American Mineralogist, Volume 84, pages 311–324, 1999

The description of Al, Si ordering in aluminosilicates using the cluster variation method

Victor L. Vinograd and Andrew Putnis*

Institute of Mineralogy, Münster University, D-48149, Münster, Germany

Abstract

The cluster variation method (CVM) for determining the thermodynamics of both short-range and long-range order in this study was extended to complex aluminosilicate minerals. Using recently developed algorithms for the maximization of cluster entropies, it is possible to construct high-order cluster variation approximations for Ising lattices of complex topology. This study gives the essentials of developing CVM models for the Al, Si nets of nepheline, leucite, feldspar, and cordierite. The robustness of the method is demonstrated by its ability to reproduce quantitatively recently published results of Monte Carlo simulations of Al, Si ordering in the Ising model with feldspar topology. The future prospects for the application of CVM to aluminosilicates are discussed.

Introduction

Many aluminosilicates exhibit Al, Si ordering that in many cases is observed macroscopically as regular alternation of average occupancies of the tetrahedral sites and subsequently as changes of lattice geometry. These macroscopically observed effects are known as long-range ordering (LRO). However, the phenomenon of ordering is more complex than can be inferred from X-ray diffraction and other methods sensitive to changes of the average structure. Recent investigations of aluminosilicates using ^{29}Si NMR spectroscopy have provided an insight into local properties of the Al, Si distributions. For example, the ^{29}Si NMR studies of synthetic cordierites and anorthites (Putnis et al. 1985, 1987; Phillips et al. 1992) have shown that significant variations in Al, Si distribution can occur even when the degree of LRO remains fixed. These variations are seen primarily as changes in the proportion of Al-O-Al, Si-O-Si, and Al-O-Si linkages in aluminosilicates, the phenomenon known as short-range ordering (SRO).

The equilibrium thermodynamic effects of Al, Si order-disorder in albite and plagioclase feldspars have been described using Landau models (Salje et al. 1985; Carpenter 1992; Holland and Powell 1992) in which variations in the thermodynamic properties of a mineral are related to a change in a macroscopic LRO parameter (Q_{od}). However, certain important characteristics of Al, Si order/disorder are beyond the scope of such an approach. In general, aluminosilicates can exhibit various degrees of SRO and LRO depending on composition and thermodynamic history of the samples. When LRO cannot develop due to compositional or kinetic constraints, the thermodynamic properties of a mineral may change as function of the degree of SRO. Therefore, interpretation of such experimental information requires modeling of Al, Si distributions as functions of the two order parameters.

When we introduce SRO into consideration, we must change our thermodynamic methodology from Landau models to Ising models. The main assumption of the Ising model is that the system of interacting particles adjust to an ordering force by changing its LRO and SRO parameters together. In the present study, we will assume that the ordering force is related to the energy difference (J_1) between the nearest-neighbor Al-O-Al, Si-O-Si, and Si-O-Al pairs:

$$J_1 = V_{AlAl} + V_{SiSi} - 2V_{SiAl} \quad (1)$$

As discussed recently by Dove et al. (1996), J_1 is the most important factor in the stabilization of ordering in aluminosilicates; however, a more correct analysis would require introducing a set of ordering constants J_n ($n = 1, 2, 3\ldots$) for pairs of first, second, and third neighbors. We will limit our analysis here to the assumption $J_1 = J$, $J_n = 0$ ($n \geq 2$).

The Al, Si distributions in various aluminosilicates have been modeled using Monte Carlo (MC) methods (Herrero 1991, 1993; Herrero and Sanz 1991; Thayaparam et al. 1994, 1996; Myers et al. 1998). However, the MC method can describe only the equilibrium (Boltzmann) distribution in which the LRO and SRO parameters have a certain functional relationship to each other and to the temperature and J. To describe non-equilibrium states of order, which are rather common in aluminosilicates, a method is needed that is capable of treating the LRO and SRO parameters as independent variables. We will consider an analytical approach to this problem based on the cluster variation method (CVM).

In recent decades, the CVM has been widely used for the description of atomic ordering in alloys (Inden and Pitsch 1991; de Fontaine 1979, 1994, and references therein). There have been relatively few studies applying

* E-mail: putnis@nwz.uni-muenster.de

0003–004X/99/0003–0311$05.00

the CVM to mineralogical problems. An excellent review of these studies has been given by Ross (1991). More recently, Herrero and Sanz (1991) proposed an elegant method for the description of Al, Si ordering in layer silicates. It was shown that using ^{29}Si NMR data, one can constrain the probability distribution (and entropy) of the four-point ''star'' cluster in the tetrahedral Al, Si layer and then use the ''star'' CVM approximation to work out the entropy of Al, Si distribution in micas. Phillips and Kirkpatrick (1994) used the same idea to calculate Al, Si entropies in leucite and analcime from spectroscopic data.

Unfortunately, the ''star'' approximation gives reasonable results only for aluminosilicates with rather low Al/(Al+Si) ratios. In our recent study (Vinograd and Putnis 1998), we have shown that for aluminosilicates with high Al/(Al+Si) ratios (e.g., anorthite, nepheline, margarite) the low-order CVM approximations (such as the star and pair) predict strongly negative entropies. Clearly, there is a need for the development of higher-order CVM approximations, i.e., those based on larger clusters. However, in the development of such approximations, certain technical problems arise.

The main problem is that one must evaluate the probability distribution (PD) of all possible configurations within the clusters and for large clusters this evaluation requires introduction of many configuration variables. This in turn causes mathematical and computational complexity. Computational difficulty has been the principal factor limiting the development of CVM in mineralogy. In fact, it has been widely assumed that there was no future for CVM in describing SRO in a structure as complex as a feldspar, even with the advent of increasing computer power. This is based on the fact that the traditional CVM techniques, when applied to clusters large enough to describe the feldspar structure adequately would require the use of hundreds of configuration variables.

In this paper, we demonstrate that using a new approach, this problem for Ising lattices disappears. The development of new computer algorithms allows avoidance of the use of variables corresponding to the high-order correlations within the clusters while continuing to treat these correlations correctly. The term ''high-order'' refers here to all groups of atoms that include pairs of second- and higher-order neighbors. The aim of the present paper is to show that by using these algorithms, it is possible to construct reasonably accurate CVM models for complex aluminosilicate lattices while avoiding most of the computational difficulties.

There is very little information published in the mineralogical literature on the details of the CVM approach, and because we demonstrate its wider applicability to aluminosilicate minerals in this paper, it is relevant to describe its main features first. Our aim is to describe the essential features of CVM pictorially, in a way that can be related relatively easily to aluminosilicate structures, and to describe the meaning of the algorithms referred to above in a similar way. We then apply this approach to show how the configurational entropy associated with Al, Si distributions in nepheline, leucite, plagioclase feldspars and cordierite can be determined. In the case of feldspars, we have the opportunity of demonstrating the virtually exact correspondence between the entropy determined by the MC method (Myers et al. 1998) and CVM.

The Ising model

The Ising model (Ising 1925; Onsager 1944) was introduced to describe magnetic ordering in solids and chemical ordering in alloys. According to this model a lattice of fixed topology is occupied by particles of two different types, A and B (which can be atoms with differently oriented spins or different atoms, or groups of atoms, such as AlO_4 and SiO_4 tetrahedra). It is assumed that nearest-neighbor particles interact with the following energies:

$$V_{AA} = V_{BB} = V_{AB} = V_{BA} = -V \qquad (2)$$

The negative sign of V corresponds to the tendency of the system to arrange spins in one particular direction (ferromagnetic ordering), whereas the positive sign favors their anti-parallel alternative arrangement (anti-ferromagnetic ordering). In cation ordering terminology, these tendencies correspond to phase separation or superlattice formation, respectively. We are interested in the case of positive V and positive J ($J = 4V$) because it allows us to describe the tendency to produce SRO by the avoidance of Al-O-Al contacts that can develop into a superlattice with regularly alternating Al-rich and Si-rich sites (LRO).

The equilibrium distribution in the Ising system is determined by the minimum of the configurational free energy. The system tends to find a compromise between energy and entropy factors. It tries to decrease the number of unfavorable nearest-neighbor linkages (to change the pair probability distribution) while keeping the entropy as high as possible under the given constraint on the pair distribution. Basically, when $J > 0$, there exists two strategies to decrease the number of the energetically unfavorable (AA) bonds. The first strategy is to increase the degree of local correlation between lattice points (increase SRO), whereas the second strategy is to develop periodic oscillations of site occupancies throughout the lattice (increase LRO). At high A/(A+B) ratios and at low temperatures, the second strategy usually prevails, whereas at low A/(A+B) ratios and high temperatures, the first strategy is favored. In many Ising lattices, a critical temperature typically exists above which the periodic oscillations of average site occupancies become unstable; this is the temperature of order/disorder transition. The critical temperature is a function of J, the A/(A+B) ratio, and the lattice topology. LRO disappears abruptly above the critical temperature, whereas the SRO generally persists up to reasonably high temperatures and disappears only in the limit of an infinitely high temperature.

These predictions of the Ising model are consistent with the existence of strong short-range ordering in alu-

minosilicates synthesized at low temperatures and with the existence of order/disorder transitions such as $I\bar{1}$-$C\bar{1}$ in anorthite-rich feldspars and *Cccm*-*P*6/*mmc* in cordierite. The Ising model also explains the high degrees of SRO in long-range disordered synthetic cordierites and anorthites measured by ^{29}Si NMR (Putnis et al. 1985, 1987; Phillips et al. 1992). The Ising model allows consideration of all these experimental observations as direct consequences of short-range interactions between Al and Si. Different temperatures of order/disorder transitions in different aluminosilicates can be attributed to different topologies of their Al, Si nets and to different Al/(Al+Si) ratios (Dove et al. 1996).

