

Models for X-ray Scattering from Random Systems

BY J. GOODISMAN AND N. COPPA

Department of Chemistry, Syracuse University, Syracuse, New York 13210, USA

(Received 16 June 1980; accepted 13 September 1980)

Abstract

A new class of models is proposed for interpretation of the small-angle X-ray scattering from catalyst-like systems, which consist of a random arrangement of two or more phases, in each of which the electron density is uniform. By considering a grid formed of identical cells, each of which is filled with one phase or another, the calculation of the correlation function, and hence the scattering intensity, is reduced to a counting process and evaluation of $P_0(r)$, the probability that a line segment of length r lies wholly within one cell. Results for simple two-dimensional grids are considered in detail, with P_0 calculated by several methods. The Fourier transform of the correlation function is calculated for the cubic lattice, and properties of the scattering intensity curve discussed. Extensions of the model discussed here, the simplest of the cell models, are indicated.

Introduction

We consider a system, such as a supported metal catalyst, which consists of a random arrangement of two or more different materials, one of which may be void. The regions occupied by each material are of varying sizes and shapes. Within each region, the electron density is uniform. The important properties which characterize such a system are the volume fractions of the phases, and the interphase surface areas. The latter may be interpreted in terms of porosities, mean particle sizes, etc. It is well known that X-ray scattering can be used to measure surface areas in such systems (Thomas & Thomas, 1967; Somorjai, Powell, Montgomery & Jura, 1967; Satterfield, 1970; Sasvári, 1976). Gas adsorption measurements, the most commonly used method, are not ideal because they depend on assumptions about the adsorption process (form of isotherm, reversibility) and properties of the adsorbing species (size, packing, chemical interactions); they have been characterized as 'essentially empirical' (Solymosi, 1976). Other methods also have disadvantages and advantages (Thomas & Thomas, 1967, Chs. 3 and 4).

The small-angle scattering of X-rays gives a measurement of the total surface, and does not depend on properties of particular molecules (indeed, comparison of surfaces from X-ray measurement with surfaces from gas adsorption would furnish important information about the nature of the surfaces). In order to use the X-ray scattering, however, one requires a model of a random system of the kind described. Usually, the models used are based on random or regular packing of particles of a specified shape, perhaps of varying size (Thomas & Thomas, 1967; Satterfield, 1970; Sasvári, 1976). While this is convenient to work with (Guinier & Fournet, 1955), it does not seem an adequate representation of our catalyst-like systems (Méring & Tchoubar, 1968). A theory which avoids the use of particles was given by Debye, Anderson & Brumberger (1957) for a two-phase system and subsequently extended to multiphase systems (Goodisman & Brumberger, 1971). The meaning of 'randomness' in the context of such models is not always clear (Goodisman & Brumberger, 1979).

A knowledge of scattering intensity $I(h)$ for large h gives the surface area for a two-phase system (Guinier & Fournet, 1955). Here, $h = 2\pi \sin \theta / \lambda$ with λ the wavelength and θ half the scattering angle. When there are several phases, there are several surfaces, and more information is needed. For an isotropic system,

$$I(h) = I_e(h) \overline{\eta^2} V \int_0^\infty \gamma(r) \frac{\sin hr}{hr} 4\pi r^2 dr, \quad (1)$$

where $I_e(h)$ is the scattering of a free electron, $\overline{\eta^2}$ is the mean-square density fluctuation, V is the illuminated sample volume, and $\gamma(r)$ is the normalized correlation function (see below), which contains the desired information about the system. For example (Guinier & Fournet, 1955, p. 81),

$$\frac{1}{2} l_c = \int_0^\infty \gamma(r) dr \quad (2)$$

defines a 'distance of heterogeneity' l_c which is the mean diameter of an isolated particle, and, for a two-phase system of the kind we discuss, is such that a line of length L which starts in phase 1 will pass through regions of phase 1 for a length $L\phi_1 + (1 - \phi_1)l_c/2$. Here ϕ_1 is the volume fraction of phase 1.

To use the X-ray scattering to measure $\gamma(r)$, we derive the form of $\gamma(r)$ from a model involving several parameters, and find the values of these parameters by comparing the experimental scattering curve $I(h)$ with that derived from $\gamma(r)$ via (1). We are presenting a new class of models, based on random filling of the cells in a grid, for interpretation of X-ray scattering from multiphase random systems. The present paper discusses the basic ideas and presents calculations on the simplest of the models. Calculation of the correlation function requires calculation of a set of probabilities, which are obtained in models of the present type by counting of discrete cells. Predictions of the simple theory will be compared with predictions of the previously presented continuous model (Debye, Anderson & Brumberger, 1957; Goodisman & Brumberger, 1971), which proceeds from quite different assumptions.

