PHYSICAL REVIEW VOLUME 76, NUMBER 7 OCTOBER 1, 1949

Quasi-Chemical Method in the Statistical Theory of Regular Mixtures

YIN-YUAN LI*
Department of Physics, Carnegie Institute of Technology, Pittsburgh, Pennsylvania

(Received May 23, 1949)

The quasi-chemical equations in the statistics of "regular" mixtures are deduced by assuming non-interference among local configurations. This method is capable of yielding higher and higher approximations by choosing a larger group of lattice sites as the local configuration under consideration. The comparative accuracy of different approximations can be judged by a simple criterion. Further applications to ferromagnetism and to the order-disorder transition in alloys, and their results are discussed. Equations for ternary mixtures or mixtures of even more components are given. The asymmetry of solubility of one solute in two immiscible solvents and the order effect of Ag_2HgI_4 serve as the interesting examples of the ternary case.

1. INTRODUCTION

A MIXTURE containing atoms (or molecules) of two or more kinds will be called a regular one if it satisfies the following conditions: (a) All the atoms are located on a set of lattice points (or sites). Each site has the same number of nearest neighbors. (b) The kinetic energy is either separable from the potential energy, or else of no importance. And (c) the fields of force of the atoms are spherically symmetric; and the interaction energy of the whole system can be expressed as the sum of closest neighbor contributions. The latter condition implies that the interaction fields of different atoms are almost equal at distances greater than that between two nearest neighbors. The typical examples of this model are of course substitutional alloys. It may also apply to the so-called "regular" liquid mixtures (of the hypothetical scheme adopted by Hildebrand[1] and later by Guggenheim[2] and others) and liquids considered from the viewpoint of the "hole" scheme[3] in which the unoccupied sites (holes) act as atoms of one kind. Other examples are ferromagnetism, solutions of hydrogen in palladium, and monolayer (localized) surface adsorptions. The existence of unoccupied sites leaves the theory under consideration perfectly unchanged as far as condition (a) is concerned.

A statistical method proposed by Alfred and Mark[4] for binary regular mixtures may be briefly stated as follows. The configuration of the system is described by specifying the occupants of the nearest pairs whose constituent sites taken together constitute the lattice, rather than by specifying the occupants of each individual lattice site. Assuming that the configurational energy of the binary system is the sum of the interaction energy of these nearest pairs multiplied by the ratio of the total number of pairs in the binary system to that of the reference set of pairs, they obtain the combinatory factor of the partition function of the binary system by considering the permutation of the reference pairs. Their result leads to an equation similar to that of a quasi-chemical equilibrium introduced first by Guggenheim to the theory of regular solution.

Recently Miller[5] gave a critical analysis of the Alfred-Mark method. By clarifying the meaning of their assumption, he has shown that it amounts to neglecting the restriction on the allocation of the nearest pairs of different kinds (namely, *A*-*A*, *B*-*B*, and *A*-*B* when the atoms of different kinds are denoted as *A*'s and *B*'s). This restriction is basically inherent in the arrangements of *A* and *B* atoms on the lattice sites. In other words, Alfred and Mark have unknowingly adopted the so-called assumption of non-interference of local configurations. The term "non-interference," which was first used by Fowler and Guggenheim, comes from the fact that actually the closest pairs are interlocked, i.e., they "interfere" with each other. Fowler and Guggenheim[6] proved even earlier (1940) that their quasi-chemical equation can be deduced by using this assumption.

By using the same assumption, the present author will develop a statistical method capable of yielding successively higher approximations and accessible to regular mixtures containing more than two different kinds of atoms. His procedure can be applied with a slight complication to the problem of superstructure of alloys. The results turn out to be the set of equations first obtained by Yang and recently generalized by Li on the general quasi-chemical theory of superstructure.[7] We shall notice that the "pair" approximation (Fowler-Guggenheim, Alfred-Mark) reaches qualitatively incorrect results for lattices with nearest interactions among the first shell sites and larger local groups offer solutions. Therefore, the general theory for successive approxima-

---

* Now at the University of Illinois, Urbana, Illinois.

[1] J. Hildebrand, J. Am. Chem. Soc. **51**, 66 (1929).

[2] E. A. Guggenheim, Proc. Roy. Soc. **A148**, 304 (1935); G. S. Rushbrooke, Proc. Roy. Soc. **A166**, 296 (1938): Also R. H. Fowler and E. A. Guggenheim, *Statistical Thermodynamics* (Cambridge University Press, London, 1939), pp. 351–366.

[3] J. Frenkel, Trans. Faraday Soc. **33**, 58 (1937). Bresler, Acta Physicochim. U.R.S.S. **10**, 491 (1939).

[4] T. Alfred and H. Mark, J. Chem. Phys. **10**, 303 (1942). This article will be cited as TA-HM.

[5] A. R. Miller, J. Chem. Phys. **15**, 513 (1947).

[6] R. H. Fowler and E. A. Guggenheim, Proc. Roy. Soc. **A174**, 189 (1940).