The CVM and other analytical methods allow investigation of the Ising model not only at equilibrium but also in the whole space of order and composition parameters. Typically, one constructs a non-equilibrium configurational free energy (F) of an A_xB_{1-x} system as a certain surface in the space of F, LRO, SRO, and x coordinates:

$$F/NJ = f = e - (kT/J)s \tag{3}$$

where N is the number of atoms in the system, $s = S/R$, and $e = E/(NJ)$ are the functions of LRO, SRO, and x parameters. S and E are the non-equilibrium configurational entropy and energy respectively. The equilibrium properties of the system are calculated by minimizing f with respect to LRO and SRO parameters at given values of x and kT/J.

Below we treat the configurational energy and entropy terms separately.

Configurational energy and pair probability distribution

The configurational energy in the nearest-neighbor Ising model can be exactly written as a function of the interaction energies V_{ij} and probabilities of the nearest-neighbor pairs, P_{ij}:

$$E = (Z/2)N \sum_{i,j} P_{ij}V_{ij} \tag{4}$$

where Z is the coordination number. Equation 4 can be applied to lattices in which the probability distribution of all the nearest-neighbor pairs is the same, and all lattice points have the same coordination number. In cases where there is more than one type of nearest-neighbor bond (e.g., T_1T_2 and T_2T_2 pairs in cordierite), Equation 4 splits accordingly into a number of terms where each term corresponds to a certain bond type.

In the general case, the pair distribution is a function of the distribution of its two point clusters and of the SRO parameter, which describes the strength of the statistical correlation between occupancies of the two points. When LRO is absent, the points have the same distribution and describing the point distribution ($P_A = x$, $P_B = 1 - x$) requires only the composition parameter x. In the case of LRO the two point distributions are different and define two sublattices, α and β. The difference between the average occupancies of α and β points can be described by the LRO parameter Q_{od}:

$$P_{A\alpha} = x(1 + Q_{od}), \quad P_{A\beta} = x(1 - Q_{od}) \tag{5}$$

When the occupancies of nearest sites are uncorrelated, the pair distribution can be found as a product of the two point probabilities (e.g., $P_{A\alpha A\beta} = P_{A\alpha}P_{A\beta}$, where $P_{A\alpha A\beta}$ is the probability of AA pair). In the case of extreme SRO (we are interested in the case where $0 < x < 0.5$), the probability of AA pairs is zero ($P_{A\alpha A\beta} = 0$). To describe the intermediate case, we introduce the SRO parameter σ ($0 < \sigma < 1$). Thus, the distribution of the ordered pair is a function of the three parameters x, s, and Q_{od} as follows:

$$\begin{aligned} P_{A\alpha A\beta} &= (1 - \sigma)P_{A\alpha}P_{A\beta}, \\ P_{B\alpha A\beta} &= P_{A\beta} - P_{A\alpha A\beta}, \\ P_{A\alpha B\beta} &= P_{A\alpha} - P_{A\alpha A\beta}, \\ P_{B\alpha B\beta} &= P_{B\alpha} - P_{B\alpha A\beta}. \end{aligned} \tag{6}$$

Table 1. Probability distribution of the pair cluster constrained by the parameters $x = 0.25$, $\sigma = 0.94$, and $Q_{od} = 0$

IJ	P_{ij}
AA	0.0025
AB	0.2475
BA	0.2475
BB	0.5025

An example of the pair distribution constrained by the parameters $x = 0.25$, $\sigma = 0.96$, and $Q_{od} = 0$ is in Table 1. Basically, the SRO parameters (or correlation functions) can be defined in a similar way for pairs of higher order or groups of neighbors. We will avoid the introduction of high order (multisite) correlation functions, because the probability distribution of any group of points in the nearest-neighbor Ising model (when $J_n = 0$, $n \geq 2$) in principle can be found as a function of x, σ, and Q_{od}. However, it is important to keep in mind that the distribution of high-order neighbors even in the nearest-neighbor Ising model is non-random. Therefore, we will use the general term ''high-order correlations'' to describe variables that could have been introduced to define non-random high-order pair probabilities. Basically, when we say ''correlations,'' we mean ''non-random pair probabilities'' (Sanchez and de Fontaine 1978, for the formal definition of the correlation functions).

Configurational entropy in the cluster variation approximation

The configurational entropy in the Ising model depends on correlations between all atoms of the system and thus is a complex function of the lattice topology. Fortunately, the statistical correlation between a pair of neighbors vanishes rapidly with the increase of the length of the sequence of bonds separating them in a lattice (Meirovitch 1977). This gives the possibility of accurately describing the entropy by considering probability distributions in

reasonably small clusters. The CVM takes advantage of this fact.

The CVM (Kikuchi 1951; Sanchez and de Fontaine 1978) allows one to write a series of approximations to the entropy as functions of cluster probabilities. Kikuchi and Brush (1967) have shown that the approximations tend to converge to the correct result as the cluster size increases. The entropy in a given approximation can be expressed as follows:

$$S = \mathrm{k}Ns = -\mathrm{k}N \sum_r \gamma(r) \sum_t p(r,t)\ln(r,t), \qquad (7)$$

where $p(r,t)$ is the probability of a certain (t) configuration of the r-cluster, k is the Boltzmann constant, and N is the total number of lattice points. The $\gamma(r)$ coefficients account for the number of clusters and for the sign of their contribution to the lattice entropy.

The clusters in the CVM are topologically distinct groups of lattice points connected by lines that denote correlations between the occupancies of the points. The cluster entropy is a simple logarithmic function of the cluster probability distribution (PD). The cluster PD is composed of probabilities of all the possible cluster configurations, where a configuration is defined by a set of $i, j, k, \ldots, n$ indices, and each index accepts either an A or a B state.

When constructing a CV approximation for a given lattice, one faces two main questions. The first question is what kinds of clusters are needed and how to calculate the $\gamma(r)$ coefficients for them. The second is how to calculate the probability distributions and entropies of the chosen set of clusters. The answers to these questions appear in the studies of Kikuchi (1951), Barker (1952), Hijmans and de Boer (1955), Sanchez and de Fontaine (1978), and Sanchez et al. (1984). These questions have been discussed traditionally only in relation to highly symmetric lattices, such as the simple cubic, fcc, and bcc, which require the use of compact clusters (e.g., cubes, tetrahedra, and octahedra). These methods become cumbersome for the less-symmetric lattices that appear in aluminosilicates.

The most difficult problem concerns the evaluation of cluster entropies. Sanchez and de Fontaine (1978) have shown that the number of correlation functions needed to describe a cluster PD is equal to the number of all distinct subclusters into which the given cluster can be decomposed. This number increases exponentially with cluster size, and for less-symmetric, clusters the increase is even faster than for the symmetric ones. As we will show, the description of aluminosilicate lattices typically requires using large and non-symmetric clusters. For example, the CVM model of the feldspar lattice must be based on clusters containing up to 16 points. The traditional technique would require using hundreds (maybe thousands) of configuration variables and then to minimize the free energy of the model with respect to these variables.

To limit the number of variables, we have developed new algorithms for evaluating cluster entropies. Some progress in this direction has already been reported by Vinograd et al. (1997). Here, we will address the problem of calculating cluster entropies in relation to the specific cluster topologies relevant to aluminosilicates.

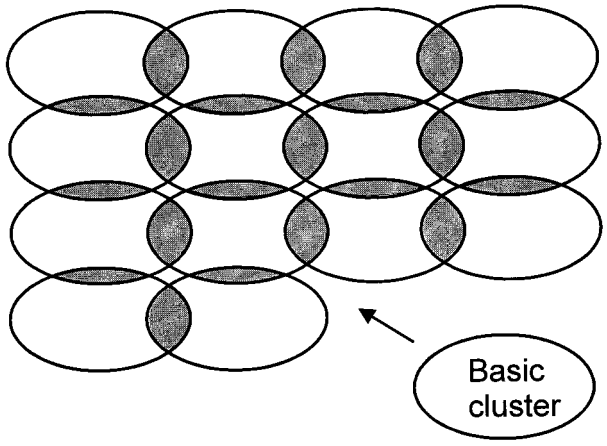


Figure 1. The scheme of construction of a lattice from basic clusters. Shaded areas correspond to subclusters of the basic clusters, which are duplicated in the construction process.

Another problem is that the methods for derivation of the $\gamma(r)$ coefficients typically are described in the literature on the CVM in a formal way that requires counting of all possible subclusters of the given basic cluster. This is appropriate for clusters of high symmetry and small size where the ennumeration of the subclusters is not very difficult. In cases of large clusters of complex topology, this procedure is cumbersome. In the present study, we found it more convenient to use a less-formal geometrical scheme for the derivation of CVM equations described by de Fontaine (1979) and illustrated for the case of the simple cubic lattice. We will show here that this scheme can be adapted readily to various aluminosilicate lattices.

