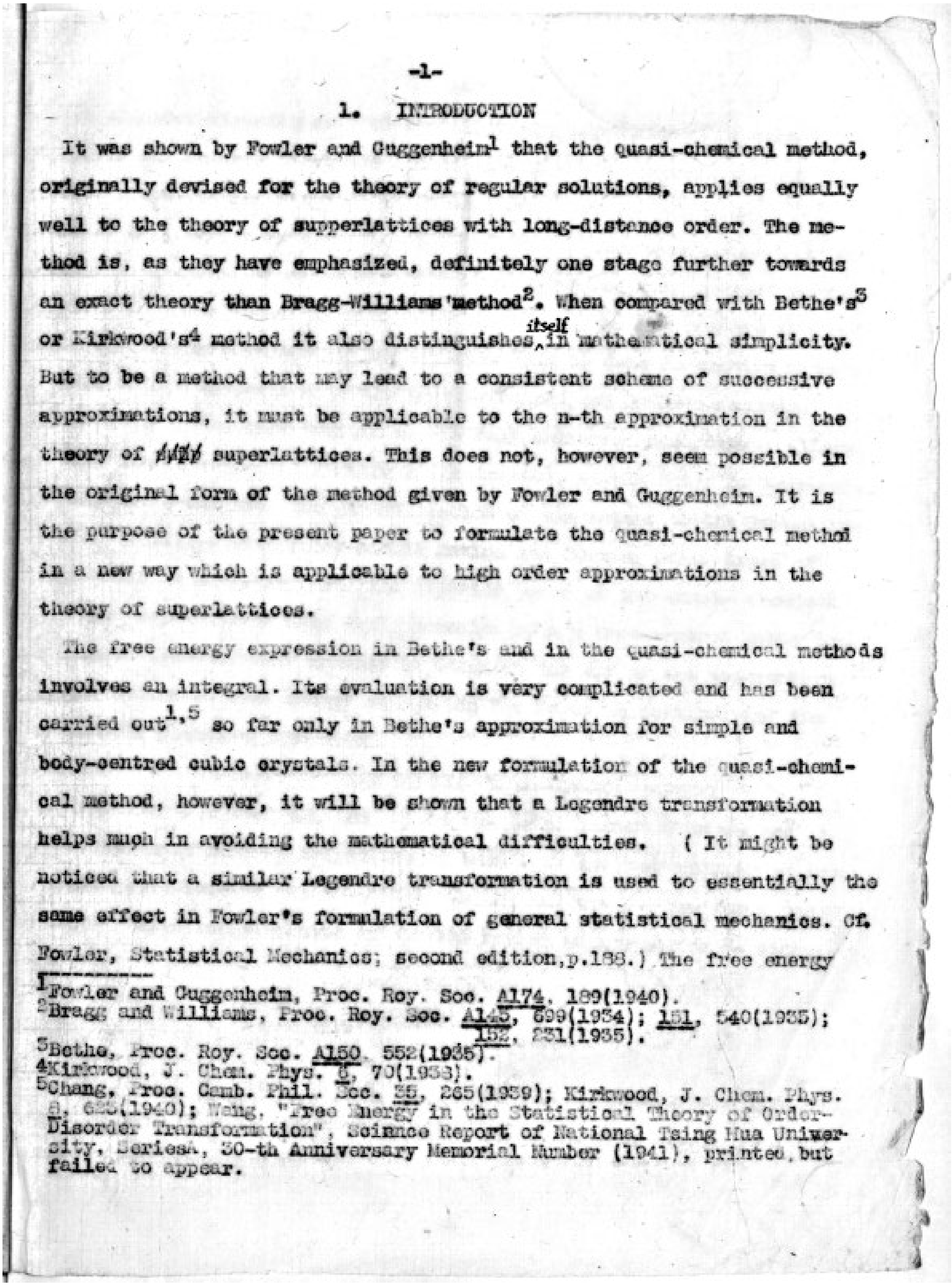
By C. N. Yang[*]
National Tsing Hua University,
Kunming, China

ABSTRACT

The quasi-chemical method introduced by Fowler and Guggenheim for the equilibrium distribution of pairs of sites in a superlattice is generalized. It is shown that by considering groups containing large numbers of sites the method may be used to obtain successive approximations of the free energy of the crystal. To analyze the fundamental assumption ~~of~~ underlying the method more closely, the hypothesis of the non-interference of local configurations is discussed. The free energy of the crystal is obtained without integration as a closed algebraic expression with the aid of a Legendre transformation. Applications of the results are then made to different approximations for simple and body-centred cubic crystals and for the ~~[illegible]~~ face-centred cubic crystal Cu_3Au. In each case the free energy is obtained and discussed.

---

[*] Research Fellow of the China Foundation for the Promotion of Education and Culture.



1. INTRODUCTION

It was shown by Fowler and Guggenheim[1] that the quasi-chemical method, originally devised for the theory of regular solutions, applies equally well to the theory of supperlattices with long-distance order. The method is, as they have emphasized, definitely one stage further towards an exact theory than Bragg-Williams' method[2]. When compared with Bethe's[3] or Kirkwood's[4] method it also distinguishes itself in mathematical simplicity. But to be a method that may lead to a consistent scheme of successive approximations, it must be applicable to the n-th approximation in the theory of ~~[illegible]~~ superlattices. This does not, however, seem possible in the original form of the method given by Fowler and Guggenheim. It is the purpose of the present paper to formulate the quasi-chemical method in a new way which is applicable to high order approximations in the theory of superlattices.

The free energy expression in Bethe's and in the quasi-chemical methods involves an integral. Its evaluation is very complicated and has been carried out[1,5] so far only in Bethe's approximation for simple and body-centred cubic crystals. In the new formulation of the quasi-chemical method, however, it will be shown that a Legendre transformation helps much in avoiding the mathematical difficulties. (It might be noticed that a similar Legendre transformation is used to essentially the same effect in Fowler's formulation of general statistical mechanics. Cf. Fowler, Statistical Mechanics; second edition, p.188.) The free energy

---

[1] Fowler and Guggenheim, Proc. Roy. Soc. A174, 189(1940).

[2] Bragg and Williams, Proc. Roy. Soc. A145, 699(1934); 151, 540(1935); 152, 231(1935).

[3] Bethe, Proc. Roy. Soc. A150, 552(1935).

[4] Kirkwood, J. Chem. Phys. 6, 70(1938).

[5] Chang, Proc. Camb. Phil. Soc. 35, 265(1939); Kirkwood, J. Chem. Phys. 8, 623(1940); Wang, "Free Energy in the Statistical Theory of Order-Disorder Transformation", Scinnce Report of National Tsing Hua University, Series A, 30-th Anniversary Memorial Number (1941), printed but failed to appear.

is obtained directly as a closed algebraic expression. Its values are given for Bethe's first and second (modified) approximations and for the face-centred alloy Cu_3Au in sections 7 and 8.

To make sure that the quasi-chemical method may actually be used to obtain a series of successively better approximations, we must investigate the free energy in high order approximations and compare it with the partition function of the crystal. This is done in section 5 together with a comparison of the quasi-chemical and Bethe's methods.

Except in the last section we are only concerned with binary alloys with atomic ratio 1:1 forming a (quadratic), simple cubic or body-centred cubic lattice. The generalization of the method to the investigation of alloys with other atomic ratios and forming other types of lattices is easy. In fact, the superior power of the quasi- chemical method appears to be even more fully revealed when a face-centred cubic lattice is treated. This problem is taken at the end of the paper where an approximate free energy expression for Cu_3Au is obtained and its critical phenomena discussed.

2. REFORMULATION OF THE QUASI-CHEMICAL METHOD

Consider a crystal AB. Let $2N$ be the total number of atoms and z the number of nearest neighbours of each . At low temperatures we can distinguish between the so-called α-sites and ~~the~~ β-sites for A and B atoms respectively. Denote by Nr the number of A atoms on α- sites. Let

$$1-r=w \quad , \quad r-w=s .$$

The partition function of the crystal is

$$\sum_r p(r,T) \quad , \qquad (1)$$

where $p(r,T)$ is equal to $\sum \exp(-E/kT)$ over all possible configurations of the crystal with the given value of r. The average energy over all these configurations is

$$\overline{E}(r,T) = kT^2\frac{\partial}{\partial T}\log p(r,T), \qquad (2)$$

But evidently,

$$p(r,\infty) = g(r) = \left[\frac{N!}{(Nr)!\,(Nw)!}\right]^2. \qquad (3)$$

Hence

$$\log p(r,T) = \log g(r) + \int_{\infty}^{T}\frac{1}{kT^2}\overline{E}(r,T)\,dT. \qquad (4)$$

The problem therefore reduces to one of finding $\overline{E}(r,T)$. Since a direct solution is very difficult we shall try to find an approximate solution by the quasi-chemical method, which is presented below in a form easily generalizable.