[7] C. N. Yang, J. Chem. Phys. **13**, 66 (1945). C. N. Yang and Y. Y. Li, J. Chinese Phys. **7**, 59 (1947). Y. Y. Li, J. Chem. Phys. **17**, 447 (1949). These articles will be cited as Ref. I, II, and III, respectively.

tions is not only useful but very necessary if this statistical method should be good also for more tightly packed lattices.

2. STATISTICAL METHOD

Let us consider a group of sites in a lattice which is occupied by the atoms of a binary system. The whole system is conceived as a superposition of such groups, all of identical form. Each group consists of a (connected) set of sites, but two different groups may have sites in common. Alfred and Mark selected the simplest and smallest groups. In addition they adopted a very particular manner of taking them by cutting the lattice of N sites into $N/2$ nearest pairs in such a way that no two pairs have a site in common. We intentionally generalize the choice of local groups to any shape and size, and also generalize the manner of taking them.

We denote[8] the total number of atoms in the binary mixture by N; $N\theta$ of them being A atoms and $N(1-\theta)$ being B atoms. N is also the total number of lattice sites. Let z be the number of nearest neighbors of each site. In a given choice of the groups let there be n sites and p nearest pairs in each group, and let the total number of groups be M. The sites in each group will be labeled by 1, 2, $\cdots$, n, in one and the same way. We shall let $q_i (i=1, 2, \cdots, n)$ specify the occupant of the ith site; $q_i=1$ if the ith site is occupied by an A atom, $q_i=0$ if it is occupied by a B atom. In a given distribution of atoms in the whole lattice let $[q_1, q_2, \cdots, q_n]M$ be the number of groups having q_iA atoms in the ith site $(i=1, 2, \cdots, n)$. Since the total number of sample groups is M, we have

$$\sum_q [q_1, q_2, \cdots, q_n]=1, \tag{2.1}$$

where the summation is taken over all possible values of the q's. When the distribution of A and B atoms on the whole lattice is purely random, the probability of finding an A atom on the ith site of any one of the M groups is simply θ. Consequently, we have

$$\sum_q q_i[q_1, q_2, \cdots, q_n]=\theta, \tag{2.2}$$

giving n equations (with $i=1, 2, \cdots, n$). The number of ways of arranging the indistinguishable groups is

$$\sum_{[q_1, \cdots, q_n]} \frac{M!}{\prod_q \{M[q_1, \cdots, q_n]\}!},$$

with the summation $\sum_{[q_1, \cdots, q_n]}$ taken over all possible values of $[q_1, q_2, \cdots, q_n]$'s which are subjected to the limitation (2.1) and (2.2).

If there were no interlocking among the local configurations (so far called groups), these M groups together should duplicate the binary system by a factor $(Nz/2Mp)$, in which $Nz/2$ and pM are, respectively, the total number of closest pairs in the binary system and that in the M groups. Therefore, the assumption of non-interference permits us to write the configurational part of the partition function of the binary mixture as

$$\begin{aligned} P(N, \theta, T) &= \sum_{[q_1, \cdots, q_n]} g(N, \theta)\left\{\frac{M!}{\prod_q (M[q_1, \cdots, q_n])!}\right\}^{Nz/2Mp} \\ &\quad\times\exp\left\{-\frac{Nz}{2Mp}\sum_q M[q_1, \cdots, q_n] \times\chi(q_1, \cdots, q_n)/kT\right\} \\ &= \sum_{[q_1, \cdots, q_n]} g(N, \theta)\frac{\left(\dfrac{Nz}{2p}\right)!}{\prod_q \left(\dfrac{Nz}{2p}[q_1, \cdots, q_n]\right)!} \\ &\quad\times\exp\left\{-\frac{Nz}{2p}\sum_q [q_1, \cdots, q_n] \times\chi(q_1, \cdots, q_n)/kT\right\} \end{aligned} \tag{2.3}$$

where χ is the configurational energy which depends on the atomic distribution $(q_1, \cdots, q_n)$. Stirling's approximation

$$\left\{\frac{(r+s)!}{r!s!}\right\}^{\alpha}=\frac{\{\alpha(r+s)\}!}{(\alpha r)!(\alpha s)!} \tag{2.4}$$

for large number r and s has been used. The factor $g(N, \theta)$ is to be adjusted so that the total number of configurations of the system is given by

$$\frac{N!}{(N\theta)!(N[1-\theta])!}=P(N, \theta, \infty). \tag{2.5}$$

It is not a surprise that $g(N, \theta)$ cannot be dispensed with, for the assumption of non-interference of local configuration should not be perfectly correct.