Decomposition of the total entropy into cluster contributions

The idea underlying the decomposition of lattice entropy into cluster terms is that a typical configuration of A and B particles in a lattice can be regarded as a complex stochastic event (Alexandrowicz 1971). The probability of this global event (P_{typ}) can be approximated as a product of probabilities of local events, where each event corresponds to an adsorption of a certain basic cluster to a set of already adsorbed clusters (Fig. 1). The inverse to the probability of the typical lattice configuration ($W = 1/P_{\mathrm{typ}}$) is the number of accessible configurations, and the entropy is given by k ln$W$. The problem basically consists of calculating the total probability of a typical configuration composed of an infinite set of overlapping cluster configurations. Clearly, one can construct a continuous lattice only from overlapping clusters. Therefore, each event of adsorption must be accompanied by the events of desorption or subtraction of certain duplicated elements (subclusters) of the basic clusters. The contribution

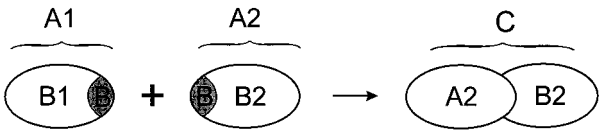


FIGURE 2. The scheme of construction of a complex cluster C from clusters A1 and A2, which have a common subcluster B.

of adsorption and desorption events to the total entropy is of opposite sign.

To illustrate the last statement, we consider a case where the total configuration (denoted as the supercluster C) is composed of two (previously isolated) clusters A1 and A2 overlapping through their common subcluster B (Fig. 2). The probability of an $i_1,i_2,i_3 \ldots i_n$ configuration of the cluster C can be considered as a probability of a stochastic event (C) composed of the events (A1) and (B2) where B2 is a part of the cluster A2 that has no common elements with B. Each of the events (A1) and (B2) corresponds to a certain subset of indexes $i_1,i_2,i_3 \ldots i_n$. Bayes formula for the probability of a complex event can be written:

$$P(\mathrm{C}) = P(\mathrm{A1})P(\mathrm{B2/A1}) \tag{8}$$

where P(B2/A1) is the probability of the event (B2) conditional on the event (A1). Similarly, the event (A2) can be considered as a combination of the events (B) and (B2):

$$P(\mathrm{A2}) = P(\mathrm{B})P(\mathrm{B2/B}) \tag{9}$$

where P(B2/B) is the probability of the event (B2) conditional on the event (B). Because the clusters A1 and A2 previously were isolated from each other, the event (B2) does not depend directly on the event (A1) but depends only on the event (B). Clearly, P(B2/A1) = P(B2/B). Now, using Equation 9, Equation 8 can be rewritten as follows:

$$P(\mathrm{C}) = P(\mathrm{A1})P(\mathrm{A2})/P(\mathrm{B}) \tag{10}$$

Equation 10 relates the probability distribution of the combination of clusters A1 and A2 to the product of their probabilities and the probability of their common subcluster B. From Equation 10 it follows that the entropy of the complex cluster C can be written as a combination of positive entropy terms related to its two overlapping subclusters A1 and A2 and the negative term related to their overlapped common element B:

$$k \ln\left\{\begin{array}{c}\mathrm{A1}\ (\mathrm{B})\ \mathrm{A2}\end{array}\right\} = k \ln \frac{\{\mathrm{A1}\}\{\mathrm{A2}\}}{\{\mathrm{B}\}} \tag{11}$$

Here the terms taken in brackets represent the so-called Kikuchi short-hand notation for the numbers of cluster configurations. Table 2 further explains this notation for the case of simple clusters consisting of points, pairs, and squares. Table 2 also shows that there is no need to evaluate the cluster combinatorial terms in the combinatorial way. It is easier to evaluate directly the logarithms of these terms. The combinatorial terms and the short-hand notations are introduced here because the CVM equations appear more compact if written in these notations. Equation 10 is well known in the CVM literature (e.g., Kikuchi and Brush 1967) as the superposition approximation. We will show below that this equation forms the basis of the CVM.

TABLE 2. Short-hand notations for the combinatorial terms and entropy formulas for different clusters

Cluster	Combinatorial term	Cluster entropy
point	$\{\bullet\} = \dfrac{N!}{\prod_i (NP_i)!}$	$k \ln\{\bullet\} = -R \sum_i P_i \ln P_i$
pair	$\{\bullet\!-\!\bullet\} = \dfrac{N!}{\prod_{i,j} (NP_{ij})!}$	$k \ln\{\bullet\!-\!\bullet\} = -R \sum_{i,j} P_{ij} \ln P_{ij}$
square	$\left\{\square\right\} = \dfrac{N!}{\prod_{i,j,k,l} (NP_{ijkl})!}$	$k \ln\left\{\square\right\} = -R \sum_{i,j,k,l} P_{ijkl} \ln P_{ijkl}$

Using Equation 10 one can construct the PD of the cluster C preserving the distribution of its parent subclusters A1 and A2. Therefore, if the subclusters have a specified distribution of nearest-neighbor pairs, then the supercluster C also will have the same distribution of nearest-neighbor pairs. By using a sequence of Equations 10 and 11, one can construct the whole lattice (and calculate its entropy) as a function of the given pair distribution.

By relating the clusters A1 and A2 to pairs and the subcluster B to their common point subcluster, one can easily derive an entropy equation for a three-point chain fragment (Fig. 3, left column). Using Equation 10 and applying it in an iterative way to the already constructed fragments, it is possible to derive a general equation for the combinatorial term (and entropy) of the chain fragment composed of n pairs:

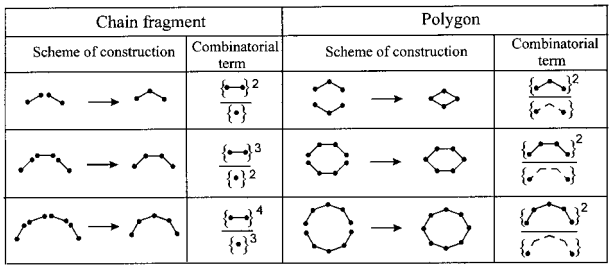


FIGURE 3. The schemes of construction of chain fragments from pair clusters and of polygons from the chain fragments. The combinatorial terms correspond to the constructed clusters. The denominators of the combinatorial terms correspond to the subclusters, which have been duplicated in the construction process. Dashed lines show the geometry of the parent cluster and indicate that the specific combinatorial term corresponds to a subcluster of the parent cluster.

$$\left\{\begin{matrix}n\text{-pair chain}\\ \text{fragment}\end{matrix}\right\} = \frac{\{\bullet\!-\!\bullet\}^n}{\{\bullet\}^{n-1}} \tag{12}$$

where $n - 1$ is the number of points. Actually, Equation 12 is valid for any finite set of connected pairs (Bethe lattice) that do not contain rings of clusters (polygons). For an infinitely long chain ($n \to \infty$), one obtains:

$$\{\bullet\!-\!\bullet\!-\!\bullet\!-\!\bullet\!-\cdots\} = \left(\frac{\{\bullet\!-\!\bullet\}}{\{\bullet\}}\right)^n. \tag{13}$$

Here n corresponds to the number of moles of pairs in the chain. Generally, we will write the CVM formulae with reference to 1 mol of a particular cluster and omit n from equations such as Equation 13. The important result is that the entropy of an infinitely long chain contains *equal* numbers of positive and negative terms. This is also true for infinitely long chains composed of clusters of other types.

The combinatorial term for a chain composed of symmetric clusters A intersecting through their common subclusters B can be written:

$$\{\text{A B A B A B}\cdots\} = \frac{\{\text{A}\}}{\{\text{B}\}} \tag{14}$$

The more general case of a chain composed of different clusters A1 and A2 intersecting through their common B1 and B2 subclusters can be written:

$$\{\text{A1 A2 B2 A1 A2 B2}\cdots\} = \frac{\{\text{A1}\}\{\text{A2}\}}{\{\text{B1}\}\{\text{B2}\}} \tag{15}$$

Equation 15 also can be applied to a chain composed of similar but differently oriented asymmetric clusters. For example, a chain of ordered pairs can be written:

$$\{\circ\!-\!\bullet\!-\!\circ\!-\!\bullet\cdots\} = \frac{\{\circ\!-\!\bullet\}^2}{\{\circ\}\{\bullet\}} \tag{16}$$

Here, the black and white dots denote point clusters with different distributions. The combinatorial formula for the chain composed of square clusters takes the form:

$$\{\cdots\} = \frac{\{\square\}}{\{\bullet\!-\!\bullet\}} \tag{17}$$

Clearly, the pair cluster plays the role of the common subcluster of two adjacent square clusters.