There are in the crystal $3N$ nearest pairs of sites α-β. Among these let there be $[q_\alpha, q_\beta]$ with their α-sites occupied by q_α(=0,1) wrong (B) atoms and their β-sites by q_β(=0,1) wrong (A) atoms. For given r the following relations hold:

$$\begin{aligned} [0,1] + [0,0] + [1,1] + [1,0] &= 3N, \\ [1,1] + [1,0] &= 3Nw, \qquad (5) \end{aligned}$$

and

$$[0,1] \qquad + [1,1] \qquad = 3Nw.$$

Upon the approximation of ~~the~~ neighbour interaction the energy of the crystal may be written as

$$E(r,T) = [0,1]V_{AA} + [0,0]V_{AB} + [1,1]V_{AB} + [1,0]V_{BB}, \qquad (6)$$

where the V's are the interaction energies between a pair of nearest neighbours.

We may give (5) and(6) a different interpretation by imagining [0,1], [0,0], [1,1], [1,0] and V_{AA}, V_{AB}, V_{AB}, V_{BB} to be respectively the numbers and the molecular internal energies of the four different kinds of molecules XZ, X, XYZ, XY of an ~~assembly~~ gaseous assembly. The interpretation of (6) is that the assembly has the same internal (non-kinetic) energy as the crystal at the given value of r. (5) would mean that there are altogether $3N$ X atoms, $3Nw$ Y atoms and $3Nw$ Z atoms in the assembly.

The quasi- chemical method consists in taking the averages $\overline{[0,1]}$,

$\overline{[0,0]}$, $\overline{[1,1]}$, $\overline{[1,0]}$ of the assembly at any temperature as approximately representing the corresponding averages of the crystal at the same temperature. Whether this approximation is good can only be judged for the present from the results it leads to.

A detail ed treatment of the problem of a gaseous assembly has been given by Fowler[6] . We are only interested in our assembly of four different kinds of molecules, for which the results may be summerized as:

$$\overline{[0,1]} = \xi\nu e^{-V_{AA}/kT}, \qquad \overline{[0,0]} = \xi e^{-V_{AB}/kT},$$
$$\overline{[1,1]} = \xi\mu\nu e^{-V_{AB}/kT}, \qquad \overline{[1,0]} = \xi\mu e^{-V_{BB}/kT}, \tag{7}$$

where ξ, μ and ν are to be determined from (6). From (7) we get

$$\overline{[0,0]}\,\overline{[1,1]} / \overline{[0,1]}\,\overline{[1,0]} = x^{-2}, \tag{8}$$

where

$$x = \exp[-\tfrac{1}{2}(V_{AA}+V_{BB}-2V_{AB})/kT]. \tag{9}$$

(8) and (6) together form the starting point of Fowler and Guggenheim's work[1]. The subsequent calculations of $E(r,T)$, $p(r,T)$ and the free energy of the crystal are straightforward and will not be repeated here. We shall see later how the free energy can be written down directly without actual integration.

3. GENERALIZATION TO GROUPS OF FOUR SITES

So far we have fixed our attention on the pairs of nearest neighbours in the crystal and have used the quasi-chemical method to obtain the average numbers of the four different kinds of pairs. Now we shall generalize the whole procedure: we shall study all the groups of sites in the crystal of an arbitrarily chosen form in, and classifying these groups according to the way they are occupied by atoms we shall obtain the average number of groups in each class by chemical analogy.

To make this clear let us consider in detail groups of four sites forming squares (as shown) in a quadratic lattice. We classify these groups into 2^4=16 classes

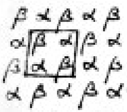
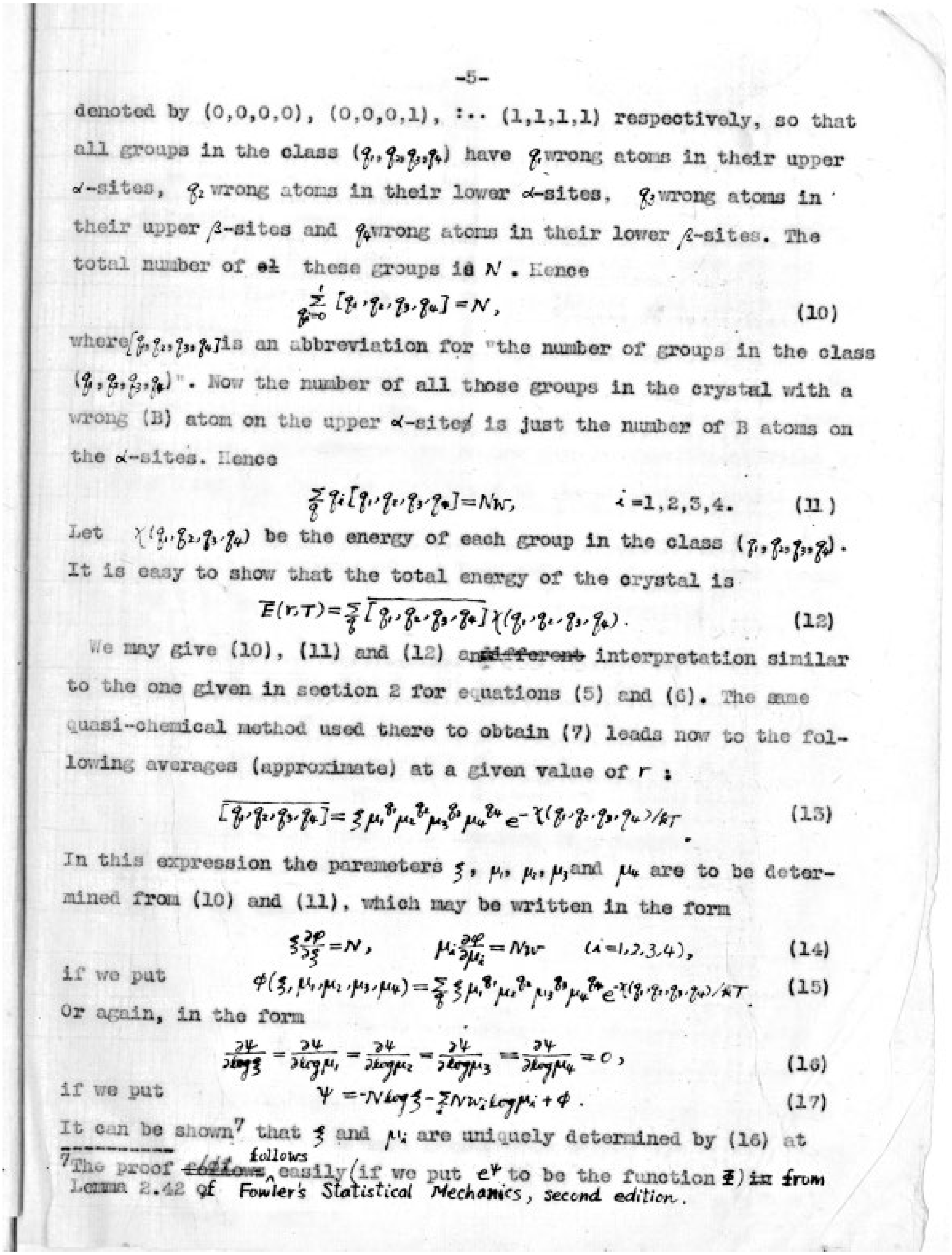


Fig. 1

---

[6] Fowler, Statistical Mechanics, second edition, pp.162-163.



denoted by (0,0,0,0), (0,0,0,1), :·· (1,1,1,1) respectively, so that all groups in the class (q_1, q_2, q_3, q_4) have q_1 wrong atoms in their upper α-sites, q_2 wrong atoms in their lower α-sites, q_3 wrong atoms in their upper β-sites and q_4 wrong atoms in their lower β-sites. The total number of ~~al~~ these groups is N. Hence

$$\sum_{q_i=0}^{1} [q_1, q_2, q_3, q_4] = N, \qquad (10)$$

where $[q_1, q_2, q_3, q_4]$ is an abbreviation for "the number of groups in the class (q_1, q_2, q_3, q_4)". Now the number of all those groups in the crystal with a wrong (B) atom on the upper α-site~~s~~ is just the number of B atoms on the α-sites. Hence

$$\sum_{q} q_i [q_1, q_2, q_3, q_4] = N w_i, \qquad i=1,2,3,4. \qquad (11)$$

Let $\chi(q_1, q_2, q_3, q_4)$ be the energy of each group in the class (q_1, q_2, q_3, q_4). It is easy to show that the total energy of the crystal is

$$E(r,T) = \sum_{q} \overline{[q_1, q_2, q_3, q_4]}\, \chi(q_1, q_2, q_3, q_4). \qquad (12)$$