The most probable (equilibrium) value of the fraction $[q_1, \cdots, q_n]$ can be obtained by maximizing the general term of (2.3) under the condition (2.1) and (2.2). Using Stirling's approximate formula and undetermined Lagrange's multipliers, we get

$$[q_1, q_2, \cdots, q_n]_{eq} = \frac{\mu_1^{q_1}\mu_2^{q_2}\cdots\mu_n^{q_n}\exp[-\chi(q_1, q_2, \cdots, q_n)/kT]}{\sum_q \mu_1^{q_1}\mu_2^{q_2}\cdots\mu_n^{q_n}\exp[-\chi(q_1, q_2, \cdots, q_n)/kT]} \tag{2.6}$$

---

[8] In this section we adopt almost the same notations as were used in Ref. II.

with $\mu_1, \cdots, \mu_n$ to be determined by (2.2). By putting

$$\varphi=\sum_q \mu_1^{q_1}\mu_2^{q_2}\cdots\mu_n^{q_n}\exp[-\chi(q_1, q_2, \cdots, q_n)/kT],$$

(2.2) and (2.4) may be written as

$$\mu_i\frac{\partial\ln\varphi}{\partial\mu_i}=\theta, \tag{2.2'}$$

and

$$[q_1, q_2, \cdots, q_n]_{eq}=\frac{1}{\varphi}\mu_1^{q_1}\mu_2^{q_2}\cdots\mu_n^{q_n} \times\exp[-\chi(q_1, q_2, \cdots, q_n)/kT]. \tag{2.4'}$$

Yang (Ref. II) has pointed out that a proof of the uniqueness of the solution for the μ's of an equation of the type (2.2′) follows easily from Lemma 2.42 in Fowler's *Statistical Mechanics*, second edition.

(2.4′) must reduce to the Alfred-Mark result when the groups are nearest pairs as taken by them. With $\mu_1=\mu_2$ given by (2.2′) for this case we have the number of *A-B*, *A-A*, and *B-B* pairs, respectively,

$$X_{AB}=\frac{Nz}{2}\{[0, 1]_{eq}+[1, 0]_{eq}\} =\frac{Nz}{\varphi}\mu_1\exp\{-V_{AB}/kT\}, \tag{2.7a}$$

$$X_{AA}=\frac{Nz}{2}[1, 1]_{eq}=\frac{Nz}{2\varphi}\mu_1^2\exp\{-V_{AA}/kT\}, \tag{2.7b}$$

$$X_{BB}=\frac{Nz}{2}[0, 0]_{eq}=\frac{Nz}{2\varphi}\exp\{-V_{BB}/kT\}, \tag{2.7c}$$

and so

$$X_{AB}^2=4X_{AA}X_{BB}\exp\{-2V/kT\} =(N_A-X_{AB})(N_B-X_{AB})\exp\{-2V/kT\}, \tag{2.8}$$

where V_{AB}, V_{AA}, and V_{AB} are, respectively, the interaction energy between *A-B*, *A-A*, and *B-B* nearest atoms; and $V=\frac{1}{2}(V_{AA}+V_{BB}-2V_{AB})$ is the "energy of mixing," or $2V$ is the amount of energy gain by creating an *A-A* pair and a *B-B* pair at the expense of two *A-B* pairs. (2.8) is seen to be the same as the quasi-chemical equation used by Guggenheim, but rather is different from Alfred-Mark result. The latter has the exponential $(-2Vz/kT)$ instead of simply $(-2V/kT)$. However, this extra factor z is erroneous due to a minor mistake in the procedure of formulating their statistical method. Unfortunately, Miller did not pay much attention to this difference between Eq. (11) of TA-HM and Guggenheim's quasi-chemical equation but just omitted the extra factor z in writing Eq. (14) of his article.

The present method is capable of yielding different approximations by choosing different groups. Consequently, we need a criterion to examine which group will give the better result. We may consider the factor $g(N, \theta)$ as representing to some degree the accuracy of the approximation. $g(N, \theta)$ should be equal to one, if the assumption on non-interference of local groups were

TABLE I. Order-disorder transitions of the second kind calculated with the quasi-chemical equations. E_0 = the energy of transformation from perfect order to complete randomness, T_c = the critical temperature of the order-disorder transition, and ΔC_v = the discontinuity of molar specific heat during the transition. The numerical values of the quantities given in the Table are for the atomic ratio $\theta=\frac{1}{2}$. The range of θ gives the limits within which the superstructure may be formed. The "pair" approximation and "center and first shell" give the same results as Bethe's method. Local configurations are named as square, rhombus, or tetrahedron, when the lattice points and the nearest interactions included in the local groups consist of, respectively, the vertices and the sides (or edges) of a square, etc.

Lattice	Quadratic $z=4$			Simple cubic $z=6$				Body-centered cubic $z=8$			
E_0	NV			$3/2\ NV$				$2\ NV$			
Local group	Pair	Center and first shell	Square	Pair	Center and first shell	Square	Unicell	Pair	Center and first shell	Rhombus	Second shell incomplete[a]
$zn/2p$	4	5/2	2	6	7/2	3	3/2	8	9/2	4	15/8
RT_c/E_0	1.443	Same as "pair"	1.386	1.644	Same as "pair"	1.624	1.613	1.738	Same as "pair"	1.736	1.698
$\Delta C_v/R$	1.93		1.99	1.78		1.80	1.83	1.70		1.70	1.80
Superstructure range of θ	0.25 −0.75		0.2696 −0.7304	0.1667 −0.8333		0.1705 −0.8295	0.1762 −0.8238	0.1250 −0.8750		0.1265 −0.8735	0.1360 −0.8640
Source of the above data	Ref. I	Unpublished elsewhere	Ref. I	Ref. I	Unpublished elsewhere	Unpublished elsewhere	Unpublished elsewhere	Ref. I	Unpublished elsewhere	Unpublished elsewhere	Unpublished elsewhere

[a] See Fig. 1.