Equation 14 can be applied even to chains in which the chain elements are in turn represented by infinitely long chains. De Fontaine (1979) has shown that the entropy of the square Ising lattice can be evaluated in the exact way as ''the chain of ladders'' (Fig. 4):

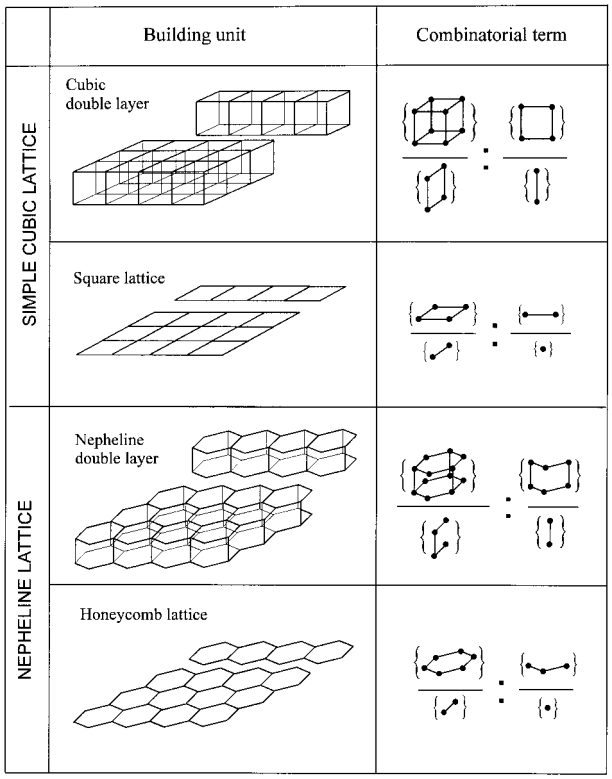
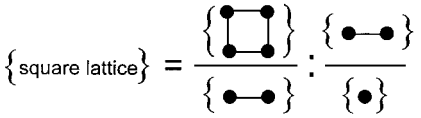


FIGURE 4. The two-dimensional building units of the simple cubic lattice and of the nepheline lattice. The combinatorial terms correspond to the relevant building units. Term is composed of a nominator and a denominator. The nominator corresponds to a chain unit (shown separately), and the denominator corresponds to a subcluster of this chain unit, which is duplicated in the process of construction of the particular two-dimensional unit from the chain units.

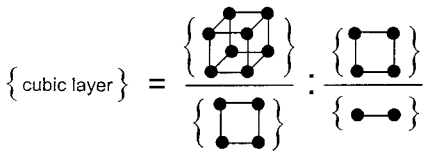
$$S_{\text{square lattice}} = k \ln \frac{\{\square\square\}}{\{\square\square\}} \tag{18}$$

The dashed lines here are intended to show that the distribution of the cluster that appears in the denominator of Equation 18 must be calculated not as an independent single chain, but rather as a subcluster of the ladder cluster that appears in the nominator. This can be done by summation of the ladder PD over the points corresponding to one row of the ladder. The exact evaluation of the PD of the ladder cluster is difficult, however. The CVM avoids this by approximating the distribution of the ladder PD as that of a chain of square clusters. The PD of the subcluster that appears in the denominator of Equation 18 is approximated as a PD of the single chain of pairs. By substituting Equations 13 and 17 into Equation 18, the combinatorial formula of Kikuchi (1951) for the square net, can be obtained:



(19)

Equation 19 is a rather accurate model for the entropy of the square net. The approximation comes primarily from the fact that the PD of the (isolated) square cluster, which appears in Equation 19 is not exactly the same as the distribution of the square cluster taken from a continuous net. De Fontaine (1979) has shown that by using similar ideas, one can construct an entropy equation for a double layer composed of cubes (Fig. 4), treating it as a chain of cubic chains:



(20)

Finally, the three-dimensional simple cubic lattice can be treated as a chain of double layers composed of cubes. Square lattices serve as the common subclusters of these layers; therefore:

$$\{\text{cubic lattice}\} = \frac{\{\text{cubic layer}\}}{\{\text{square lattice}\}}. \qquad (21)$$

It is important to note that the entropy term that appears in the denominator of Equation 21 is not exactly the same as the entropy of the square in Equation 19 where it was not related to the cubic lattice. The difference is that when a square lattice is treated as a common subcluster of the cubic layers, the PD of the square must be calculated as that of the square subcluster of the cubic cluster but not as the PD of the isolated square.

Derivation of entropy formulae for aluminosilicates

A similar construction scheme based on Equations 19, 20, and 21 applies to Al, Si nets of aluminosilicates. The key point is to define and construct a ''double layer'' starting from a polyhedron (basic cluster) or several polyhedra that play the same role as the cubic cluster in the cubic lattice.

The nepheline lattice can be considered as a chain of double layers formed by two interconnected honeycomb lattices. Flat honeycomb units [layers parallel to (001)] represent the common subclusters of the adjacent layers. The analogy between structural units of simple cubic and nepheline lattices (Fig. 4) can be written:

$$\{\text{nepheline lattice}\} = \frac{\left\{\begin{matrix}\text{nepheline}\\ \text{double layer}\end{matrix}\right\}}{\{\text{honeycomb lattice}\}}. \qquad (22)$$

The combinatorial terms corresponding to the double layer and to the honeycomb unit are expanded further in Figure 4. These units in turn are considered as replications of smaller chain units. The double layer is constructed from polyhedral chains that overlap through polygonal chains. The honeycomb lattice in turn is considered as a chain of chains composed of hexagons. The hexagonal chains overlap through single chains. The combinatorial term of the honeycomb net is thus composed of the two terms corresponding to the hexagonal and single chains (Fig. 4).

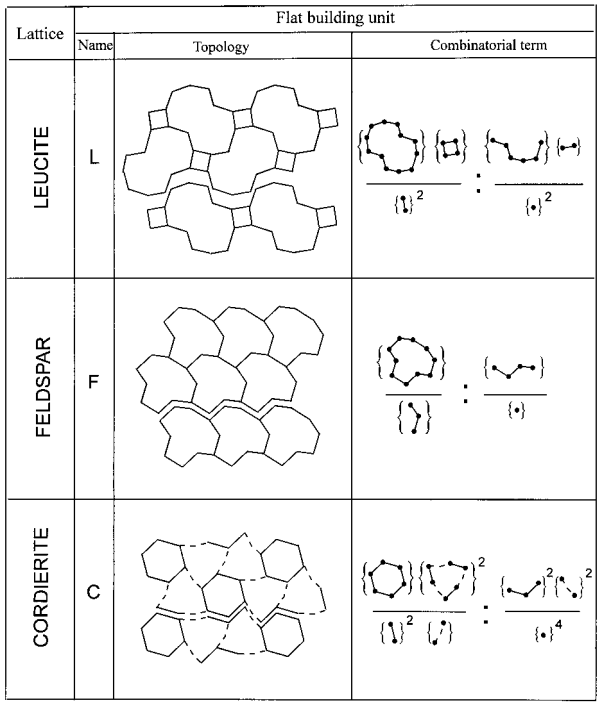


Figure 5. The flat building units of Al, Si nets in aluminosilicates; the notation is the same as in Figures 3 and 4.

When using Kikuchi's notation for subclusters of the basic clusters, we will usually draw them preserving their exact geometry as elements of the basic clusters. In some cases, we will use dashed lines to indicate the geometry of the parent basic cluster. It is important to realize that the PD of a given subcluster is obtained from the PD of the parent basic cluster by summation over the ''missing'' points, i.e., points that are not included in the subcluster. The dashed lines help to understand over which points the summation was performed. The summation technique will be further explained in more detail.

Figures 5 and 6 present similar developments of the entropy formulae for leucite, feldspar, and cordierite. Figure 5 considers the scheme of evaluation of combinatorial equations for flat building units, and Figure 6 considers the topological units produced when these flat building units are stacked on top of each other to form double layers. Here, we will stress again that the polygons that form the flat units are considered as subclusters of the basic clusters that form double layers.

The combinatorial terms of the flat building units are considered as the products of terms related to polygonal

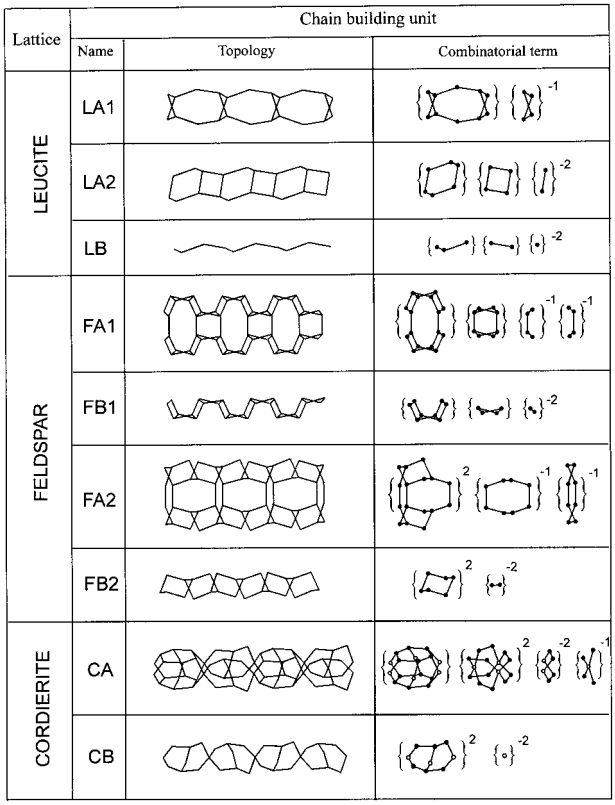


FIGURE 6. The chain A-units (e.g., LA1 and LA2 in leucite) and their subclusters B-units (e.g., LB in leucite) needed for the construction of the double layers of the Al, Si nets. The combinatorial terms that correspond to A- and B-units are given in the right part of the figure. (In the cordierite clusters, white dots designate T_1 sites.)

and single chains. These terms are exactly analogous to the two terms of Equation 19. In the same way, the combinatorial terms of the double layers are the products of terms related to ''polyhedral'' chains (A-chains), and to ''polygonal'' chains (B-chains), which play the role of common subclusters of two adjacent polyhedral chains. Figure 6 shows that the topology of the elements of A-chains in some cases can be described better not by polyhedral units but by polygons. Similarly, the topology of B-chains in some cases can be represented better by single chains. This happens because the coordination number of aluminosilicate nets is lower than that of the cubic lattice. Because of the low coordination number, the clusters in aluminosilicates in fact may have a geometry that is intermediate to polygons and polyhedra.