We may give (10), (11) and (12) an ~~different~~ interpretation similar to the one given in section 2 for equations (5) and (6). The same quasi-chemical method used there to obtain (7) leads now to the following averages (approximate) at a given value of r :

$$\overline{[q_1, q_2, q_3, q_4]} = \xi \mu_1^{q_1} \mu_2^{q_2} \mu_3^{q_3} \mu_4^{q_4} e^{-\chi(q_1, q_2, q_3, q_4)/kT} \qquad (13)$$

In this expression the parameters ξ, μ_1, μ_2, μ_3 and μ_4 are to be determined from (10) and (11), which may be written in the form

$$\xi \frac{\partial \phi}{\partial \xi} = N, \qquad \mu_i \frac{\partial \phi}{\partial \mu_i} = N w_i \qquad (i=1,2,3,4), \qquad (14)$$

if we put

$$\phi(\xi, \mu_1, \mu_2, \mu_3, \mu_4) = \sum_{q} \xi \mu_1^{q_1} \mu_2^{q_2} \mu_3^{q_3} \mu_4^{q_4} e^{-\chi(q_1, q_2, q_3, q_4)/kT}. \qquad (15)$$

Or again, in the form

$$\frac{\partial \psi}{\partial \log \xi} = \frac{\partial \psi}{\partial \log \mu_1} = \frac{\partial \psi}{\partial \log \mu_2} = \frac{\partial \psi}{\partial \log \mu_3} = \frac{\partial \psi}{\partial \log \mu_4} = 0, \qquad (16)$$

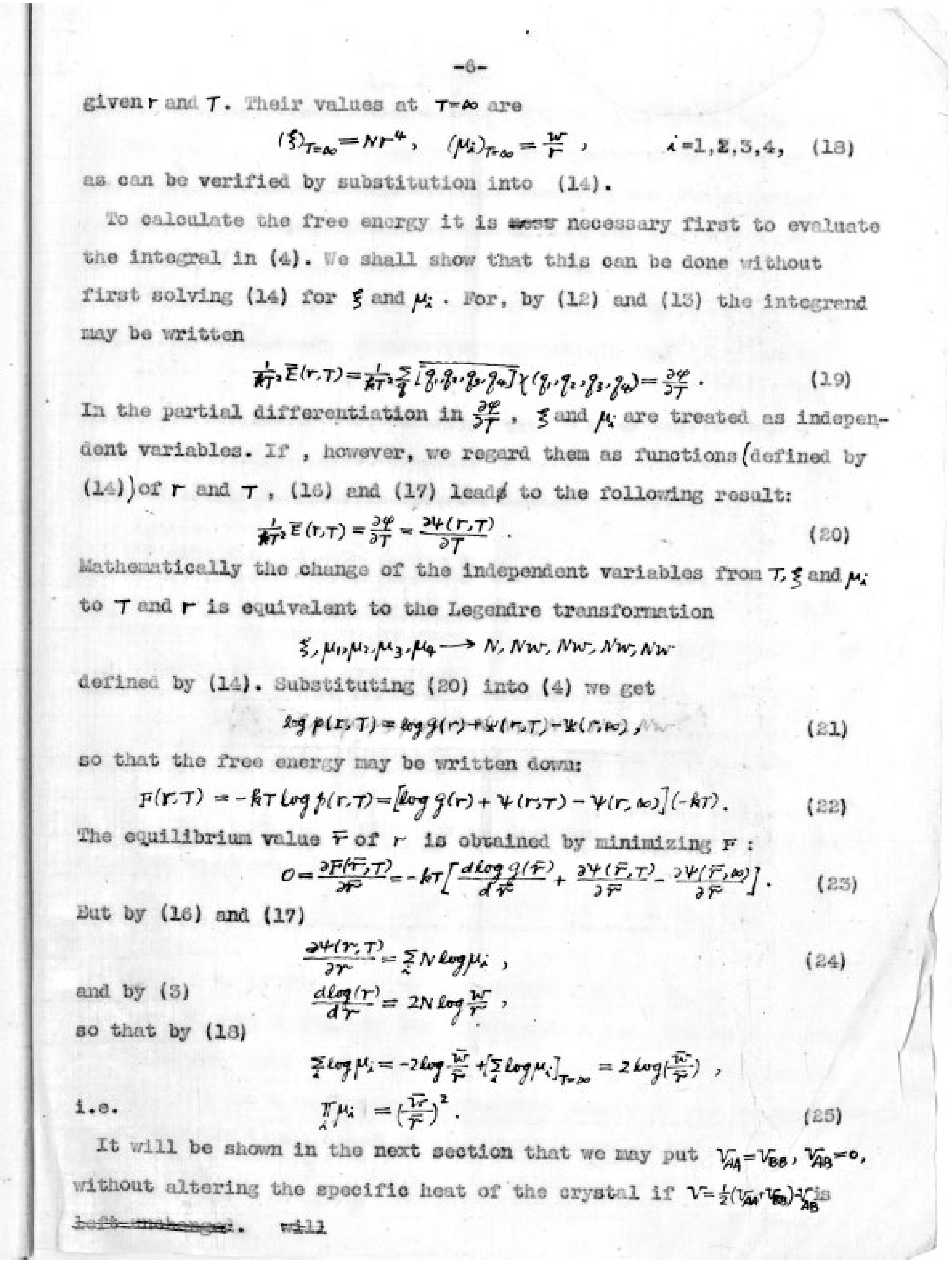
if we put

$$\psi = -N \log \xi - \sum_{i} N w_i \log \mu_i + \phi. \qquad (17)$$

It can be shown[7] that ξ and μ_i are uniquely determined by (16) at

---

[7] The proof ~~follows~~ follows easily (if we put e^{ψ} to be the function Φ) ~~is~~ from Lemma 2.42 of Fowler's Statistical Mechanics, second edition.



given r and T. Their values at $T=\infty$ are

$$(\xi)_{T=\infty}=Nr^{4}, \quad (\mu_i)_{T=\infty}=\frac{w}{r}, \qquad i=1,2,3,4, \tag{18}$$

as can be verified by substitution into (14).

To calculate the free energy it is ~~nece~~ necessary first to evaluate the integral in (4). We shall show that this can be done without first solving (14) for ξ and μ_i. For, by (12) and (13) the integrand may be written

$$\frac{1}{kT^2}\bar{E}(r,T)=\frac{1}{kT^2}\sum_{q}\overline{[q_1,q_2,q_3,q_4]}\chi(q_1,q_2,q_3,q_4)=\frac{\partial\varphi}{\partial T}. \tag{19}$$

In the partial differentiation in $\frac{\partial\varphi}{\partial T}$, ξ and μ_i are treated as independent variables. If, however, we regard them as functions (defined by (14)) of r and T, (16) and (17) lead~~s~~ to the following result:

$$\frac{1}{kT^2}\bar{E}(r,T)=\frac{\partial\varphi}{\partial T}=\frac{\partial\psi(r,T)}{\partial T}. \tag{20}$$

Mathematically the change of the independent variables from T, ξ and μ_i to T and r is equivalent to the Legendre transformation

$$\xi, \mu_1, \mu_2, \mu_3, \mu_4 \longrightarrow N, Nw, Nw, Nw, Nw$$

defined by (14). Substituting (20) into (4) we get

$$\log p(r,T)=\log g(r)+\psi(r,T)-\psi(r,\infty), \tag{21}$$

so that the free energy may be written down:

$$F(r,T)=-kT\log p(r,T)=[\log g(r)+\psi(r,T)-\psi(r,\infty)](-kT). \tag{22}$$

The equilibrium value $\bar{r}$ of r is obtained by minimizing F:

$$0=\frac{\partial F(\bar{r},T)}{\partial\bar{r}}=-kT\left[\frac{d\log g(\bar{r})}{d\bar{r}}+\frac{\partial\psi(\bar{r},T)}{\partial\bar{r}}-\frac{\partial\psi(\bar{r},\infty)}{\partial\bar{r}}\right]. \tag{23}$$

But by (16) and (17)

$$\frac{\partial\psi(r,T)}{\partial r}=\sum_i N\log\mu_i, \tag{24}$$

and by (3)

$$\frac{d\log g(r)}{dr}=2N\log\frac{w}{r},$$

so that by (18)

$$\sum_i\log\mu_i=-2\log\frac{w}{\bar{r}}+\left[\sum_i\log\mu_i\right]_{T=\infty}=2\log\left(\frac{w}{\bar{r}}\right),$$

i.e.