TABLE II. Order-disorder transitions of the first kind calculated by the quasi-chemical equations. Q is the latent heat of the order-disorder transition. The numerical values of Q and RT_c/E_0 given above are for the stoichiometric atomic ratio $\theta=\frac{1}{4}$ and $\frac{1}{2}$, respectively, for the two different lattices.

Lattice	Face-centered cubic $\alpha\beta_3$-type			Face-centered tetragonal[a] $\alpha\beta$-type		
E_0	$3\,NV/4$			NV		
Local group	Pair	Rhombus	Tetrahedron	Pair	Rhombus	Tetrahedron
$zn/2p$	12	6	4	12	6	4
RT_c/E_0	No superstructure can be shown by this approximation	Same as "pair" approximation	1.097	No superstructure can be shown by this approximation	Same as "pair" approximation	0.7306
Q			$0.294E_0$			$0.288E_0$
Superstructure range of θ			0.17–0.34 and 0.66–0.83			0.43–0.57
Source of the above data	Unpublished elsewhere	Unpublished elsewhere	Ref. I, II	Unpublished elsewhere	Unpublished elsewhere	Ref. III

[a] Only a slight deviation from cubic form is treated as a cubic lattice in its energy interactions among atoms.

correct. From (2.5) we obtain

$$\ln g(N,\theta)=-N\theta\ln\theta-N(1-\theta)\ln(1-\theta)$$
$$+\frac{Nz}{2p}\sum_q[q_1,\cdots,q_n]_{eq}(T=\infty)$$
$$\times\ln[q_1,\cdots,q_n]_{eq}(T=\infty)$$
$$=N\left(\frac{nz}{2p}-1\right)[\theta\ln\theta+(1-\theta)\ln(1-\theta)].\quad(2.9)$$

In the derivation of the foregoing expression we have replaced as usual the summation of (2.3) by its maximum term and used (2.2), (2.4′) and the solution of (2.2′) at $T=\infty$, i.e.,

$$(\mu_i)_{T=\infty}=\theta/1-\theta,\quad(2.10)$$

and

$$(\varphi)_{T=\infty}=1/(1-\theta)^n.\quad(2.11)$$

Now we see that $g(N,\theta)=1$ only when $nz/2p=1$, or $p=np/2$, i.e., only when the local group is comparable in size to the binary system. In practical applications of the present theory, the use of a large group introduces troublesome calculations; therefore, we have to work with groups having $nz/2p$ rather larger than 1. The criterion which offers a guide in choosing the group, may be stated as follows: *A given approximation yields a result that is not less accurate than another if the former has a smaller* $nz/2p$. It is quite interesting to notice the following consequences: (a) A group containing a given number of sites may yield a result less accurate than another group with *less* sites, since the ration n/p acts as the gauge of accuracy not n itself. (b) A new group obtained by adding a corner site[9] to another group, with $p>n$, yields a less accurate result, since

$$\frac{z(n+1)}{2(p+1)}-\frac{zn}{2p}=\frac{z(p-n)}{2p(p+1)}.$$

By a corner site is meant a site having only one nearest neighbor in the group.

Finally we have the free energy function

$$F(N,\theta,T)=-kT\ln P$$
$$=-\frac{Nz}{2p}kT\left\{\ln\varphi-\sum_i\theta\ln\mu_i+\left(n-\frac{2p}{z}\right)\right.$$
$$\left.\times[\theta\ln\theta+(1-\theta)\ln(1-\theta)]\right\}.\quad(2.12)$$

3. THEORY OF SUPERSTRUCTURE

It is commonly observed phenomena that many substitutional alloys[10] exhibit order at not too high tem-

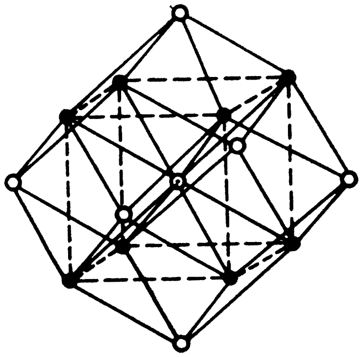

FIG. 1. A local group taken from a body-centered lattice containing a center site, denoted by ⊙ in the figure, its first shell sites, ●'s, and 6 sites, ○'s, from the second shell. Solid lines stand for the nearest interactions and dotted lines for the lattice frame.

---

[9] For a special consideration on corner sites see Theorem (iv) of Ref. I.