The basic ideas of the cell models are discussed in §I. The information necessary to obtain the correlation function, from which X-ray scattering intensities can be derived, will be identified. In §II, the two-dimensional, two-phase system is illustrated with a specific example of the completely random system. In §III, the correlation function is obtained for this system, to show how the surface-to-volume ratios enter. Our goal in this work is of course the use of scattering intensities to obtain interphase surface areas for three-dimensional systems. The method of §III is applied to the three-dimensional two-phase system in §IV, and the results discussed. In §V, we discuss the generalizations of our formulas to multiphase systems and possibilities for future work on these models.

I. The cell models

The systems of interest to us are random assemblies of regions of different electron density, and the X-ray scattering yields information on the surfaces between the regions. The regions are of varying size, shape and orientation. The electron density is homogeneous within each region, so that the system is characterized by the electron densities and volume fractions for the different phases. As a whole, the system is isotropic and homogeneous.

In the cell models, we consider the space of the system to be divided into cells, all of the same size and shape. Each cell is filled with matter belonging to one of the phases (including void) making up the system. The way in which one decides which phase to place in a given cell will differ from model to model, but the fraction of cells filled with a particular phase is given by a predetermined volume fraction or composition. In the simplest model, which we discuss in the present paper, each cell is independent, and is assigned randomly to one phase or another. By filling cells according to some












recipe, one generates inclusions, for each phase, of varying size and shape.

For example, Fig. 1 shows a random two-phase, two-dimensional system generated by taking each cell in turn and deciding on the basis of a random event whether to fill it in or not. There are no correlations between cells. The predetermined volume fractions, $\frac{1}{4}$ and $\frac{3}{4}$, give the probabilities for filling that were used in making the decision for each cell. Some of the simpler figures generated are listed in Table 1.

The surface-to-volume ratio in two dimensions would be the length of boundary line between filled and unfilled squares, divided by the area of filled squares ($\frac{1}{4}$ the total area). We will consider in §II what this is expected to be for a given set of composition numbers ($\frac{1}{4}$ and $\frac{3}{4}$ in the present case). The surface-to-volume ratio could be obtained by summing up contributions of the different figures, but we will be able to give a general formula. It will also be interesting to consider predicted frequencies for each figure of Table 1, to compare to what was found experimentally. Study of the shapes generated in this simple case is helpful in developing a feeling for what occurs in more complicated systems.

The correlation function γ , in general, is defined as the normalized average, over orientations and positions of the vector \mathbf{r}_0 , of $\eta(\mathbf{r}_0) \eta(\mathbf{r}_0 + \mathbf{r})$, where $\eta(\mathbf{r})$ is the deviation of the electron density at \mathbf{r} from the average value. Clearly, γ depends on r alone for an isotropic system. The normalization of $\gamma(r)$ is such that $\gamma(0) = 1$.

Table 1. *The smaller configurations, with theoretical probabilities, for the two-dimensional square grid*

	n	s	Configuration		Orientations
<i>a</i>	1	4		0.07910	1
<i>b</i>	2	6		0.01112	2
<i>c</i>	3	7		0.00209	4
<i>d</i>	3	8		0.00156	2
<i>e</i>	4	8		0.00039	1
<i>f</i>	4	8		0.00039	?
<i>g</i>	4	8		0.00039	2
<i>h</i>	4	8		0.00039	4
<i>i</i>	4	9		0.00029	4
<i>j</i>	4	9		0.00029	4
<i>k</i>	4	10		0.00022	2

If $P_{ij}(r)$ represents the probability (averaged over positions and orientations) that a line segment AB of length r has end A in phase i and end B in phase j , one can easily show (Goodisman & Brumberger, 1971)

$$\gamma(r) = \frac{\sum_{ij} P_{ij} n_i n_j - \bar{n}^2}{\bar{n}^2 - \bar{n}^2}. \quad (3)$$

If the volume fraction or composition variable of phase i is represented by φ_i , the average electron density is

$$\bar{n} = \sum_i \varphi_i n_i \quad (4)$$

and the average squared electron density

$$\bar{n}^2 = \sum_i \varphi_i n_i^2. \quad (5)$$

In the simplest cell model, there is no correlation between different cells, so that $P_{ij}(r) = \varphi_i \varphi_j$ if the line segment spans two or more cells and $P_{ij}(r) = \varphi_i \delta_{ij}$ if the line segment lies wholly within one cell ($\delta_{ij} = 0$, $i \neq j$; $\delta_{ij} = 1$, $i = j$).