$$\prod_i\mu_i=\left(\frac{w}{\bar{r}}\right)^2. \tag{25}$$

It will be shown in the next section that we may put $V_{AA}=V_{BB}$, $V_{AB}=0$, without altering the specific heat of the crystal if $V=\frac{1}{2}(V_{AA}+V_{BB})-V_{AB}$ is ~~left unchanged~~. will left unchanged. When this is done , ϕ will be symmetrical with respect to ~~the~~ μ_1, μ_2, μ_3 and μ_4 , and we conclude that all the μ's are equal from the facts that (i) equation (14) has only one set of solution[7], and (ii) if the conclusion is true (14) becomes, with all μ_i put equal to μ ,

$$\xi\frac{\partial\phi}{\partial\xi} = N \quad , \qquad \mu\frac{\partial\phi}{\partial\mu} = 4Nw, \tag{26}$$

which does have[7] a set of solution in ξ and μ . Now ϕ is given by

$$\phi = \xi\left[1+4\mu x^2+(4\mu^2x^2+2\mu^2x^4)+4\mu^3x^2+\mu^4\right], \tag{27}$$

where x is defined by (9). On eliminating ξ from (26) we obtain

$$(1+s)\mu^4+(2+4s)x^2\mu^3+2sx^2(x^2+2)\mu^2+(4s-2)x^2\mu+(s-1)=0. \tag{28}$$

The free energy is given by (21) and (18):

$$-\frac{F(r,T)}{2NkT} = r\log r+w\log w-2w\log\mu+\tfrac{1}{2}\log(1+4\mu x^2+4\mu^2x^2+2\mu^2x^4+4\mu^3x^2+\mu^4), \tag{29}$$

and the condition of equilibrium by (25):

$$\mu=\sqrt{\frac{1-s}{1+s}} \ . \tag{30}$$

To obtain the critical temperature, we expand (28) in powers of s and find after identifying coefficients

$$\log\mu = -\frac{1+6x^2+x^4}{2+2x^2}s+Ks^3+\cdots ,$$

which is the only real solution for $\log\mu$. Next we expand (30):

$$\log\mu = -s-\tfrac{1}{3}s^3\cdots .$$

At the critical value x_c of x , these last two equations have a multiple solution at s=0. Hence

$$-\frac{1+6x_c^2+x_c^4}{2+2x_c^2} = -1 \quad ,$$

i.e.

$$x_c = (\sqrt{5}-2)^{1/2} = .4858 \ .$$

4. GENERAL FORM OF THE QUASI-CHEMICAL METHOD

Let us now take a group of any size and form. Let it have a α-sites b β-sites and γ pairs of nearest neighbours. The procedures to obtain an approximate expression for the free energy of the crystal follow exactly the same line as in the special case considered in the last section. Equations (13), (14) and (18) are essentially unchanged:

$$\overline{[q_1, q_2, \cdots]} = \xi \mu_1^{q_1} \mu_2^{q_2} \cdots e^{-\chi/kT}, \tag{31}$$

$$\phi = \sum_q \xi \mu_1^{q_1} \mu_2^{q_2} \cdots e^{-\chi/kT}, \quad \xi \frac{\partial \phi}{\partial \xi} = N, \quad \mu_i \frac{\partial \phi}{\partial \mu_i} = N w_i \tag{32}$$

and

$$(\mu_i)_{T=\infty} = \frac{w}{r}. \tag{33}$$

But (12) should be corrected by a factor $\frac{\gamma}{3}$ to account for the duplications in calculating E from the sum of the energies of all the groups in the crystal:

$$\overline{E} = \frac{3}{\gamma} \sum_q \overline{[q_1, q_2, \cdots]}\, \chi(q_1, q_2, \cdots) = \frac{3}{\gamma} kT^2 \frac{\partial \phi}{\partial T}. \tag{34}$$

Hence (22) becomes*

$$F(r,T) = -kT\left[\log g(r) + \frac{3}{\gamma}\psi(r,T) - \frac{3}{\gamma}\psi(r,\infty)\right], \tag{35}$$

or more explicitly, by (3), (17), and (33):

$$F(r,T) = -\frac{3NkT}{\gamma}\left[\log N + \left(a+b-\frac{2\gamma}{3}\right)(r\log r + w\log w) - \log\xi - w\sum_i \log\mu_i\right]. \tag{36}$$

The derivative is

$$\frac{\partial}{\partial r}F(r,T) = -\frac{3NkT}{\gamma}\log\left[\left(\prod_i \mu_i\right)\left(\frac{r}{w}\right)^{a+b-\frac{2\gamma}{3}}\right], \tag{37}$$

so that the condition of equilibrium is

$$\prod_i \mu_i = (\overline{w}/\overline{r})^{a+b-\frac{2\gamma}{3}}. \tag{38}$$

In actual calculations the following points may prove helpful:

(1) The free energy is changed by a constant if V_{AA} and V_{BB} are both replaced by $\frac{1}{2}(V_{AA}+V_{BB}) - V_{AB}$, and V_{AB} by 0. To prove this let z_i be the number of sites in the group neighbouring to the site i. Let χ be changed into χ' by the replacement. It is evident that

$$\chi' - \chi = -\gamma V_{AB} + \frac{V_{AA} - V_{BB}}{2}(\text{no. of B-B pairs} - \text{no. of AA pairs}),$$

and that

$$\sum_{\alpha\text{-sites}} q_i z_i - \sum_{\beta\text{-sites}} q_i z_i = \text{no. of B-B pairs} - \text{no. of A-A pairs}.$$

Hence

$$\overline{[q_1, q_2, \cdots]} = \xi \mu_1^{q_1} \mu_2^{q_2} \cdots e^{-\chi/kT} = \xi' \mu_1'^{q_1} \mu_2'^{q_2} \cdots e^{-\chi'/kT},$$

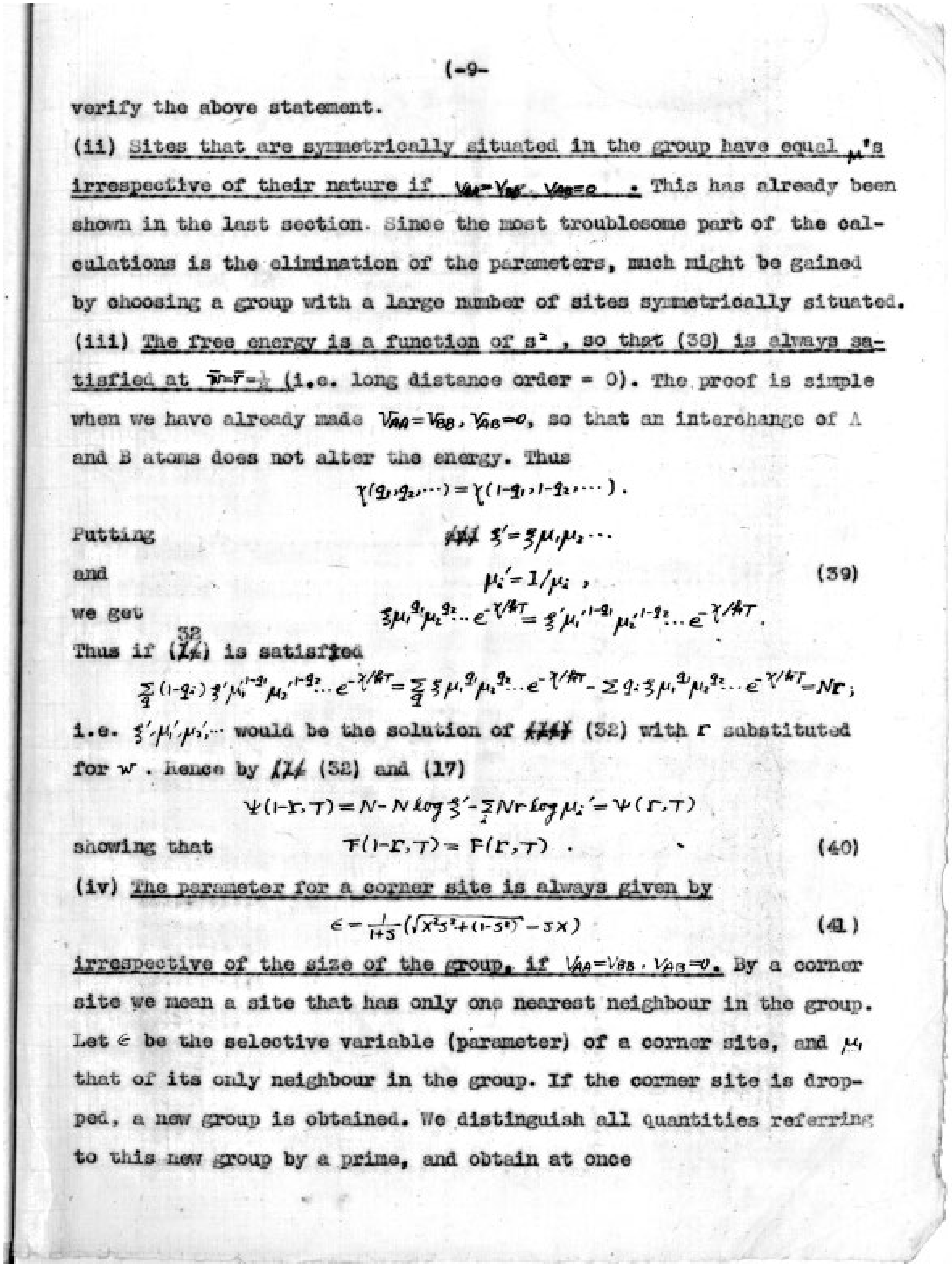
if we put

$$\xi' = \xi e^{-\gamma V_{AB}/kT}, \quad \mu_i' = \mu_i e^{\pm z_i (V_{AA} - V_{BB})/2kT}$$

where the + sign or the - sign is to be taken according as the site i is an α or a β site. We can now calculate the new free energy and

---

*Care must be taken when the theory is extended to the case when the atomic ratio is not 1:1. The function Ψ in (35) must then be replaced by $\frac{1}{2}\Psi + \frac{1}{2}\Psi'$ where Ψ' is the function Ψ for the case when the group of interest has the same form as the original one but with α and β sites interchanged.



verify the above statement.