In the case of leucite, the double layer is constructed from two different chains LA1 and LA2, which intersect through their common subcluster (LB-chain). The chain LA2 is an example of a polygonal chain that plays the role of the polyhedral chain in the construction scheme. The combinatorial term of the leucite double layer can be written as $\{LA1\}\{LA2\}/\{LB\}^2$.

In the feldspar lattice, the flat building unit F (Fig. 5) forms double layers parallel to (001) of two types (Fig. 6). The first layer can be described by replicating the chain FA1, and the second one can be constructed by replicating the chain FA2. The entropies of these layers are {FA1}/{FB1} and {FA2}/{FB2} respectively.

Similar ideas can be used to construct the double layer and the flat building unit C of cordierite. The double layer, parallel to (001), is constructed from chains CA composed of polyhedra of two different shapes. The cordierite structure is usually visualized as rings of T_2 tetrahedra joined by T_1 tetrahedra (Putnis 1992). The first polyhedron (cordierite cage) is formed by two $6T_2$ rings connected through six $4T_22T_1$ rings. The second polyhedron is formed by three $4T_22T_1$ rings connected through T_1 corners or by two 9-member $6T_23T_1$ rings connected through their subcluster $3T_1$. The flat unit C is formed by $6T_2$ hexagons and by $6T_2$ subclusters of the nine-member $6T_23T_1$ rings.

To arrive at the final expressions for the entropies of the 3D lattices, one has to consider these lattices as chains composed of double layers, which intersect through their common subclusters represented by the flat lattices on Figure 5. Therefore, the combinatorial terms that correspond to the double layers must be divided by the terms that correspond to the flat building units. The combinatorial term for the feldspar lattice takes the form $(\{FA1\}/\{FB1\})(\{FA2\}/\{FB2\})/\{F\}^2$. The choice of the double layers and the way they can be decomposed into cluster terms is, of course, arbitrary, and thus one can construct a number of CVM approximations for a given lattice. However, if one uses basic clusters of similar size and correctly accounts for their overlaps, the result would be virtually the same.

The total entropy is thus composed of the positive entropy terms corresponding to the basic clusters and the terms of different sign corresponding to the overlapping elements (subclusters) of the basic clusters. The PD of the subclusters are trivially obtained from the PD of the basic clusters through summations over the missing points. The distribution of the polygons that form the flat building units usually can be found directly by the appropriate summations over the points of basic polyhedra. However, in the model for leucite, the flat building unit contains the 10-point cluster which is not included as a part in any of the leucite basic clusters. The PD of this cluster is then constructed from chain subclusters of the basic clusters in a manner similar to the scheme of Figure 3. The main problem is to calculate the PD of the basic clusters.

CALCULATION OF ENTROPIES OF THE BASIC CLUSTERS

The main task here is to construct the PD of the basic clusters as products of probabilities of smaller cluster elements using Equation 10. Basically, one can start from the point and pair clusters and construct chain fragments. Using chain fragments as starting elements, one can then construct polygons as shown in Figure 3, and starting

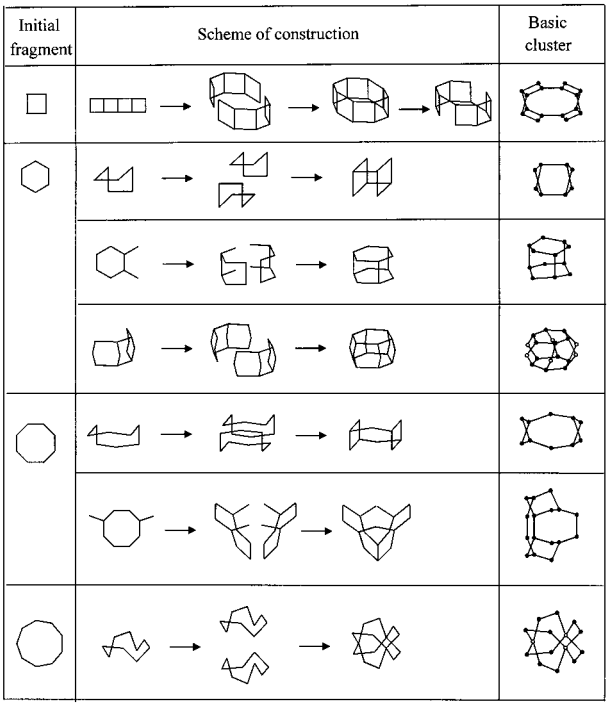


FIGURE 7. The schemes of construction of basic clusters from fragments of polygonal chains.

from polygons, one can arrive at PD of different polyhedra (Fig. 7). The question then is why one needs the concept of the basic cluster if the PD of a basic cluster can be found as a function of point and pair probabilities. The problem is that the schemes of derivation shown in Figures 3 and 7 lead to cluster probability distributions that do not have the maximum entropy.

To arrive at a correct value of the total entropy from a CVM equation, one has to maximize entropies of the basic clusters. From the above derivations, it is evident that the total entropy of a lattice in a CVM approximation is a function of the probability distributions of the basic clusters. Maximum total entropy, therefore, then coincides with the maximum entropy of the basic clusters. The final requirement is to calculate the minimum value of the free energy. However, at a fixed set of the order and composition parameters (and, consequently, fixed value of the configurational energy), the minimum of the free energy coincides with the maximum of the entropy.

At a fixed set of point and pair probabilities, the entropy of a basic cluster can still vary as a function of high-order correlations. However, construction schemes of Figures 3 and 7 result in asymmetric probability distributions where the high-order correlation functions do not have optimum values. The basic clusters are thus needed as the objects on which the procedure of optimization of the high-order correlations (maximization of the cluster entropies) can be carried out.

The traditional approach (Sanchez and de Fontaine 1978) involves evaluating the cluster entropies as symmetric analytical functions of a generally large set of configuration variables. These variables are then determined from the condition of the minimum of the free energy of a CVM model using standard optimization methods for multi-variable functionals. Our approach is different.

In the Ising system with nearest-neighbor interactions, one can divide the total number of configuration variables into two unequal sets. The small set includes the variables whose equilibrium values depend on both the energy and entropy factors. These are the three variables that determine the probabilities of the nearest-neighbor pairs. The large set includes the variables that constrain the correlations between higher-order neighbors. These variables depend only on the entropy factor, because the configurational energy (Eq. 4) is insensitive to the high order correlations. Therefore, the procedure of minimization of the free energy can be fulfilled in two steps: fix the variables of the small set and maximize the entropy of a CVM model (i.e., the entropy of the basic clusters) with respect to the large set of variables; minimize the free energy with respect to the small set of variables. The simplification arises from the fact that the maximization of the cluster entropies can be achieved without introducing the variables of the large set.

The maximization of cluster entropies can be done in two stages. First we construct initial cluster probability distributions according to the schemes of Figures 3 and 7. Then we reconstruct (anneal) these clusters and increase their entropy and symmetry up to the maximum.

CONSTRUCTION STAGE

An analysis of Figures 4, 5, and 6 shows it is necessary to evaluate probability distributions and entropies of clusters of three groups. These are the chain fragments (we include fragments of branched chains in this group), polygons, and polyhedra. We also need to know the entropies of the subclusters of these clusters. If the cluster distribution is known, the distribution of its subcluster can be calculated easily by taking a sum over a set of points of the cluster that are not included in the subcluster. For example, the distribution of the 2-point I..L subcluster of the chain fragment IJKL can be found as follows:

$$P_{i..l} = \sum_{j,k} P_{ijkl} \tag{23}$$

The constructive stage is based entirely on Equations 10 and 11. To use these equations, one must be sure that both the parent clusters A1 and A2 have a commom subcluster (the subcluster B must have identical PD in both the A1 and A2). This condition is always met when A1 and A2 intersect through a point or through a nearest-neighbor pair (we assume that the A1 and A2 are always constrained to have the correct PD of points and pairs). One can also be certain that the condition is met if A2 and A1 have identical PD. This means that the cluster A2 can be considered as a symmetrical image of A1. Usually, many clusters can be considered as combinations of two specific symmetrically related elements. The PD of these

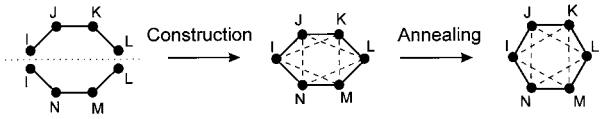


FIGURE 8. Two stages of evaluation of the probability distribution of the hexagon. The first ''construction'' stage results in the initial PD with asymmetric second- and third-neighbor correlations. The second (annealing) stage alters the distribution of the second and the third neighbors and makes them symmetric. Dashed lines denote second-order correlations. The dotted line represents the ''reflection operation,'' which forms the basis of the construction scheme.

clusters thus can be constructed easily using Equation 10. When A2 is related symmetrically to A1, we will refer to Equation 10 as the ''reflection operation.''

Figure 3 shows that probability distributions and the entropies of even-member rings (polygons) can be calculated by sequential application of the reflection operation, first to pairs and then to chain fragments. The probability distribution of the common subcluster of two chain fragments, which is needed in this procedure, can be determined using equations analogous to 23. For example, the PD of a hexagon IJKLMN (Fig. 8) is the following:

$$P_{ijklmn} = P_{ijkl}P_{inml}/P_{i..l} \tag{24}$$

where

$$P_{ijkl} = P_{ij}\,P_{jk}\,P_{kl}/(P_j\,P_k) \tag{25}$$

and where $P_{i..l}$ is given by Equation 23. Equation 25 follows from Equation 10 when it is applied successively to pair fragments. Figure 7 shows that by using similar reflection operations, one can construct various polyhedra starting from polygons, polygonal chains, or from combinations of polygons with fragments of single chains.