[10] See the review article by F. C. Nix and W. Shockley, Rev. Mod. Phys. **10**, 1 (1938).

peratures. We shall consider a binary system and designate the atoms by A and B. We may regard that the lattice sites as consisting of different sets of sublattices. Under suitable conditions the A atoms and B atoms distribute themselves unevenly on the different sublattices. In case of highly ordered states, some of the sublattices have their sites completely occupied by A atoms and the others by B atoms. The degree of the so-called long-range order can be described by giving the probability of finding A (or B) atoms on the sites of sublattices. The division of lattice sites into suitable sublattices must be made in such a manner that the resulting free energy successfully describes the system. For simpler cases only two sublattices (so called α- and β-sites) are needed. Here we leave the number of sublattices, m, unspecified.

Let[11] the number of sites on the sublattices L_h be $Nr_h (h=1, 2, \cdots, m)$. In a given distribution of atoms among the lattice sites, the number of A atoms on L_h is denoted by $Nr_h\theta_h$. The M groups may now fall into l different types with Mc_λ groups of the type $\lambda (\lambda=1, \cdots, l)$, because it may happen that the ith site is on L_1 in one group but on L_2 in another group. In the place of (2.1) and (2.2) we now get the limitation on the distribution of atoms in the M groups at given θ_h's in the form

$$\sum_q [q_1, \cdots, q_n]_\lambda = c_\lambda, \tag{3.1}$$

$$\sum_q q_i [q_1, \cdots, q_n]_\lambda = c_\lambda \theta_{\lambda i}. \tag{3.2}$$

Where $M[q_1, \cdots, q_n]_\lambda$ is, of course, the number of groups of type λ with q_i A atom in the ith site. $\theta_{\lambda i} = \theta_h$ when the ith site of the group of type λ is on L_h. Further, we have

$$\sum_h r_h \theta_h = \theta. \tag{3.3}$$

By the same procedure used in the preceding section, we get

$$\begin{aligned} P(N, \theta_1, \cdots, \theta_m, T) &= \sum_{[q_1, \cdots, q_n]_\lambda} g(N, \theta_1, \cdots, \theta_m) \\ &\quad \times \frac{\dfrac{Nz}{2p}}{\prod_{q,\lambda} \left(\dfrac{Nz}{2p} [q_1, \cdots, q_n]_\lambda \right)!} \\ &\quad \times \exp\left\{ -\frac{Nz}{2p} \sum_{q,\lambda} [q_1, \cdots, q_n]_\lambda \right. \\ &\quad \left. \times \chi(q_1, \cdots, q_n)/kT \right\}, \end{aligned} \tag{3.4}$$

$$\begin{aligned} [q_1, \cdots, q_n]_{\lambda(eq)} &= \frac{c_\lambda}{\varphi_\lambda} \mu_{\lambda 1}{}^{q_1} \cdots \mu_{\lambda n}{}^{q_n} \\ &\quad \times \exp\{-\chi(q_1, \cdots, q_n)/kT\}, \end{aligned} \tag{3.5}$$

where

$$\varphi_\lambda = \sum_q \mu_{\lambda 1}{}^{q_1} \cdots \mu_{\lambda n}{}^{q_n} \exp\{-\chi(q_1, \cdots, q_n)/kT\}, \tag{3.6}$$

with μ's to be determined by

$$\mu_{\lambda i} \frac{\partial \ln \varphi_\lambda}{\partial \mu_{\lambda i}} = \theta_{\lambda i}.$$

Further, we have

$$g(N, \theta_1, \cdots, \theta_m) \left\{ \frac{\left(\dfrac{Nz}{2p} \right).}{\prod_{q,\lambda} \left(\dfrac{Nz}{2p} [q_1, \cdots, q_n]_{\lambda(eq)}! \right)} \right\}_{T=\infty} = \frac{N!}{\prod_h (N\theta_h)!(N[1-\theta_h])!},$$

or

$$\begin{aligned} \ln g(N, \theta_1, \cdots, \theta_m) &= N \left(\frac{nz}{2p} - 1 \right) \sum_h r_h \\ &\quad \times [\theta_h \ln\theta_h + (1-\theta_h) \ln(1-\theta_h)]. \end{aligned} \tag{3.7}$$

The derivation of the foregoing is quite similar to (2.9) but one more relation

$$\sum_{\lambda, i} c_\lambda f(\theta_{\lambda i}) = n \sum_h r_h f(\theta_h),$$

has been used. Finally, we replace the summation of (3.4) by its maximum term as usual and write the configurational free energy function

$$\begin{aligned} F(N, \theta_1, \cdots, \theta_m, T) &= -kT \ln P \\ &= -\frac{Nz}{2p} kT \left\{ \sum_\lambda \ln\varphi_\lambda - \sum_{\lambda, i} \theta_{\lambda i} \ln\mu_{\lambda i} \right. \\ &\quad + \left(n - \frac{2p}{z} \right) \sum_h r_h [\theta_h \ln\theta_h \\ &\quad \left. + (1-\theta_h) \ln(1-\theta_h)] \right\}. \end{aligned} \tag{3.9}$$

We see that our Eqs. (3.2′), (3.6), and (3.9) coincide with those of Ref. III. The reader is referred to Ref. III for further work of finding equilibrium long-range order and allied physical quantities.