Let $P_0(r)$ be the probability that a line segment of length r does not cross a cell boundary, *i.e.* lies wholly within one cell. The probability P_{11} is given by $P_0 \varphi_1 + (1 - P_0) \varphi_1^2$; P_{12} is given by $(1 - P_0) \varphi_1 \varphi_2$, and so on. Thus, for the two-phase system with no correlations between cells,

$$\begin{aligned} \sum_{ij} P_{ij} n_i n_j &= P_0(\varphi_1 n_1^2 + \varphi_2 n_2^2) \\ &+ (1 - P_0)(\varphi_1^2 n_1^2 + \varphi_2^2 n_2^2 + 2\varphi_1 \varphi_2 n_1 n_2). \end{aligned}$$

Therefore the correlation function for the two-phase system is

$$\gamma(r) = \frac{P_0(\varphi_1 n_1^2 + \varphi_2 n_2^2) + (1 - P_0)\bar{n}^2 - \bar{n}^2}{\bar{n}^2 - \bar{n}^2} = P_0 \quad (6)$$

and our problem reduces to the determination of P_0 . This will be dealt with in §§ III and IV, after consideration of a specific example in two phases and two dimensions.

Consider now a random system involving more than two phases. The generalization of the above arguments is straightforward, and

$$\begin{aligned} \gamma(r) &= \left[P_0 \left(\sum_i \varphi_i n_i^2 \right) + (1 - P_0) \left(\sum_i \varphi_i^2 n_i^2 \right) \right. \\ &\quad \left. + \sum'_{ij} (1 - P_0) \varphi_i \varphi_j n_i n_j - \bar{n}^2 \right] \\ &\quad \times \left[\sum_i \varphi_i n_i^2 - \left(\sum_i \varphi_i n_i \right)^2 \right]^{-1}, \end{aligned}$$

where the prime on the summation means i and j must be different. We obtain

$$\gamma(r) = \frac{P_0 \left[\sum_i \varphi_i n_i^2 - \left(\sum_{ij} \varphi_i \varphi_j n_i n_j \right) \right]}{\sum_i \varphi_i n_i^2 - \left(\sum_i \varphi_i n_i \right)^2} = P_0 \quad (7)$$

independently of the number of phases. This implies that, for the uncorrelated random system, the shape of the scattering intensity curve $I(h)$ is independent of the number of phases. Its magnitude will change because of the factor of \bar{n}^2 in (1). For example, if one formed a random system of three phases by partly filling the void phase of a two-phase system with a third phase (Goodisman & Brumberger, 1979), the shape of the scattering would not change, assuming the filling was done in a totally random way.

It should be emphasized that the division of space into cells of identical size and orientation is a convenience for counting volumes and surface areas. It is not meant as a faithful representation of the system; averaging over orientations must be performed. However, it can be used to give some insight into the structure according to percolation theory (Shante & Kirkpatrick, 1971).

A two-phase (solid and void) system would be mechanically unstable unless a solid cluster of infinite size existed. The minimum fractional occupation number of solid which guarantees such a cluster is the percolation probability, which is 0.31 for a lattice of cubes (Powell, 1979). The actual system would not be made of cubic particles; it would be more realistic to use spheres on the cubic lattice. The actual volume fraction would then be $(0.31) (\pi l^3/6)/l^3 = 0.16$, a number which in fact varies little with the type of lattice used (Powell, 1979). Since the spheres of diameter l just touch, structural stability requires that larger ones be used, raising the volume fraction. On the other hand, for a catalyst one would like the largest void volume consistent with structural stability, to make the interior surfaces maximally accessible. Indeed, many catalysts seem to be produced with volume fractions of 18–20% for the solid phase.

In a more sophisticated model, we could consider placing solid spheres or other shapes within the cells, on a random basis, to produce a desired volume fraction. In the present paper, however, we use cubes, with the occupation probabilities chosen to equal the volume fractions, and average over orientations to produce a correlation function depending only on one length parameter. Surface areas are calculated for the figures formed from the cubes.

II. Example: Two dimensions, two phases

The two-dimensional grid shown in Fig. 1 was generated by randomly assigning each grid square a

color, black or white. The only constraint on the assignment was that the grid squares would be assigned white three times as often as black giving a three to one ratio white to black. This was achieved by selecting marbles one-by-one from a stirred lot consisting of thirty black and ninety white marbles. The selection of a black marble resulted in assigning the grid square a black color.