(ii) Sites that are symmetrically situated in the group have equal μ's irrespective of their nature if $V_{AA}=V_{BB}$, $V_{AB}=0$. This has already been shown in the last section. Since the most troublesome part of the calculations is the elimination of the parameters, much might be gained by choosing a group with a large number of sites symmetrically situated.

(iii) The free energy is a function of s^2 , so that (38) is always satisfied at $\bar{w}=\bar{r}=\frac{1}{2}$ (i.e. long distance order = 0). The proof is simple when we have already made $V_{AA}=V_{BB}$, $V_{AB}=0$, so that an interchange of A and B atoms does not alter the energy. Thus

$$\chi(q_1,q_2,\cdots)=\chi(1-q_1,1-q_2,\cdots).$$

Putting $$\xi'=\xi\mu_1\mu_2\cdots$$

and $$\mu_i'=1/\mu_i\ , \tag{39}$$

we get $$\xi\mu_1^{q_1}\mu_2^{q_2}\cdots e^{-\chi/kT}=\xi'\mu_1'^{1-q_1}\mu_2'^{1-q_2}\cdots e^{-\chi/kT}.$$

Thus if ($\overset{32}{\cancel{14}}$) is satisfied

$$\sum_q(1-q_i)\xi'\mu_1'^{1-q_1}\mu_2'^{1-q_2}\cdots e^{-\chi/kT}=\sum_q\xi\mu_1^{q_1}\mu_2^{q_2}\cdots e^{-\chi/kT}-\sum q_i\xi\mu_1^{q_1}\mu_2^{q_2}\cdots e^{-\chi/kT}=Nr;$$

i.e. $\xi',\mu_1',\mu_2',\cdots$ would be the solution of ~~(14)~~ (32) with r substituted for w . Hence by ~~(14)~~ (32) and (17)

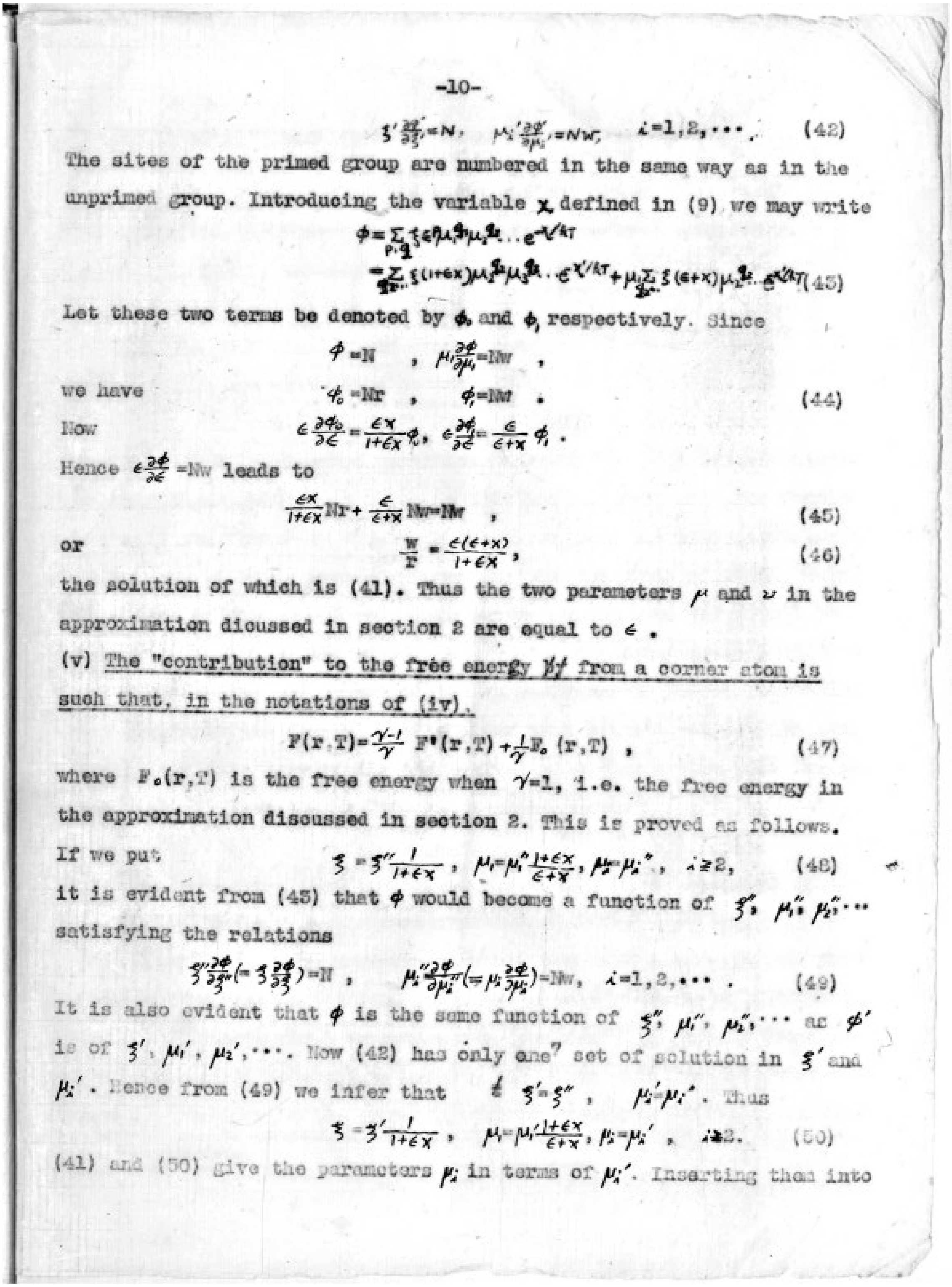
$$\Psi(1-r,T)=N-N\log\xi'-\sum_i Nr\log\mu_i'=\Psi(r,T)$$

showing that $$F(1-r,T)=F(r,T)\ . \tag{40}$$

(iv) The parameter for a corner site is always given by

$$\epsilon=\frac{1}{1+s}\left(\sqrt{x^2s^2+(1-s^2)}-sx\right) \tag{41}$$

irrespective of the size of the group, if $V_{AA}=V_{BB}$, $V_{AB}=0$. By a corner site we mean a site that has only one nearest neighbour in the group. Let ϵ be the selective variable (parameter) of a corner site, and μ_1 that of its only neighbour in the group. If the corner site is dropped, a new group is obtained. We distinguish all quantities referring to this new group by a prime, and obtain at once



-10-

$$\xi'\frac{\partial\phi'}{\partial\xi'}=N,\quad \mu_i'\frac{\partial\phi'}{\partial\mu_i'}=Nw,\qquad i=1,2,\cdots. \tag{42}$$

The sites of the primed group are numbered in the same way as in the unprimed group. Introducing the variable x defined in (9) we may write

$$\begin{aligned}\phi&=\sum_{p,q}\xi\epsilon^{p}\mu_1^{q_1}\mu_2^{q_2}\cdots e^{-V/kT}\\&=\sum_{q_2\cdots}\xi(1+\epsilon x)\mu_2^{q_2}\mu_3^{q_3}\cdots e^{V/kT}+\mu_1\sum_{q_2\cdots}\xi(\epsilon+x)\mu_2^{q_2}\cdots e^{V/kT}\end{aligned}\tag{43}$$

Let these two terms be denoted by ϕ_0 and ϕ_1 respectively. Since

$$\phi=N\quad,\quad \mu_1\frac{\partial\phi}{\partial\mu_1}=Nw\quad,$$

we have
$$\phi_0=Nr\quad,\quad \phi_1=Nw\quad. \tag{44}$$

Now
$$\epsilon\frac{\partial\phi_0}{\partial\epsilon}=\frac{\epsilon x}{1+\epsilon x}\phi_0,\quad \epsilon\frac{\partial\phi_1}{\partial\epsilon}=\frac{\epsilon}{\epsilon+x}\phi_1\ .$$

Hence $\epsilon\frac{\partial\phi}{\partial\epsilon}=Nw$ leads to

$$\frac{\epsilon x}{1+\epsilon x}Nr+\frac{\epsilon}{\epsilon+x}Nw=Nw\quad, \tag{45}$$

or
$$\frac{w}{r}=\frac{\epsilon(\epsilon+x)}{1+\epsilon x}, \tag{46}$$

the solution of which is (41). Thus the two parameters μ and ν in the approximation dicussed in section 2 are equal to ϵ.