---

[11] In this section we adopt almost the same notation as was used in Ref. III.

In the opinion of this author the argument of the present theory is obviously more straightforward than that of Yang's theory which employed a chemical analogy and a mathematical trick of solving an integration by a Legendre transformation. Nevertheless, the fundamental assumptions of these two methods are only superficially different. By drawing an analogy between the local groups and the molecules in an ideal gaseous system, Yang implicitly followed the exact line of the hypothesis of non-interference of local configurations. He also tried to deduce his quasi-chemical method by adopting the hypothesis although with little success (§6 of Ref. I).

Although the statistical theory presented in this paper does not obtain new equations it leads to a useful criterion on the successive approximations. Because the factor $g(N, \theta_h)$ is hidden behind the scenes in Yang's deduction, no criterion has been established in Ref. I–III. In order to show the effectiveness of the criterion, we collect the calculated results of different approximations for several different kinds of lattices in Tables I and II. (See also Fig. 1.) For the procedure of calculations involved our readers are referred to Ref. I–III. In order to conserve space no details can be given here.

Those with the order-disorder *transition of the second kind* are shown in Table I which exhibits that in all cases T_c and $nz/2p$ vary in the same direction and ΔC_v in the opposite direction. Therefore, we may expect that higher approximations would yield larger ΔC_v in better agreement with experimental data. In Table II the simplest group which shows the superstructure phenomenon of a face-centered lattice is the tetrahedron. Both the "pair" approximation and "rhombus" are absolutely uncapable of describing the phenomena. It is interesting to see that these two local groups do not include in themselves the characteristic feature of a face-centered lattice that there are nearest neighbors among the first-shell sites of any site. On the other hand, the simple cubic or the body-centered cubic lattice has no nearest interactions among the nearest neighbors of any site. This difference in their lattice structure might be interpreted as the reason why the order-disorder transition of a face-centered cubic lattice is a transition of the first kind[12] (a sudden vanishing of long-range order with the appearance of latent heat), while that of a simple or a body-centered cubic lattice is a transition of the second kind (continuous change to disorder without latent heat). Calculations with different methods (Bethe, Kirkwood, Fowler-Guggenheim, and Yang-Li) have shown that an order-disorder transition is of the second kind whenever there is no nearest interactions among the nearest neighbors of any lattice site. It would be interesting if we could prove *in general* that *the order-disorder transition of a superlattice with nearest interactions among the nearest neighbors of a site occurs as one of the first kind.* It roughly means that the order-disorder transition happens much more abruptly than would be the case if there were no nearest interactions among the first shell sites. Work has been done only for a particular case, Cu-Au alloy system (Table II or Ref. III). Another important example is the Mg-Cd alloy system (close-packed hexagonal structure).

Incidentally, we may mention that the usual classification of the superstructures into those of $\alpha\beta$-type (designated as "with equivalent sublattices" by Fowler and Guggenheim[13]) and those of $\alpha\beta_3$-type ("with non-equivalent sublattices") is unessential and leads one to the incorrect conclusion that order-disorder transition of that with equivalent sublattices occurs as one of the second kind for any crystal structure. We have shown the incorrectness of the latter conclusion (Ref. III).

4. FERROMAGNETISM IN THE ISING MODEL

As remarked a few years ago by Peierls,[14] the theory of ferromagnetism can be presented in parallel with that of the order-disorder of alloys and this parallel is even closer when the Ising model[15] is assumed. In this model the spin of $\frac{1}{2}$ per atom is oriented either parallel or antiparallel to a specified direction. It is equivalent to disregarding the quantum mechanical combination of two spins $\sigma=\sigma_1+\sigma_2=1$ and $\sigma_z=0$, with the z axis being the specified direction. Now we can think of two kinds of atoms, A with a spin of $+\frac{1}{2}$ and B with a spin of $-\frac{1}{2}$, although they are actually of the same chemical element. The potential interaction of a nearest pair is taken as

$$\begin{aligned} V&=-2JS_1S_2-2(S_1+S_2)\beta H,\\ V_{AA}&=-\tfrac{1}{2}J-2\beta H,\\ V_{BB}&=-\tfrac{1}{2}J+2\beta H,\\ V_{AB}&=\tfrac{1}{2}J, \end{aligned} \tag{4.1}$$

where J is the exchange integral,[16] β the Bohr magneton $e\hbar/2mc$, and H the external field. Evidently we may shift the zero level of energy by replacing V_{AA}, V_{BB}, and V_{AB}, respectively, with $-J-2\beta H$, $-J+2\beta H$, and 0 in order to make the evaluation simpler.

$(2\theta-1)N\beta/v$ gives the magnetization M with v as the volume of the domain under consideration. The atomic ratio θ now varies with the temperature instead of remaining constant as it does in chemical mixtures. Elsewhere nothing is different from that discussed in Section 2. Therefore, the equilibrium value of θ can be determined by

$$\partial F/\partial\theta=0 \quad \partial^2F/\partial\theta^2>0, \tag{4.2}$$

---

[12] In a transition of the nth kind, the free energy function of the two phases have the same derivatives up to the $(n-1)$th order, whereas the nth derivatives are different. The zeroth derivative of a function is, of course, the function itself.