This random assignment of black and white colors to the grid resulted in the formation of variously shaped aggregates of neighboring black grid spaces. Specific polygons hereafter are referred to as configurations. Most configurations will have several orientations, each distinguished from one another by rotations in the plane of the grid. Superimposable rotations are considered as one orientation.

The probability of finding a polygon composed of n black grid squares and s neighboring white squares will be equal to the probability of finding a black grid ($\frac{1}{4}$) raised to the n th power times the probability of finding a white grid unit ($\frac{3}{4}$) raised to the s th power.

$$P_0^{(n,s)} = \left(\frac{1}{4}\right)^n \left(\frac{3}{4}\right)^s, \quad (8)$$

where $P_0^{(n,s)}$ is the theoretical probability of finding a polygon of n black grid squares neighboring s white squares. Only grid squares that share a side, and not squares situated diagonally to each other, are considered to be neighbors. Thus

$$P_0^{(1,4)} = \left(\frac{1}{4}\right) \left(\frac{3}{4}\right)^4 = 0.07910$$

is the probability for an isolated black square and

$$P_0^{(2,6)} = \left(\frac{1}{4}\right)^2 \left(\frac{3}{4}\right)^6 = 0.01112$$

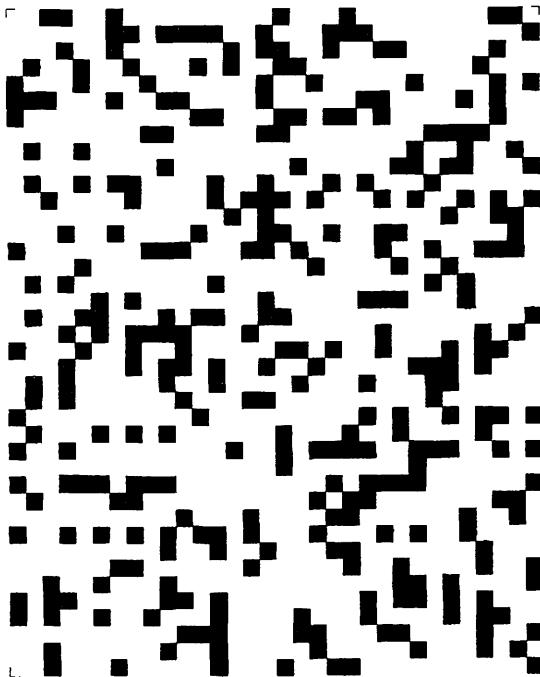


Fig. 1. Sample two-dimensional grid, formed by independent random coloring of squares.

for a two-square domino. For n greater than 2, different configurations exist, and for polygons made up of more than one grid square different orientations must be considered. Table 1 lists the various configurations, orientations and their respective theoretical probabilities. The theory of random clumping (Roach, 1968) gives a similar algorithm for calculation of the probabilities.

As we have defined it, $P_0^{(n,s)}$ is the probability that an oriented figure having indices n and s will be located at an arbitrarily chosen square on the grid. To make this definition complete, we define, for each figure, a 'key square' whose position gives the location of the figure. The key square of an orientation could be any square within the orientation so long as it is not changed throughout the treatment. Now the product of $P_0^{(n,s)}$ and the number of squares on the grid should give the number of times a given configuration and orientation is found on the grid.

For comparison with (8), the experimental probabilities $P^{(n,s)}$ were determined from the number of times a specific configuration and orientation occurred within the boundaries of the grid. Figures or apparent parts of figures lying on the grid border were given special consideration. If a configuration shares a side with the grid border, the neighboring grid squares outside the boundary may or may not meet the criteria for defining the shape. The probability that they do is multiplied by the number of occurrences. If part of a figure (including its key square) is on the grid, similar considerations are applied to the squares outside the border to give a probability to be multiplied by the number of occurrences. The total number of times a specific configuration and orientation occurs (including the contributions of the bordering possibilities), divided by the total number of grid squares, gives the experimental probability.