(v) The "contribution" to the free energy from a corner atom is such that, in the notations of (iv),

$$F(r,T)=\frac{\gamma-1}{\gamma}F^{*}(r,T)+\frac{1}{\gamma}F_0(r,T)\ , \tag{47}$$

where $F_0(r,T)$ is the free energy when $\gamma=1$, i.e. the free energy in the approximation discussed in section 2. This is proved as follows. If we put

$$\xi=\xi''\frac{1}{1+\epsilon x}\ ,\quad \mu_1=\mu_1''\frac{1+\epsilon x}{\epsilon+x},\ \mu_i=\mu_i''\ ,\quad i\geq 2, \tag{48}$$

it is evident from (43) that ϕ would become a function of ξ'', μ_1'', μ_2'', $\cdots$ satisfying the relations

$$\xi''\frac{\partial\phi}{\partial\xi''}\left(=\xi\frac{\partial\phi}{\partial\xi}\right)=N\ ,\quad \mu_i''\frac{\partial\phi}{\partial\mu_i''}\left(=\mu_i\frac{\partial\phi}{\partial\mu_i}\right)=Nw,\quad i=1,2,\cdots\ . \tag{49}$$

It is also evident that ϕ is the same function of ξ'', μ_1'', μ_2'', $\cdots$ as ϕ' is of ξ', μ_1', μ_2', $\cdots$. Now (42) has only one^{7} set of solution in ξ' and μ_i'. Hence from (49) we infer that $\xi'=\xi''$, $\mu_i'=\mu_i''$. Thus

$$\xi=\xi'\frac{1}{1+\epsilon x}\ ,\quad \mu_1=\mu_1'\frac{1+\epsilon x}{\epsilon+x},\ \mu_i=\mu_i'\ ,\quad i\geq 2. \tag{50}$$

(41) and (50) give the parameters μ_i in terms of μ_i'. Inserting them into

(36) we obtain

$$F(r,T) = -NkT\frac{z}{\gamma}\Big[\log N + \Big(a+b-\frac{2\gamma}{z}\Big)(r\log r + w\log w) - \log\xi' - w\sum\log\mu_i' + r\log(1+\epsilon x) + w\log(\epsilon+x) - w\log\epsilon\Big]$$

$$= \frac{\gamma-1}{\gamma}F'(r,T) - NkT\frac{z}{\gamma}\Big[\Big(1-\frac{2}{z}\Big)(r\log r + w\log w) + r\log(1+\epsilon x) + w\log\frac{\epsilon+x}{\epsilon}\Big]. \quad (51)$$

If the original (unprimed) group is a pair of nearest neighbours, we have $\gamma=1$, and (51) reduces to the expression for the free energy in the approximation discussed in section 2:

$$F_0(r,T) = -zNkT\Big[\Big(1-\frac{2}{z}\Big)(r\log r + w\log w) + r\log(1+\epsilon x) + w\log\frac{\epsilon+x}{\epsilon}\Big]. \quad (52)$$

Inserting this back into (51) we get (47).

5. COMPARISON WITH BETHE'S METHOD

The so-called local grand partition function[1,5,8] in Bethe's approximations with long-distance order is identical in form with our funtion ϕ when all the "interior sites" in the group have the same parameter μ. For the case of equal concentrations for the two kinds of atoms, which is the case so far considered, this parameter has been put ~~e~~equal to unity by Bethe. Since the different terms of the local grand partition function stand for the probabilities of occurance of the corresponding local groups in the crystal, it is clear that Bethe's method with long-distance order is essentially ~~the~~ equivalent to our method plus the assumption that the free energy (35) has a minimum when

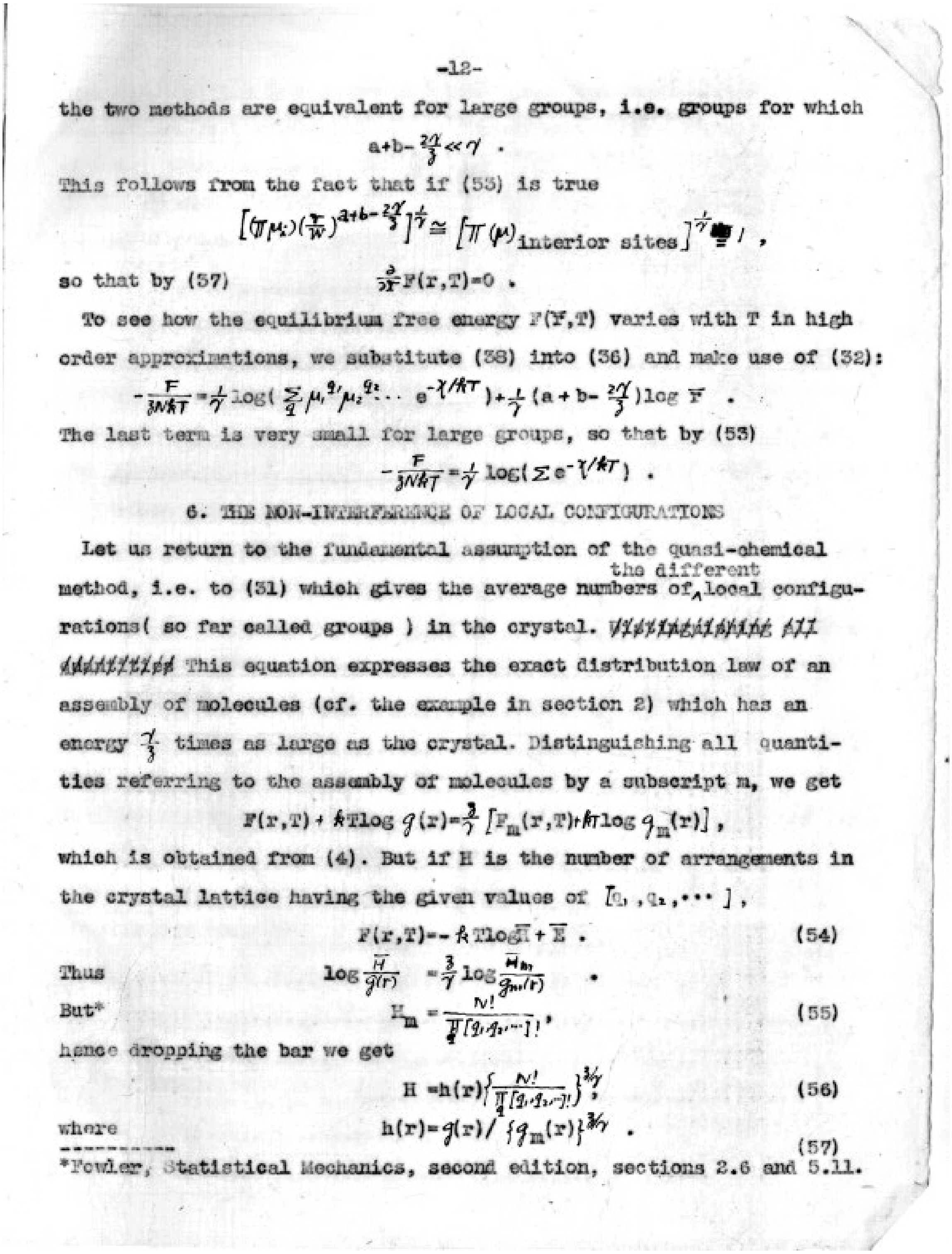
$$(\mu)_{\text{interior sites}} = 1. \quad (53)$$

But as we have shown that (38) gives the condition of a minimum of the free energy, the complete* identification of Bethe's and the quasi-chemical methods in any approximation reduces to the mathematical proof of the equivalence of (38) and (53). While this presents no difficulty at all for Bethe's first approximation (section7), a general proof is by no means easy. We can only satisfy ourselves with the assertion that

---

*"complete" as far as the probabilities of occurrance of the local configurations are concerned. The energy calculations are different in the two methods.

[8] Easthope, Proc. Camb. Phil. Soc. **33**, 502(1937).



-12-

the two methods are equivalent for large groups, i.e. groups for which

$$a+b-\frac{2\gamma}{3} \ll \gamma \;.$$

This follows from the fact that if (53) is true

$$\left[(\Pi\mu_i)\left(\frac{r}{W}\right)^{a+b-\frac{2\gamma}{3}}\right]^{\frac{1}{\gamma}} \cong \left[\Pi(\mu)_{\text{interior sites}}\right]^{\frac{1}{\gamma}} \; [illegible] \; 1 \;,$$

so that by (57) $\frac{\partial}{\partial r}F(r,T)=0$.