[13] R. H. Fowler and E. A. Guggenheim, *Statistical Thermodynamics* (Cambridge University Press, London, 1939), pp. 566, 598.

[14] R. Peierls, Proc. Roy. Soc. **A154**, 207 (1936).

[15] E. Ising, Zeits. f. Physik **31**, 253 (1925).

[16] See J. H. Van Vleck, *Electric and Magnetic Susceptibility* (Oxford University Press, Teddington, 1932), Chapter XII.

where F is given by (2.12). We find directly that

$$\sum_i \ln\mu_i=\left(n-\frac{2p}{z}\right)\ln\frac{\theta}{1-\theta}, \tag{4.3a}$$

and

$$\sum_i \frac{\partial\ln\mu_i}{\partial\theta}>\left(n-\frac{2p}{z}\right)\frac{1}{\theta(1-\theta)}, \tag{4.3b}$$

in the place of (4.2).

When the external field H does not exist, $\theta=\frac{1}{2}$ (and so $M=0$) offers a solution of (4.3a) for any temperature and this is the only solution for $T=\infty$. The mathematical proof of this statement can be easily obtained from the following relations, which are derived from (2.2′).

$$\mu_i(\theta)\mu_i(1-\theta)=1 \quad i=1, 2, \cdots, n, \tag{4.4}$$

and

$$\mu_i(T=\infty)=\frac{\theta}{1-\theta}.$$

Consequently, we conclude that at very high temperatures ferromagnetism must disappear and it can only exist below a critical temperature at which a root of (4.3a) other than $\frac{1}{2}$ starts to show the lowest free energy. Above the critical temperature magnetization only appears when an external field exists (paramagnetism). The transition at the Curie temperature T_c occurs as one of the second kind if $\theta=\frac{1}{2}$ is a double root of (4.3a) at T_c; i.e., when

$$\sum_i\left(\frac{\partial\mu_i}{\partial\theta}\right)_{\theta=\frac{1}{2}}=4\left(n-\frac{2p}{z}\right). \tag{4.5}$$

However, the transition may occur as one the first kind, if two minima of the free energy $F(\theta=\frac{1}{2}, T)$ and $F(\theta\neq\frac{1}{2}, T)$ become equal at a temperature higher than that for which $\theta=\frac{1}{2}$ gives a double root of (4.3a). For examples of the latter kind found in the order-disorder transition see Fig. 7 of Ref. III.

Let us first take the "pair" approximation ($n=2$, $p=1$). The two sites are symmetric and so $\mu_1=\mu_2=\mu$. We have

$$\varphi=X+2\mu+\mu^2X, \tag{4.6}$$

where

$$X=\exp(J/kT), \tag{4.7}$$

(2.2′) reduces to

$$\mu(1+\mu X)=\theta(X+2\mu+\mu^2X). \tag{4.8}$$

From (4.5) the critical temperature T_c is given by

$$J/kT_c=\ln(z/z-2), \tag{4.9}$$

which indicates that a positive exchange interaction is necessary for the ferromagnetism to appear. Using (4.3a) we find that the equilibrium θ below T_c is determined by the implicit equation

$$\frac{(2\theta-1)\theta^{1-1/z}(1-\theta)^{1/z}}{[\theta^{1-1/z}(1-\theta)^{1/z}]^2-\theta(1-\theta)}=X. \tag{4.10}$$

The reduced saturation magnetization $M(T)/M(T=0)$ is simply equal to $2\theta-1$.

From these established results we draw the following conclusions: (a) Ferromagnetic transitions are transitions of the second kind; the spontaneous magnetization of domains vanishes gradually at a certain Curie temperature. (b) The linear chain ($z=2$) should not exhibit ferromagnetism at any temperature and two-dimensional lattices should show ferromagnetism. And (c) different lattice structures with equal coordinate number z, such as the hexagonal layer and the simple cubic lattice, should have the same ferromagnetic characteristics. Weiss[17] obtained similar results by applying Bethe's method to the Ising model. Our method is capable of seeking higher approximations with ease. Higher approximations have been tried. The results are more sensitive to the lattice geometry; hence, conclusion (c) is incorrect owing to the roughness of the "pair" approximation. (a) and (b) remain true in higher approximations. However, the improvements offered by higher approximations are not very significant, because we are working with the Ising model which does not agree with reality and only serves to show the characteristics of the mathematical theory of ferromagnetism. For example, (b) has been disproved by both Bloch's method of spin function[18] and the Bethe-Weiss method. They show that only three-dimensional lattices can ever exhibit ferromagnetism.

It remains to be seen how we can modify the quasi-chemical method to fit the quantum-mechanical spin vector. This work will not only offer an alternative method to the Bethe-Weiss method but will be more effective in applications. Because the interactions among the first shell sites introduce involved calculations, Weiss did not carry out the work on the face-centered cubic and the close-packed hexagonal lattice. From Ref. I–III we have seen that the quasi-chemical method solves the face-centered cubic lattice without much more trouble than the simple cubic or body-centered cubic lattice.