For example, consider the horizontal orientation of the (2,6) ('domino') configuration. Let the key square be the one on the left. In the grid of Fig. 1 there are nine non-bordering horizontal dominoes. There are three with a long edge along the boundaries of the grid; the probability that these cases fulfil the criteria for this configuration is $\left(\frac{3}{4}\right)^2$. On the right edge there are nine single black squares that each can possibly be part of a horizontal configuration of the type (2,6), with probability $\left(\frac{1}{4}\right) \left(\frac{3}{4}\right)^3$. Single squares on the left boundary cannot belong to a horizontal domino on the grid because the key square is outside. Finally there is one corner black square that is assigned the weight $\left(\frac{1}{4}\right) \left(\frac{3}{4}\right)^4$. The probability for horizontal (2,6) configurations found on the grid is thus

$$\frac{9 + 3\left(\frac{3}{4}\right)^2 + 9\left(\frac{1}{4}\right) \left(\frac{3}{4}\right)^3 + \left(\frac{1}{4}\right) \left(\frac{3}{4}\right)^4}{\text{number of squares}} \\ = \frac{11.71582}{1280} = 0.00915.$$

Table 2 lists a comparison of the experimental and theoretical probabilities.

For each configuration there is a definite perimeter-to-area ratio, equal to the sum of line segments about each configuration divided by the configuration area. The overall perimeter-to-area ratio could theoretically be calculated by summing up contributions of all the orientations of all the configurations (probabilities times perimeter-to-area ratios). The number of possible configurations as n increases becomes large rapidly, making this approach extremely tedious.

A general formula for the perimeter-to-area ratio may be derived, avoiding the difficulties of counting configurations. The probability that any line segment of the grid is part of the perimeter, *i.e.* that it separates a white grid square from a black grid square, is $2(\frac{1}{4})(\frac{3}{4})$. This probability times the total number of line segments in the grid will give the theoretical perimeter found on the entire grid. Thus, in general, the perimeter-to-area ratio is

$$\frac{P}{A} = \frac{4mnp(1-p)l}{mnl^2} = \frac{4(1-p)p}{l}, \quad (9)$$

where m and n are the overall dimensions of the grid, p = the probability a grid square is black, $(1-p)$ = the probability a grid square is white, and l = length of a

Table 2. Comparison between theoretically calculated and measured occurrence probabilities for shapes on a two-dimensional square grid

	Configuration	Theory	Experiment*
a		0.07910	0.0750
b		0.01112	0.0109
c		0.00209	0.0019
d		0.00156	0.0024
e		0.00039	(0.0000)
f		0.00039	(0.0008)
g		0.00039	(0.0000)
h		0.00029	(0.0007)
i		0.00029	(0.0007)
j		0.00029	(0.0000)
k		0.00022	0.0003

*These represent averages over orientations. Numbers in parentheses represent occurrences of less than 1 on the 1280-space grid, and hence are not statistically significant.

line segment. For the grid of Fig. 1 the ratio of perimeter to filled area was found to be 2.96; the perimeter-to-filled area ratio found using (9) is 3.00.

The extension of the calculations of this section to three dimensions is obvious. One can calculate occurrence probabilities for three-dimensional shapes in terms of the number of cubes making up a shape and the number of neighboring cubes. The perimeter and area become respectively surface and volume. Consider a three-dimensional grid made of cubes of length l with occupation probabilities p and $1-p$, with dimensions k , m and n . The theoretical number of squares that are part of the surface is $3kmn$ times $2p(1-p)$, so the surface-to-volume ratio is

$$\frac{S}{V} = \frac{6kmnp(1-p)l^2}{kmnl^3} = \frac{6(1-p)p}{l}. \quad (10)$$

The correlation function is essentially the probability of non-crossing (equations 6 and 7). We calculate it for a two-dimensional square grid in the next section, and for the three-dimensional cubic grid in § IV.

III. Crossing probabilities in two dimensions

Consider a square of edge length L , lying between O and L on the x and y axes. The probability that a line segment of length r , with one end (A) in the square, crosses out of the square, is obtained by calculating the probability when A is located at (x,y) and then averaging over all (x,y) within the square. It suffices to consider $x \leq L/2$, and crossing of the top of the square ($y = L$) only. If $p(r)$ is the averaged probability for a crossing of the top, $P_0(r) = 1 - 4p(r)$.

The probability for a particular location (x,y) and length is the fraction of orientations of the line segment which lead to a crossing. Let $u = L - y$. There are four cases:

- (1) If $u \geq r$, the probability $p(r)$ is zero for all x .
- (2) If $u^2 + x^2 \geq r^2 > u^2$ or $r^2 - x^2 \leq u^2 < r^2$, the probability is $\pi^{-1} \cos^{-1}(u/r)$.
- (3) If $(L-x)^2 + u^2 \geq r^2 > x^2 + u^2$ or $r^2 - (L-x)^2 \leq u^2 < r^2 - x^2$, the probability is $(2\pi)^{-1}[\cos^{-1}(u/r) + \cot^{-1}(u/x)]$.
- (4) If $r^2 > (L-x)^2 + u^2$ or $u^2 < r^2 - (L-x)^2$, the probability is $(2\pi)^{-1}[\cot^{-1}(u/x) + \cot^{-1}(u/L-x)]$.