ANNEALING STAGE

The probability distributions of the basic clusters which can be calculated according to the schemes of Figures 3 and 7, are only approximations to the desired results. We need to construct PD with maximum entropy but the constructed clusters have lower entropy due to the asymmetry of the construction scheme. The construction scheme (reflection operation) constrains only the probabilities of the nearest-neighbor pairs to be identical, but the probabilities of second-neighbor and higher-order pairs in the constructed supercluster may not correspond to its full symmetry. For example, in the IJKLMN hexagon resulting from Equation 24, the probabilities of second- and third-neighbor pairs are not all the same. Due to the special property of the entropy function, the asymmetric distribution has lower entropy than the symmetric one.

The aim of the annealing algorithm is to increase the entropy of the initially asymmetric cluster by changing the distribution of its high-order neighbors. This change is achieved in the course of a series of reconstructions (self-transformations) of the cluster. The idea is based entirely on the superposition operation (Eq. 10), which is now applied to the initially constructed cluster. In fact, using appropriate summation operations, one can ''decompose'' the supercluster into its subclusters A1, A2, and B (Fig. 9), then recombine them into the supercluster using Equation 10. In most cases, there are many ways of choosing the set of A1, A2, and B, and if one performs a series of different decompositions and recombinations, it is possible to increase the symmetry and entropy of the supercluster. The fact is that during the annealing step (decomposition + recombination) the entropy of the supercluster can only increase or remain unchanged. During the decomposition (summation operations) the correlations between the points of the B1 and B2 subclusters become lost. During the recombination, these correlations are then restored but in a different way. In fact, in the process of recombination, the A1 and A2 subclusters already are considered as totally independent (isolated) subclusters, and the correlation between the points of B1 and B2 is restored only as their common dependence on points of the B subcluster. The independent correlations between the points of B1 and B2 (if they existed in the original supercluster) are now substituted by the dependent correlations. The dependent correlations are generally less rigid then the independent ones, and thus the total entropy of the supercluster increases.

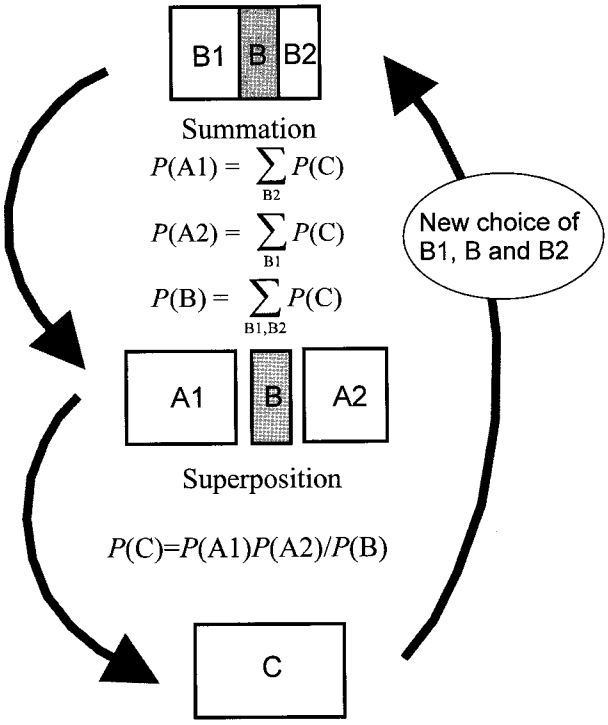


FIGURE 9. The general annealing scheme for basic clusters. During the annealing process, the PD of the supercluster C is decomposed into the PD of its A1, A2, and B subclusters, then restored using the superposition operation. The indices in the summation equations show that the summations are performed over the points of the C included in the specified subclusters.

TABLE 3. Probability distribution of the square cluster constrained by the parameters $x = 0.25$, $\sigma = 0.94$, and $Q_{od} = 0$

IJKL	P_{ijkl} before annealing	P_{ijkl} after annealing
BBBB	0.1948599890	0.2101048623
BBBA	0.1418150110	0.1449879443
BBAB	0.1633863971	0.1449879443
BABB	0.1418150110	0.1449879443
ABBB	0.1633863971	0.1449879443
BBAA	0.0024386029	0.0024192492
ABBA	0.0024386029	0.0024192492
AABB	0.0024386029	0.0024192492
BAAB	0.0024386029	0.0024192492
BABA	0.1032099890	0.1000524393
ABAB	0.0816500076	0.1000524393
BAAA	0.0000363971	0.0000403673
ABAA	0.0000249924	0.0000403673
AABA	0.0000363971	0.0000403673
AAAB	0.0000249924	0.0000403673
AAAA	0.0000000076	0.0000000163

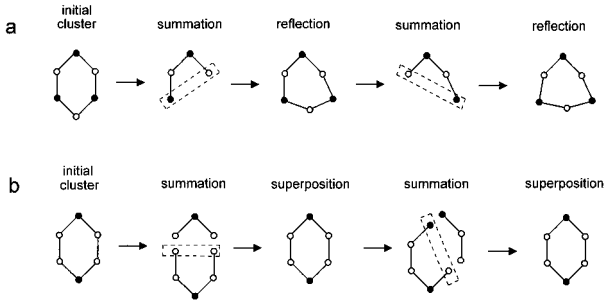


FIGURE 10. Different schemes of annealing algorithms for the hexagon. The scheme (**a**) is based on the reflection operation, whereas the scheme (**b**) is based on the superposition operation. The latter scheme is referred to as the general annealing algorithm in the text. The clusters included by the dashed lines are the ''B'' subclusters referred to in Equation 10.

Table 3 illustrates the change of the PD of the square cluster as the result of the annealing algorithm. The second column of Table 3 shows the PD of the asymmetric square cluster resulting from the application of the reflection operation (Fig. 3) on a three-point chain fragment. The third column shows how the square distribution changes as a result of the annealing process. One can see that the probabilities of all symmetrically related configurations of the square become equal. Vinograd et al. (1997) have shown that the entropy of the square cluster resulting from the annealing algorithm, if substituted into Equation 19, results in the correct entropy of the square lattice (''square'' CVM approximation). The subsequent minimization of the free energy of the square Ising model with respect to x and σ parameters then leads to the value of the Curie temperature, which coincides with the result of Kikuchi (1951). This coincidence proves that the entropy of the square cluster reaches the unique thermodynamic maximum during the annealing algorithm.

The annealing algorithm allows the number of independent variables describing the basic cluster distribution to be reduced greatly. For example, using the annealing algorithm, it is possible to evaluate the ''square'' approximation for the square lattice Ising ferromagnet using of just two parameters x and σ (Vinograd et al. 1997), whereas the traditional approaches (Kikuchi 1951; Sanchez and de Fontaine 1978) would require use of five independent ''fraction variables'' or correlation functions to treat the same problem. The advantages of the annealing algorithm become more evident in cases of approximations based on large clusters.

The annealing algorithm based on the sequential application of the reflection operation [$P(C) = P(A1)^2/P(B)$] is illustrated in Figure 10a for the hexagon. It also can be applied to any even-member ring, if the LRO scheme of the cluster is consistent with the reflection operation. Some ordering schemes (e.g., Fig. 10b) and some cluster topologies prevent the use of the reflection operation. In these cases, one can use a general annealing algorithm based on the superposition operation (Eq. 10). The idea is to choose any point in the C supercluster and consider it as a B1 subcluster, then the A1 subcluster is defined as the B1 plus all its nearest-neighbors (the A1 thus has the geometry of a star cluster). The B subcluster then includes all points of the star excluding the central (B1) point, and the A2 subcluster includes all points of the C supercluster excluding the single point of B1. The annealing step is then performed according to the given definition of A1, A2, and B subclusters and then any other point of C can be chosen for the next annealing step. The cluster distribution converges to that of the maximum entropy if the algorithm runs over all points of the C. A few cycles over the points of C ordinarily are enough for the entropy to converge. This generalized annealing algorithm was used in the present study to calculate the PD of two 16-point and one 8-point basic clusters that appear in the CVM model of the feldspar lattice.

It is important to mention that the actual connectivity of the cluster can be modified inside the annealing algorithm by defining specific coordination of each point of the cluster. Therefore, the connectivity of the initial cluster can differ from the final connectivity. However, the connectivity of the initial cluster must not be lower than that of the final cluster. For example, the 16-point basic cluster shown in the first row of Figure 7 can be constructed from the eightfold prism. The correlations between two pairs of points can be suppressed during the annealing. Similarly, the cluster shown in the second row of Figure 7 can be constructed from a cube.