Urbain *et al.*[19] have found that the ferromagnetism of metallic gadolinium (close-packed hexagonal structure) disappears suddenly at about 16°C (a transition of the first kind). This phenomenon cannot be understood by various theories so far established. Among them Stoner's collective electron theory,[20] which employs the band scheme of electron energy on a "standard" form, predicts in any case a continuous vanishing of ferromagnetism at the critical temperature. Bloch's method, being rigorous only at very low temperatures, has nothing to do with phenomena near the critical temperature. The Bethe-Weiss method suffers from involved calculations in the case of a closed-packed hex-

---

[17] P. R. Weiss, Phys. Rev. **74**, 1493 (1948); also U. Firgau, Ann. d. Physik **40**, 295 (1941).

[18] F. Bloch, Zeits. f. Physik **61**, 206 (1930).

[19] Urbain, Weiss, and Trombe, Comptes Rendus **200**, 2132 (1935).

[20] E. C. Stoner, Proc. Roy. Soc. **A165**, 372 (1938).

agonal structure, although this recent work improved Heisenberg's theory quite a lot by taking the actual energy levels instead of a Gaussian distribution. Perhaps the future work of a modified quasi-chemical method could explain the phenomenon mentioned above. Otherwise this experiment finding should be re-examined.

5. REGULAR MIXTURES CONTAINING MORE THAN TWO COMPONENTS

Ternary alloys (substitutional) and complex salts such as Ag_2HgI_4 (with Ag, Hg atoms and vacancies, or so called holes, distributed on a face-centered cubic lattice[21]) are the important examples of the systems to be considered in this section. So far no theoretical work has been done with them. Now let us consider a k component mixture in general, with k unspecified. We denote the atom of the Sth kind with A_s and their number by $N\theta_S(S=1, \cdots, k)$. The designation of distribution $(q_1, \cdots, q_n)$ of Section 2 has to be replaced by $(As_1, \cdots, As_n)$ when the ith site of the group is occupied by an A_{Si} atom $(S_i=1, \cdots, k)$. Here we may confine ourselves to the case of a random distribution, since the theory of order can be formed by merely a refinement. The conditions (2.1) and (2.2) now become

$$\sum_A [As_1, \cdots, As_n]=1, \tag{5.1}$$

$$\sum_{S_i=S} [As_1, \cdots, As_n]=\theta_s, \tag{5.2}$$

where $\sum_A$ is taken over all possible local configurations and $\sum_{S_i=S}$ over the possible configurations with the ith site occupied by an A_s. By an almost the same procedure as that employed in Section 2 we obtain the following equations:

$$[As_1, \cdots, As_n]=\frac{1}{\varphi}\mu(As_1)\cdots\mu(As_n) \times\exp[-\chi(As_1, \cdots, As_n)/kT], \tag{5.3}$$

$$\varphi=\sum_A \mu(As_1)\cdots\mu(As_n) \times\exp[-\chi(As_1, \cdots, As_n)/kT], \tag{5.4}$$

and

$$\mu_i(A_s)\frac{\partial \ln\varphi}{\partial\mu_i(A_s)}=\theta_s. \tag{5.5}$$

Each $\mu_i(A_s)$ may be regarded as a k-dimensional quantity and we may choose $\mu_i(A_k)=1(i=1, \cdots, n)$ without any loss of generality (as we did in Section 2). Altogether we have $n(k-1)$ parameters μ's.

The ternary equilibrium with a solute in two immicible solvents usually has an asymmetric solubility. A theoretical approach to this problem can be attempted by applying the equations obtained in this section under the supposition that all the atoms involved are non-electrolytic and that the model of "strict regular" liquids may be adopted.

The practical calculation of a problem with atoms of more than two different kinds would be much more involved than the applications we have made to binary systems. For binary mixture we can write χ as $V(=\frac{1}{2}[V_{AA}+V_{BB}-2V_{AB}])$ times an integer number by suitably choosing the zero level of the potential energy, but in the case of ternary mixtures χ consists of terms of the type $V_1=\frac{1}{2}(V_{AA}+V_{BB}-2V_{AB})$, $V_2=\frac{1}{2}(V_{BB}+V_{CC}-2V_{BC})$, and $V_3=\frac{1}{2}(V_{CC}+V_{AA}-2V_{CA})$. It is mainly the difficulties arising from this source instead of those arising from the presence of more variable μ's that makes the practical calculations for ternary mixtures cumbersome.

It is a pleasure to express my sincere thanks to both Dr. F. Seitz and Dr. J. S. Koehler for their interest in this work and their kind help.

---

[21] J. A. A. Ketelaar, Zeits. f. physik. Chemie **26B**, 327 (1934); **30B**, 53 (1935). At low temperatures Ag, Hg atoms, and "holes" form a highly ordered superlattice which gives the minimum energy when nearest interactions ($V_{Hg \cdot Hg} > V_{Hg \cdot Ag} > V_{Ag \cdot Ag}$) are counted.