The different regions of r and x we must consider are shown in Fig. 2. The detailed calculation is given in the Appendix, with the result in Fig. 3. The same method of calculation can be applied to the triangle (unpublished calculations) and the result (shown in Fig. 3) used to discuss a model for a random n -phase system based on a lattice of equilateral triangles.

There exists another method of calculating P_0 for a square, which is applicable to the cube as well. Considering a square grid formed by equally spaced

verticals and horizontals, we ask for the number of crossings made by a line segment of length r oriented at an angle of θ to the vertical ($0 \leq \theta < \pi$). Along the stick direction, the horizontals have a spacing of $L/|\cos \theta|$, so the average number of horizontals intercepted is $r|\cos \theta|/L$. If this is greater than 1, the stick must cross. If it is less than 1, the probability of the line segment intercepting a horizontal is $r|\cos \theta|/L$. For the verticals, the same statements can be made, with $\sin \theta$

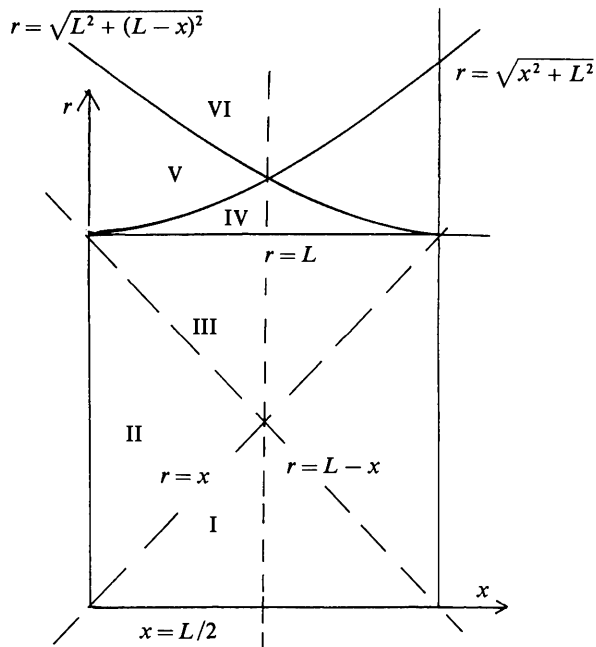


Fig. 2. Different regions of r and x for calculation of intersect probability for a square.

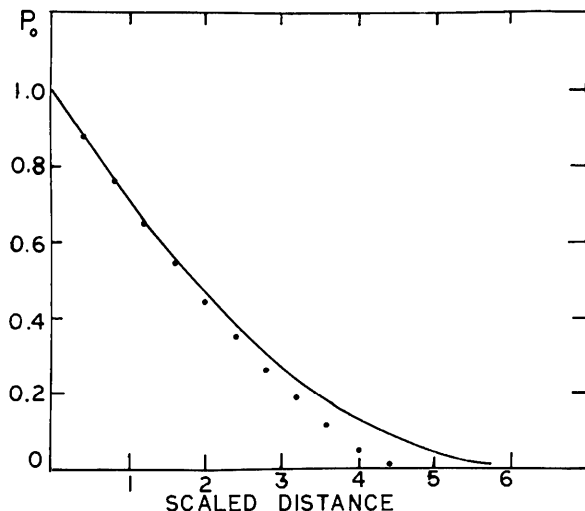


Fig. 3. Correlation function for a square grid (non-crossing probability) plotted (dots) against the scaled distance $4r/l$. Solid line shows the corresponding quantity for an equilateral triangle, plotted against the scaled distance $4r\sqrt{3}/l$.

substituted for $|\cos \theta|$. In order for the line segment to cross neither a vertical nor a horizontal, we require $r|\cos \theta| < L$ and $r \sin \theta < L$. If both conditions are fulfilled,

$$\bar{P}_0 = \left(1 - \frac{r|\cos \theta|}{L}\right) \left(1 - \frac{r \sin \theta}{L}\right). \quad (11)$$

We have to average \bar{P}_0 over θ , considering only the allowed angles, to get P_0 . The allowed angles are those for which L/r exceeds the larger of $\sin \theta$ and $|\cos \theta|$.