COOPERATIVE ORDER/DISORDER TRANSITION IN THE FELDSPAR LATTICE

The equation for feldspar entropy

$$S = k \ln(\{FA1\}\{FA2\}/\{FB2\}/\{FB1\}/\{F\}^2) \quad (26)$$

has been derived above for the case of the absence of LRO. In fact, it can be easily generalized for the LRO case if one starts constructing the basic clusters from ordered pairs. The pair distribution (Eq. 6) is a function of

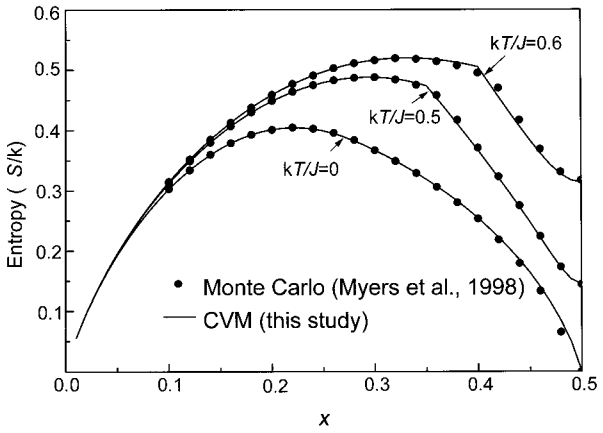


FIGURE 11. Equilibrium configurational entropy of Al-Si in the feldspar lattice at different scaled temperatures. Solid lines are the result of the present CVM approximation, and the dots are the results of the Monte Carlo study (Myers et al. 1998). The kinks in the lines correspond to the onset of the order/disorder phase transition.

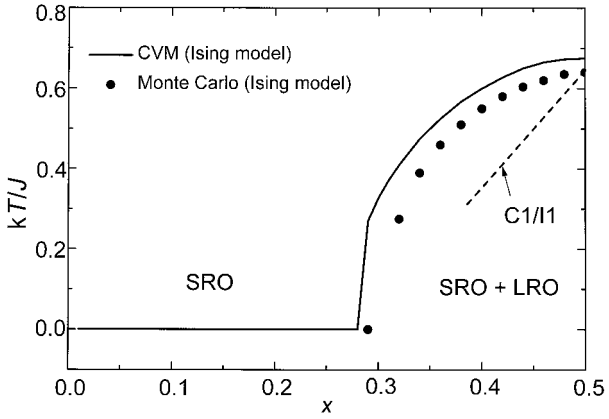


FIGURE 12. The dependence of the critical ordering temperature on the Al/(Al+Si) ratio in plagioclase according to theoretical calculations (CVM and Monte Carlo) and experimental data. The dashed line shows the experimental data of Carpenter and McConnell (1984) on the location of $I\bar{1}$-$C\bar{1}$ transition in plagioclase (scaled to kT/J). The J value was calculated on the assumption that the transition in pure anorthite occurs at 2315 K [extrapolation of the Carpenter and McConnell data by Holland and Powell (1992)]. Other symbols are the same as in Figure 11.

the three parameters x, σ, and Q_{od}. If one follows the construction schemes of Figures 3, 5, 6, 7, and 9, it is possible to derive the expression for the total entropy of the feldspar lattice as a function of the same set of parameters. The details of this evaluation are omitted here.

The non-equilibrium configurational free energy of the Al_xSi_{1-x} solid solution in the feldspar lattice can be calculated by substituting Equations 4 and 26 into Equation 3. By minimizing f with respect to σ and Q_{od}, it is possible to study the equilibrium behavior of the system at any x = Al/(Al+Si) and kT/J.

Here we consider the Al_xSi_{1-x} solid solution in the range of $0 < x < 0.5$, which is wider than the actual range of plagioclase compositions ($0.25 < x < 0.5$). Calculations of the equilibrium configurational entropy were made at compositional intervals of $\Delta x = 0.01$, and the results are shown as the solid lines in Figure 11. The entropy of the Al_xSi_{1-x} solid solution (S/R) normalized to one lattice point is plotted at three different temperatures kT/J = 0.6, kT/J = 0.5 and kT/J = 0. The kinks in the curves correspond to an order/disorder transition, which one can associate with the $I\bar{1}$-$C\bar{1}$ transformation. The black dots in Figure 11 are equilibrium entropies calculated using the Monte Carlo method (Myers et al. 1998). The comparison between the CVM and the Monte Carlo calculations is discussed below.

In Figure 12, the onset of LRO at the $I\bar{1}$-$C\bar{1}$ phase transition, as determined from the CVM model, is plotted as a function of temperature and composition, again shown as the solid line. The model predicts the phase transition of the anorthite composition (x = 0.5) at kT/J = 0.675. The phase transition can be observed only at compositions that approach the stoichiometry of x = 0.5 and a certain critical value of the composition x exists below which LRO cannot be stabilized at any temperature. Myers et al. (1998) estimated this value in the feldspar lattice as x = 0.3 using the Monte Carlo method. According to the present CVM approximation, this value lies between x = 0.28 and x = 0.29 (Fig. 12).

DISCUSSION

The superposition approximation

The present study shows that the CVM models for various Ising lattices can be constructed using the same principle, based on the sequential application of the superposition approximation. This approximation also plays an important role in the annealing algorithm. There is a certain analogy between the annealing algorithm and the Natural Iteration (NI) method of Kikuchi (1974), which is commonly used in the CVM as a tool for minimizing free energy of the cluster models. Like the NI method, the annealing algorithm is based on iterative application of the superposition approximation, which is understood as a relation between probabilities of clusters and their subclusters. In the NI method, the superposition approximation is derived from the condition of the free energy minimum of the model, and the cluster distributions are constrained to converge to the condition of the free energy minimum. The present study applies a similar iteration scheme (annealing algorithm), not to the whole model but to a single basic cluster, and constrains its distribution to converge to the maximum entropy at a given pair distribution. The derivation of the superposition approximation in the present study does not require differentiation of the free energy expression, and that simplifies the mathematics significantly. In fact, the superposition approximation (as it is understood in the present study) is a consequence of the Bayes formula for a complex stochastic event. The nature of the conversion is also dif-

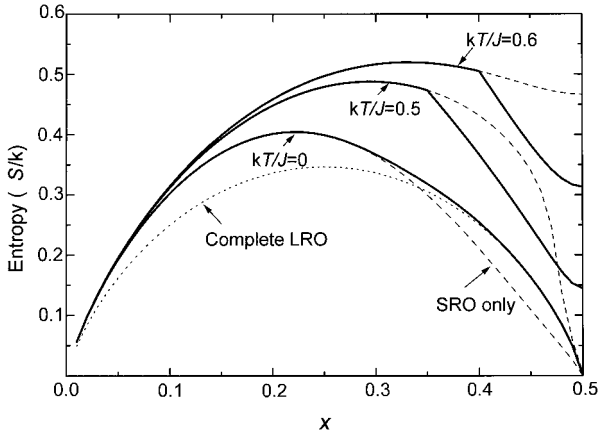


FIGURE 13. Equilibrium and non-equilibrium configurational entropy of Al-Si in the feldspar lattice as predicted by the present CVM model. The extrapolations into the non-equilibrium regions (dashed lines) have been made by fixing the LRO parameter (Q_{od}) at the zero value and minimizing the free energy of the CVM model only with respect to the SRO parameter (σ). The dotted line shows the entropy at complete LRO, defined by allowing Al atoms to occupy only one sublattice.

ferent. The annealing algorithm represents a stochastic process that converges to a steady state in a way similar to a Markov chain.

Comparison with Monte Carlo

In Figures 11 and 12, the present results are compared with the recent results of Monte Carlo simulations by Myers et al. (1998) performed using the same assumption that only nearest neighbors interact. The results of the two approaches are in excellent agreement, demonstrating that the clusters chosen to describe the feldspar structure and the methods described above to determine the configurational entropy are essentially correct. The slight difference in the calculated critical temperatures and compositions can be attributed to the internal limitation of the CVM method. The approximation of choosing clusters of finite size to represent an infinite lattice becomes less accurate in the vicinity of the phase transition (Sanchez et al. 1982). However, at temperatures or compositions only a few percent away from the phase transition the results of both methods almost coincide. This near coincidence has two important implications.

First, it is clear that in many applications the CVM can produce a result where accuracy is similar to the equilibrium Monte Carlo simulations. Being analytical, the CVM model can be used to study non-equilibrium states of ordering that are not easily achieved by the direct MC method. An example of the calculation of the non-equilibrium entropy constrained by the condition $Q_{od} = 0$ is shown in Figure 13 (dashed lines). The system is allowed to adjust to the change of temperature and x but only by varying its SRO parameter. Using such calculations, it is possible to investigate the thermodynamic behavior of a system where LRO cannot develop for kinetic reasons. Such a situation is illustrated by the metastable crystallisation of synthetic anorthite in the disordered $C\bar{1}$ form within the stability field of the ordered $I\bar{1}$ form (Kroll and Müller 1980; Carpenter 1991).

The second implication is that the CVM can be used in combination with the Monte Carlo method. In fact, using the CVM allows the extraction of values of configurational entropy from equilibrium Monte Carlo simulations. To calculate entropy, it is necessary to know the probability distribution of certain clusters at the particular kT/J value. These cluster distributions can be calculated easily from Monte Carlo simulated configurations by counting the frequencies of different cluster configurations. In the same way the CVM provides a means for extracting entropies from the inverse Monte Carlo simulations (Dove and Heine 1996). Using the inverse Monte Carlo method, one can simulate a non-equilibrium distribution of atoms that corresponds to particular spectroscopic characteristics of the atomic distribution in a sample rather than to a particular temperature. Recently, the inverse Monte Carlo method was applied to the simulation of the distribution of Al and Si in cordierite, leucite, and analcime consistent with the available ^{29}Si NMR data (Dove and Heine 1996). The CVM formulae for cordierite and leucite derived in the present study could be directly used for such calculations.