To see how the equilibrium free energy $F(\bar{r},T)$ varies with T in high order approximations, we substitute (38) into (36) and make use of (32):

$$-\frac{F}{3NkT}=\frac{1}{\gamma}\log\left(\sum_{q}\mu_1^{q_1}\mu_2^{q_2}\cdots e^{-\chi/kT}\right)+\frac{1}{\gamma}\left(a+b-\frac{2\gamma}{3}\right)\log\bar{r} \;.$$

The last term is very small for large groups, so that by (53)

$$-\frac{F}{3NkT}=\frac{1}{\gamma}\log\left(\sum e^{-\chi/kT}\right) \;.$$

6. THE NON-INTERFERENCE OF LOCAL CONFIGURATIONS

Let us return to the fundamental assumption of the quasi-chemical method, i.e. to (51) which gives the average numbers of the different local configurations(so far called groups) in the crystal. ~~[illegible]~~ This equation expresses the exact distribution law of an assembly of molecules (cf. the example in section 2) which has an energy $\frac{\gamma}{3}$ times as large as the crystal. Distinguishing all quantities referring to the assembly of molecules by a subscript m, we get

$$F(r,T)+kT\log g(r)=\frac{3}{\gamma}\left[F_m(r,T)+kT\log g_m(r)\right],$$

which is obtained from (4). But if H is the number of arrangements in the crystal lattice having the given values of $[q_1, q_2, \cdots]$,

$$F(r,T)=-kT\log\bar{H}+\bar{E} \;. \qquad (54)$$

Thus

$$\log\frac{\bar{H}}{g(r)} = \frac{3}{\gamma}\log\frac{\bar{H}_m}{g_m(r)} \;.$$

But*

$$H_m=\frac{N!}{\prod_q [q_1, q_2, \cdots]!}, \qquad (55)$$

hence dropping the bar we get

$$H = h(r)\left\{\frac{N!}{\prod_q [q_1, q_2, \cdots]!}\right\}^{3/\gamma}, \qquad (56)$$

where

$$h(r)=g(r)/\{g_m(r)\}^{3/\gamma} \;. \qquad (57)$$

---

*Fowler, Statistical Mechanics, second edition, sections 2.6 and 5.11.

Equation (56) has been referred to in Fowler and Guggenheim's paper[1] as the mathematical expression of the "hypothesis of the non-interference of local configurations", because when $\frac{\gamma}{\delta}=1$, the number of arrangements in the crystal consistent with the distribution law $[q_1, q_2, \cdots]$ for the groups of sites is, except for the factor h(r), equal to

$$H_m=\frac{N!}{\prod_q [q_1, q_2, \cdots]!},$$

which is the number of arrangements in the crystal for the given values of $[q_1, q_2, \cdots]$ if the N groups in the crystal are imagined to be separated and are filled independently with atoms. The term "non-interference" comes from the fact that actually the N groups are not separated but are interlocked and cannot be filled independently with atoms, i.e. they "interfere" with each other.

To find the value of $g_m(r)$ we notice that by definition $g_m=\Sigma H_m$. But ΣH_m is the number of arrangements in the N separated groups considered above if they are to be so filled with atoms that Nw of them have wrong atoms on the sites i, $i=1,2,\cdots$. Among the N sites i of the N groups $\frac{N!}{(Nr)!(Nw)!}$ different arrangements are possible. Hence*

$$g_m=\Sigma H_m=\Sigma\frac{N!}{\prod_q [q_1, q_2, \cdots]!}=\left[\frac{N!}{(Nr)!(Nw)!}\right]^{a+b}. \tag{58}$$

Thus

$$h(r)=\left[\frac{N!}{(Nr)!(Nw)!}\right]^{2(a+b)\frac{\delta}{\gamma}} \tag{59}$$

The free energy of the crystal may be obtained from (54), (56) and (59):

$$F(r,T)=\overline{E}-\frac{3NkT}{\gamma}\left\{(a+b-\frac{2\gamma}{\delta})(r\log r+w\log w)+\log N-\frac{1}{N}\sum_q\overline{[q_1, q_2, \cdots]}\log[q_1, q_2, \cdots]\right\}$$

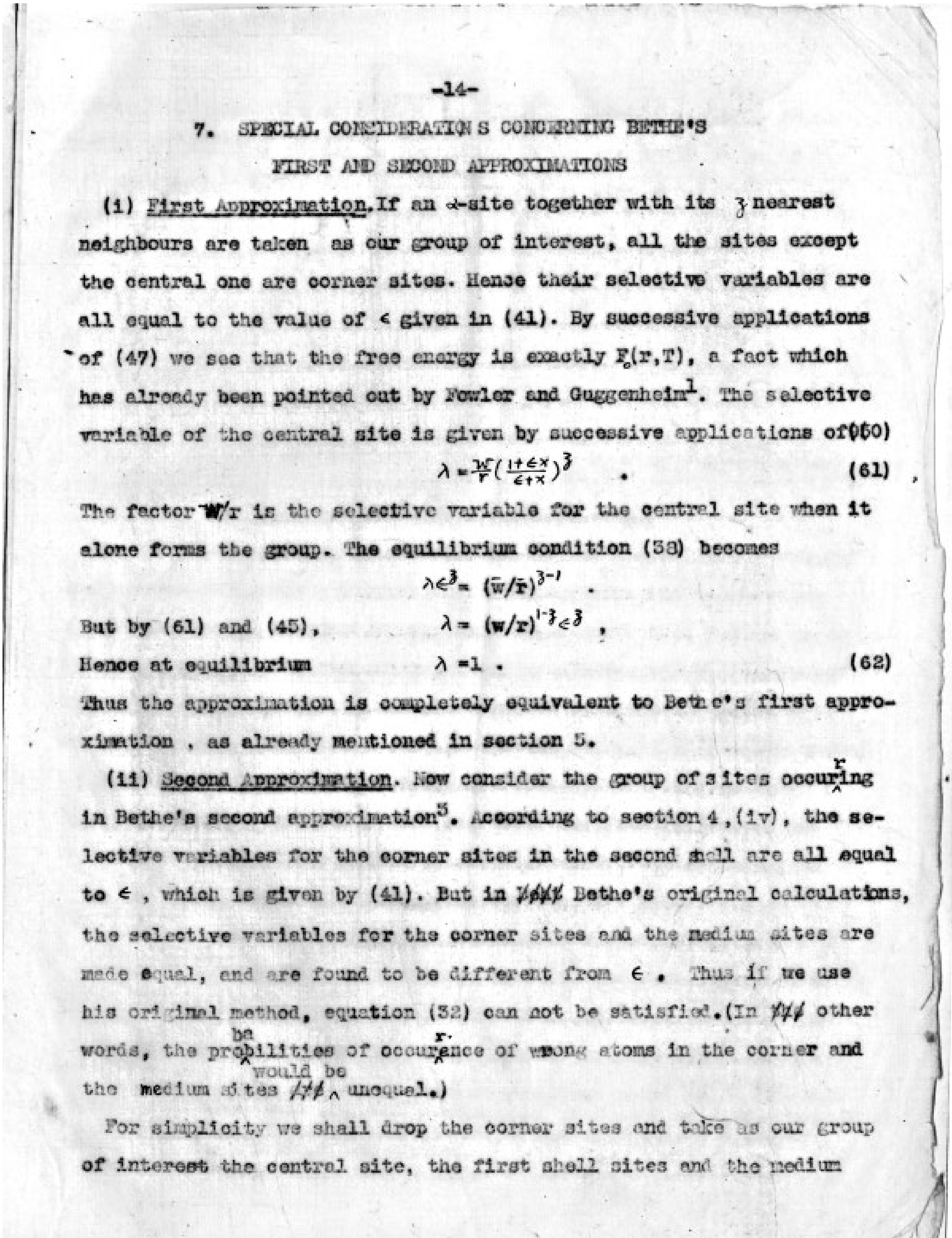
which has been obtained above by integration.

---

*It might be mentioned in passing that for the special case considered in section 2, (58) gives directly the value of the sum Σ_x in equation (8.5) of Fowler and Guggenheim's paper if their r and q are equal. The generalization to the case $r\neq q$ is however easy. The result is

$$\Sigma_x\frac{[3N]!}{[3N(r-x)]![3Nx]![3N(1-r-q+x)]![3N(q-x)]!}=\frac{[3N]!}{[3Nr]![3N(1-r)]!}\cdot\frac{[3N]!}{[3Nq]![3N(1-q)]!} \tag{60}$$

which is exact. The value of $\log\Sigma_x$ given by (60) reduces to the approximate expression that Fowler and Guggenheim obtained by identifying Σ_x with its maximum term when N is large.