For $L/r < 1/\sqrt{2}$ no angles are allowed, and the probability P_0 of no crossing is zero. For $L/r > 1$, all angles θ are allowed, and the probability P_0 is

$$\begin{aligned} \frac{1}{\pi} \int_0^\pi d\theta \left(1 - \frac{r|\cos \theta|}{L}\right) \left(1 - \frac{r \sin \theta}{L}\right) \\ = \frac{2}{\pi} \left(\frac{\pi}{2} - \frac{2r}{L} + \frac{r^2}{2L^2}\right). \end{aligned}$$

For $1 \geq L/r \geq 1/\sqrt{2}$, we get

$$\begin{aligned} P_0 = \frac{1}{\pi} \left[2 \int_{\cos^{-1}(L/r)}^{\pi/4} d\theta \left(1 - \frac{r \cos \theta}{L}\right) \left(1 - \frac{r \sin \theta}{L}\right) \right. \\ \left. + 2 \int_{\sin^{-1}(L/r)}^{\pi/4} d\theta \left(1 - \frac{r \cos \theta}{L}\right) \left(1 - \frac{r \sin \theta}{L}\right) \right] \\ = \frac{2}{\pi} \left[\sin^{-1}\left(\frac{L}{r}\right) - \cos^{-1}\left(\frac{L}{r}\right) + 2 \left(\frac{r^2 - L^2}{L^2}\right)^{1/2} \right. \\ \left. - 1 - \frac{r^2}{2L^2} \right] \end{aligned}$$

which checks our previous result. The correlation function $\gamma = P_0$ is shown in Fig. 3.

IV. Correlation function and scattering for cubic lattice

We may now consider the correlation function for a randomly filled lattice of cubes of length l . To calculate $P_0(r)$, the probability that a line segment of length r does not cross a cube boundary, we use the second of the methods given for squares. In three dimensions, we have that the probability of no crossing is 0 whenever r exceeds any of $l/|\cos \theta_1|$, $l/|\cos \theta_2|$, or $l/|\cos \theta_3|$, where θ_1 , θ_2 , and θ_3 are the angles made by the direction of r to the three perpendicular axes defining the series of planes which delineate the lattice. If r is less than all three quantities, the probability of no crossing is

$$\bar{P}_0 = \left(1 - \frac{r|\cos \theta_1|}{l}\right) \left(1 - \frac{r|\cos \theta_2|}{l}\right) \left(1 - \frac{r|\cos \theta_3|}{l}\right). \quad (12)$$

We then average \bar{P}_0 over angles to get P_0 ; it suffices to consider the first octant: $0 \leq \theta < \pi/2$ and $0 \leq \varphi \leq \pi/2$.

The details of the calculation are given elsewhere (Goodisman, 1980). It should be noted that Stokes & Wilson (1942) have calculated scattering intensities for rectangular parallelepipeds and other shapes; $P_0(r)$ can be obtained from the quantity V_x in their paper by averaging over all directions. Méring & Tchoubar (1968) have given P_0 for a cube for $r < l$. Miller & Schmidt (1962; Schmidt, 1965; Miller, 1961) gave asymptotic (large h) expansions for the scattering intensity of collections of randomly oriented right cylinders of arbitrary cross section, which includes prisms. The intensity for cubes is checked by our formulas (14) and (15) below.

We find a non-crossing probability $P_0(r)$ which decreases monotonically from a value of unity at $r = 1$ to a value of zero at $r = l\sqrt{3}$ (length of cube diagonal). $P_0(r)$ and its first derivative are continuous functions, but the second derivative of $P_0(r)$ has large discontinuities at $r = l$ and $r = l\sqrt{2}$ (cube face diagonal). Explicitly,