The CVM also allows extraction of thermodynamic information from NMR data in a more direct way. Such calculations for layer silicates and cordierites have been described by Vinograd and Putnis (1998) and Putnis and Vinograd (1999). The basis for the calculation resides in the theoretical evaluation of the probability distribution of star-like clusters as functions of the order parameters. From these star-distributions, one can calculate the theoretical NMR spectrum. The SRO and LRO parameters are then determined by comparing the theoretical and experimental spectra. The optimal set of the order parameters is used to constrain a high-order CVM entropy equation. This method for evaluating the Al, Si entropies avoids the prediction of negative values for Al-rich phases, which can occur when using the star approximation.

Comparison with experiment

The order/disorder transition that can be simulated in the present model corresponds to the $I\bar{1}$-$C\bar{1}$ phase transition, which has been observed experimentally in plagioclase feldspars at anorthite-rich compositions (Carpenter and McConnell 1984). However, according to the data of Carpenter and McConnell (1984), the transition is characterized by a steeper slope in T-x space. We have plotted the data of Carpenter and McConnell (1984) in Figure 12 as the dashed line, scaled to a transition temperature in anorthite of 2315 K.

The disagreement between the calculations and the experimental data probably can be reduced by taking into account the interactions between higher-order Al and Si neighbors. When second- and third-neighbor interactions are allowed, tendencies to other conflicting ordering

schemes will develop. For example, the tendency to avoid Al-O-Si-O-Al contacts (Dempsey rule) is in conflict with the anorthite ordering scheme that makes these contacts favorable. These conflicting ordering schemes can occur at Al/(Al+Si) ratios < 0.5. At the 0.5 ratio, the regular alternation scheme of anorthite has no alternatives. However, as the Al/(Al+Si) ratio decreases, the alternative ordering schemes begin to compete for the most effective decrease of the free energy. This competition is likely to shift the slope of the computed $I\bar{1} \leftrightarrow C\bar{1}$ transition line in the correct sense. We are currently investigating this hypothesis.

Acknowledgments

This work was supported by a grant from the DFG (Deutsche Forschungsgemeinschaft, Kr 768-13, Pu 153-4). We thank M.T. Dove, R.J. Kirkpatrick, and an anonymous reviewer for constructive comments. The visualization of the clusters needed to describe the aluminosilicate structures was made considerably easier using the program CrystalMaker (Palmer 1994).

References cited

Alexandrowicz, Z. (1971) Stochastic models for the statistical description of lattice systems. Journal of Chemical Physics, 55, 2765–2779.

Barker, J.A. (1952) Methods of approximation in the theory of regular mixtures. Proceedings of the Royal Society, London, A216, 45–56.

Carpenter, M.A. (1991) Mechanisms and kinetics of Al-Si ordering in anorthite: I. Incommensurate structure and domain coarsening. American Mineralogist, 76, 1110–1119.

———(1992) Equilibrium thermodynamics of Al/Si ordering in anorthite. Physics and Chemistry of Minerals, 19, 1–24.

Carpenter, M.A. and McConnell, J.D.C. (1984) Experimental delineation of the $C\bar{1} \leftrightarrow I\bar{1}$ transformation in intermediate plagioclase. American Mineralogist, 69, 112–121.

de Fontaine, D. (1979) Configurational thermodynamics of solid solutions. Solid State Physics, 34, 73–274.

———(1994) Cluster approach to order disorder transformations in alloys. Solid State Physics, 47, 33–176.

Dove, M.T. and Heine, V. (1996) The use of Monte Carlo methods to determine the distribution of Al and Si cations in framework aluminosilicates from ^{29}Si MAS NMR data. American Mineralogist, 81, 39–44.

Dove, M.T., Thayaparam, S., Heine, V., and Hammonds, K. (1996) The phenomenon of low Al-Si ordering temperatures in aluminosilicate framework structures. American Mineralogist, 81, 349–362.

Herrero, C. P. (1991) Short-range order of the Si, Al distribution on the faujasite framework. Journal of Physical Chemistry, 95, 3282–3288.

———(1993) Monte Carlo simulations of the Si, Al distribution in A-type zeolites. Journal of Physical Chemistry, 97, 3238–3343.

Herrero, C.P., and Sanz, J. (1991) Short-range order of the Si, Al distribution in layer silicates. Journal of Physics and Chemistry of Solids, 52, 1129–1135.

Hijmans, J. and De Boer, J. (1955) An approximate method for order-disorder problems I. Physica, 21, 471–484.

Holland, T.J.B. and Powell, R. (1992) Plagioclase feldspars: activity-composition relations based upon Darken's quadratic formalism and Landau theory. American Mineralogist, 77, 53–61.

Inden, G. and Pitsch, W. (1991) Atomic ordering. In R.W. Cahn, P. Haasen, and E.J. Kramer, Eds., Materials Science and Technology, Phase Transformations in Materials, 5, 497–549. VCH, Weiheim, New York, Basel, Cambridge.

Ising, E. (1925) Beitrag zur Theorie des Ferromagnetismus. Zeitschrift für Physik, 31, 253–258.

Kikuchi, R.A. (1951) Theory of cooperative phenomena. Physical Review, 81, 988–1003.

———(1974) Superposition approximation and the natural iteration calculation in cluster-variation method. Journal of Chemical Physics, 60, 1071–1080.

Kikuchi, R.A. and Brush, S.G. (1967) Improvement of the cluster-variation method. Journal of Chemical Physics, 47, 195–203.

Kroll, H. and Müller, W.F. (1980) X-ray and electron-optical investigation of synthetic high-temperature plagioclases. Physics and Chemistry of Minerals, 5, 255–277.

Meirovitch, H. (1977) Calculation of entropy with computer simulation methods. Chemical Physics Letters, 45, 389–392.

Myers, E., Heine, V., and Dove, M.T. (1998) Some consequences of Al/Al avoidance in the ordering of Al/Si tetrahedral framework structures. Physics and Chemistry of Minerals, 25, 457–464.

Onsager, L. (1944) Crystal statistics. I. A two-dimensional model with an order-disorder transition. Physical Review, 65, 117–149.

Palmer, D. (1994) CrystalMaker: Interactive crystallography for Macintosh and Power Macintosh. Lynxvale, Cambridge, U.K.

Phillips, B.L. and Kirkpatrick, J.R. (1994) Short-range Al-Si order in leucite and analcime: determination of the configurational entropy from ^{27}Al and variable-temperature ^{29}Si NMR spectroscopy of leucite, its Cs- and Rb- exchanged derivatives, and analcime. American Mineralogist, 79, 1025–1031.

Phillips, B.L., Kirkpatrick, R.J., and Carpenter, M.A. (1992) Investigation of short-range Al, Si order in synthetic anorthite by ^{29}Si MAS NMR spectroscopy. American Mineralogist, 77, 484–494.

Putnis, A. (1992) Introduction to Mineral Sciences, 457 p. Cambridge University Press.

Putnis, A. and Vinograd, V.L. (1999) Principles of solid state NMR spectroscopy and applications to determining local order in minerals. In K. Wright and C.R.A. Catlow, Eds., Microscopic Properties and Processes in Minerals Series. NATO ASI Reidel, (in press).

Putnis, A., Fyfe, C.A., and Gobbi, G.C. (1985) Al, Si ordering in cordierite using ''Magic Angle Spinning'' NMR. I: ^{29}Si spectra of synthetic cordierites. Physics and Chemistry of Minerals, 12, 211–216.

Putnis, A., Salje, E., Redfern, S., Fyfe, C.A., and Strobl, H. (1987) Structural states of Mg-cordierite I: order parameters from synchrotron X-ray and NMR data. Physics and Chemistry of Minerals, 14, 446–454.

Ross, C.R., II (1991) Ising model and geological applications. In J. Ganguly, Ed., Advances in Physical Geochemistry, vol 8, Diffusion atomic ordering, and mass transport, p. 51–90. Springer-Verlag, New York.

Salje, E., Kuscholke, B., Wruck, B., and Kroll, H. (1985) Thermodynamics of sodium feldspar II: experimental results and numerical calculations. Physics and Chemistry of Minerals, 12, 99–107.

Sanchez, J.M. and de Fontaine, D. (1978) The fcc Ising model in the cluster variation approximation. Physical Review B, 17, 2926–2936.

Sanchez, J.M., de Fontaine, D., and Teiter, W. (1982) Comparison of approximate methods for the study of antiferromagnetism in the fcc lattice. Physical Review B, 26, 1465–1468.

Sanchez, J.M., Ducastelle, F., and Gratias, D. (1984) Generalized cluster description of multicomponent systems. Physica, 128A, 334–350.

Thayaparam, S., Dove, M.T., and Heine, V. (1994) A computer simulation study of Al/Si ordering in gehlenite and the paradox of low transition temperature. Physics and Chemistry of Minerals, 21, 110–116.

Thayaparam, S., Heine, V., Dove, M.T., and Hammonds, K. (1996) A computational study of Al/Si ordering in cordierite. Physics and Chemistry of Minerals, 23, 127–139.

Vinograd, V.L. and Putnis, A. (1998) Calculation of the configurational entropy of Al, Si in layer silicates using the cluster variation method. Physics and Chemistry of Minerals, (in press).

Vinograd, V.L., Saxena, S.K., and Putnis, A. (1997) Calculation of the probability distribution of basic clusters involved in cluster variation approximations to the Ising model. Physical Review B, 56, 11493–11502.

Manuscript received May 14, 1998
Manuscript accepted October 15, 1998
Paper handled by James G. Blencoe