-14-

7. SPECIAL CONSIDERATIONS CONCERNING BETHE'S FIRST AND SECOND APPROXIMATIONS

(i) First Approximation. If an α-site together with its z nearest neighbours are taken as our group of interest, all the sites except the central one are corner sites. Hence their selective variables are all equal to the value of ϵ given in (41). By successive applications of (47) we see that the free energy is exactly $F_0(r,T)$, a fact which has already been pointed out by Fowler and Guggenheim[1]. The selective variable of the central site is given by successive applications of (60)

$$\lambda = \frac{w}{r}\left(\frac{1+\epsilon x}{\epsilon+x}\right)^{z} \quad . \qquad (61)$$

The factor w/r is the selective variable for the central site when it alone forms the group. The equilibrium condition (38) becomes

$$\lambda\epsilon^{z} = (\bar{w}/\bar{r})^{z-1} \quad .$$

But by (61) and (45), $\lambda = (w/r)^{1-z}\epsilon^{z}$,

Hence at equilibrium $\lambda = 1$. (62)

Thus the approximation is completely equivalent to Bethe's first approximation , as already mentioned in section 5.

(ii) Second Approximation. Now consider the group of sites occurring in Bethe's second approximation[5]. According to section 4, (iv), the selective variables for the corner sites in the second shell are all equal to ϵ , which is given by (41). But in Bethe's original calculations, the selective variables for the corner sites and the medium sites are made equal, and are found to be different from ϵ . Thus if we use his original method, equation (32) can not be satisfied. (In other words, the probabilities of occurrence of wrong atoms in the corner and the medium sites would be unequal.)

For simplicity we shall drop the corner sites and take as our group of interest the central site, the first shell sites and the medium sites; with selective variables μ , ν and λ respectively.(The contribution ~~to the free energy of the~~ by the corner sites can be included in the free energy by simple addition as shown in section 4 (v).) With the notations n, m and g_{nm} of Bethe[3] we find

$$\phi = \xi \sum_n (x^n + \mu x^{z-n}) P_n(x,\nu,\lambda) , \tag{63}$$

where
$$P_n(x,\nu,\lambda) = \nu^n \sum_m g_{nm}[(1+\lambda)x]^m (x^2+\lambda)^{(\frac{z}{2}-1)n-\frac{m}{2}}(1+\lambda x)^{(\frac{z}{2}-1)(z-n)-\frac{m}{2}}.$$

After eliminating ξ and μ , (32) becomes

$$zw=\frac{2\sum_{l,n} x^{l-n}(rP_n\lambda\frac{\partial}{\partial\lambda}P_l + wP_l\lambda\frac{\partial}{\partial\lambda}P_n)}{(z-2)(\sum_n x^n P_n)(\sum_n x^{-n}P_n)}=\frac{\sum_{l,n} x^{l-n}(rP_n\nu\frac{\partial}{\partial\nu}P_l + wP_l\nu\frac{\partial}{\partial\nu}P_n)}{(\sum_n x^n P_n)(\sum_n x^{-n}P_n)}. \tag{64}$$

The free energy is obtained from (36):

$$F(r,T)=-\frac{NkT}{z-1}\Big[\frac{1}{2}(z^2-4z+4)(r\log r+w\log w)+r\log\sum_n x^n P_n+w\log\sum_n x^{z-n}P_n-zw\log\nu-wz(\tfrac{z}{2}-1)\log\lambda\Big]. \tag{65}$$

8. APPLICATION TO THE CRYSTAL Cu_3Au

For the face-centred crystal Cu_3Au, we may of course follow Peierls[9] and take as our group a central site together with its twelve first shell neighbours. The free energy expression would then contain seven selective variables*, four of which can be eliminated. The resultant expression is very cumbersome and numerical calculations would be laborious. We therefore make a simpler approximation: the group is taken to be four nearest neighbours forming a tetrahedron. A little geometrical consideration assures us that all such tetrahedrons contain an α-site (for gold atoms) and three β-sites (for copper atoms), an interesting conclusion showing that the tetrahedron might be regarded as a sort of "molecular" structure in a face-centred lattice with atomic ratio 1:3. Our approximation may thus be reasonably expected to reveal the more important features of the order-disorder transformation in such

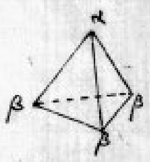


Fig.2

---

*For α-centred groups, three, and for β-centred, four, parameters are necessary. Both these two kinds of groups must be considered because otherwise the energy of the crystal cannot be easily obtained from the energy of the groups in the crystal.

[9] Peierls, Proc. Roy. Soc. London A154, 207(1936).

alloys.

Let μ and ν be the parameters (for wrong atoms) of the β-sites and the α-sites respectively. Let there be altogether 4N atoms. It is easy to see that there are 8N groups in the crystal. When Nw atoms on the α-sites are wrong, the equations determining the parameters are

$$8N\phi=\xi[x^3+3x^2\mu+3x^3\mu^2+x^6\mu^3+\nu(x^6+3x^3\mu+3x^2\mu^2+x^3\mu^3)], \tag{66a}$$

$$8Nw=\nu\frac{\partial\phi}{\partial\nu}=\xi\nu(x^6+3x^3\mu+3x^2\mu^2+x^3\mu^3), \tag{66b}$$

and $$8N(\tfrac{w}{3})+8N(\tfrac{w}{3})+8N(\tfrac{w}{3})=\mu\frac{\partial\phi}{\partial\mu}=3x^2\mu\xi[1+2x\mu+x^4\mu^2+\nu(x+2\mu+x\mu^2)], \tag{66c}$$

where x is defined by (9). The energy of the crystal is (cf. (34)),

$$E=\frac{1}{2}kT^2\frac{\partial\phi}{\partial T}+\text{constant} ; \tag{67}$$

so that the free energy becomes (cf. (35))

$$F(w,T)=-kT\left[\log g(w)+\tfrac{1}{2}(\phi-8N\log\xi-8Nw\log\nu-8Nw\log\mu)_{T=\infty}^{T}\right] .$$

But $$\log g(w)=-N\left\{(1-w)\log(1-w)+w\log w+w\log\frac{w}{3}+(3-w)\log[(3-w)/3]\right\} .$$

and at $T=\infty$ $$\nu=\frac{w}{1-w} , \qquad \mu=\frac{w}{3-w} , \qquad \xi=8N(1-w)(1-\tfrac{w}{3})^3 .$$

Hence $$-\frac{F(w,T)}{NkT}=-9\log 3+4\log 8N+6w\log w+3(1-w)\log(1-w)+3(3-w)\log(3-w)$$
$$-4\log\xi-4w\log\mu-4w\log\nu . \tag{68}$$

Since ξ and ν can be very easily solved from (66), numerical calculations are quite simple. The equilibrium value of w is given by (cf. (37) and(38))

$$0=3\log\frac{(1-\bar{w})(3-\bar{w})}{\bar{w}^2}+4\log\mu\nu . \tag{69}$$

This is always satisfied at $\bar{w}=\frac{3}{4}$.* Actual calculation shows that the absolute minimum of the free energy is or is not at $\bar{w}=\frac{3}{4}$ according as $x\geq .2965$ or $x<.2965$. The values of the free energy is plotted in Fig.3. From the form of the graph it is seen that the crystal has a critical temperature at which the long-distance order and(hence)the energy are discontinuous. The critical temperature T_c and the latent heat Q are found to be

---

*This is not evident from (69) directly. But if we divide the whole crystal into four sublattices which are all simple cubic and introduce a w for each sublattice so that Nw_i is the number of A atoms on the i-th sublattice (i=1,2,3,4), it is obvious that the free energy is symmetrical in the w's. From this we infer that (69) is satisfied at $\bar{w}=\frac{3}{4}$.

$$T_c = .8228\,\frac{1}{k}\left[\tfrac{1}{2}(V_{AA}+V_{BB})-V_{AB}\right] \quad , \quad Q = .8824N\left[\tfrac{1}{2}(V_{AA}+V_{BB})-V_{AB}\right] \ .$$

In terms of the total energy change from $T=0$ to $T=\infty$:

$$E_0 = 3N\left[\tfrac{1}{2}(V_{AA}+V_{BB})-V_{AB}\right]$$

these quantities become

$T_c = 1.097E_0/R$, ($T_c = 2.19E_0/R$ in Bragg-Williams' approximation and
$T_c \cong 1.3E_0/R$ in Peierls' approximation.)

$Q = .2941E_0$, ($Q = .218E_0$ in Bragg-Williams' approximation and
$Q \cong .36E_0$ in Peierls' approximation.)

where R stands for $4Nk$.

It will be noticed that due to the lack of a free energy expression Peierls[9] did not give the exact values of these quantities.

Figure

In conclusion, the author wishes to express his thanks to Prof. J. S. Wang for valuable criticism ~~discuss~~ and advice.

Fig.3