$$\begin{aligned}
 P_0(r) &= 1 - \frac{3r}{2l} + \frac{2r^2}{\pi l^2} - \frac{1}{4\pi} \frac{r^3}{l^3}, \quad 0 \leq r < l \\
 P_0(r) &= -2 + \left(\frac{3}{2} - \frac{1}{4\pi} \right) \frac{l}{r} + \frac{3r}{2\pi l} + \frac{6r}{\pi l} \cos^{-1} \frac{l}{r} \\
 &\quad - \left(\frac{2}{\pi} + \frac{4r^2}{\pi l^2} \right) \left(1 - \frac{l^2}{r^2} \right)^{1/2} \\
 &\quad + \frac{r^3}{2\pi l^3}, \quad l \leq r < l\sqrt{2} \\
 P_0(r) &= \frac{2}{\pi} \left\{ \frac{l}{r} \left[\frac{3\pi}{4} + \cos^{-1} \frac{l}{(r^2 - l^2)^{1/2}} - \frac{5}{8} \right] \right. \\
 &\quad + \frac{3\pi r}{4l} - \frac{2\pi}{3} + \tan^{-1} \left[\frac{r(r^2 - 2l^2)^{1/2}}{l^2} \right] \\
 &\quad - \frac{2l}{r} \tan^{-1} \left[\frac{2l(r^2 - 2l^2)^{1/2}}{3l^2 - r^2} \right] \\
 &\quad - \frac{3r}{l} \cos^{-1} \frac{l}{(r^2 - l^2)^{1/2}} \\
 &\quad + \left(1 + \frac{r^2}{l^2} \right) \left(1 - \frac{2l^2}{r^2} \right)^{1/2} - \frac{1}{8} \frac{r^3}{l^3} \\
 &\quad \left. - \frac{3r}{4l} - \tan^{-1} \left(\frac{a + \sqrt{3}b}{b - \sqrt{3}a} \right) \right\}, \\
 &\quad l\sqrt{2} \leq r < l\sqrt{3},
 \end{aligned}$$

where $a^2 = 4l^4(r^2 - 2l^2)$ and $b = r^3 - 2l^2r - l^4/r$. This function is shown in Fig. 4. A plot of the second derivative, which is proportional to the intersect distribution for a cube, shows some interesting features (Goodisman, 1980).

In order to determine the scattering intensity, we need the Fourier transform of $\gamma(r)$ or $P_0(r)$:

$$\begin{aligned}
 &\int_0^{\infty} r^2 P_0 \frac{\sin hr}{hr} dr \\
 &= \frac{1}{h} \left[\int_0^l \left(1 - \frac{3r}{2l} + \frac{2r^2}{\pi l^2} - \frac{1}{4\pi} \frac{r^3}{l^3} \right) (\sin hr)r dr \right. \\
 &\quad \left. + \int_l^{l\sqrt{2}} \gamma(r) (\sin hr)r dr + \int_{l\sqrt{2}}^{l\sqrt{3}} \gamma(r) (\sin hr)r dr \right]. \quad (13)
 \end{aligned}$$

Here, we can perform the integration analytically from $r = 0$ to $r = l$; in each of the other ranges of r , we approximate the correlation function by fitting $r\gamma(r)$ to a power series in r , so that the integration can be done analytically as well. It is important to ensure that the values and slopes of γ at the ends of each range are correct, since spurious discontinuities in $\gamma(r)$ or $\gamma'(r)$ lead to terms in the Fourier transform behaving as h^{-2} and h^{-3} . The result for (13) is as follows (different polynomials were tried to check accuracy of the formula):

$$\int_0^{\infty} r^2 P_0 \frac{\sin hr}{hr} dr = F(hl)/h^3 \quad (14)$$

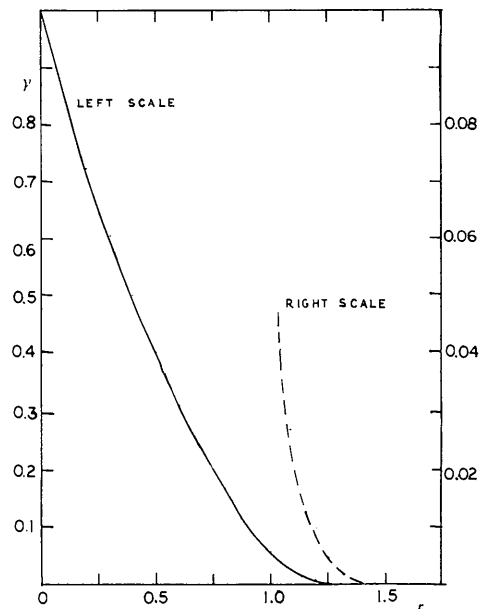


Fig. 4. Correlation function (non-crossing probability) for a cube, plotted against distance in units of edge length.

where

$$\begin{aligned}
 F(x) = & \frac{3}{x} - \frac{1.9099}{x^3} - \left(\frac{2.2920}{x} - \frac{18.2051}{x^3} \right) \cos x \\
 & - \frac{9.9655}{x^4} \sin x \\
 & + \left(\frac{0.1576}{x} - \frac{16.2952}{x^3} \right) \cos(x\sqrt{2}) \\
 & + \frac{1.0367}{x^4} \sin(x\sqrt{2}) - \frac{0.0253}{x} \cos(x\sqrt{3}) \\
 & + \frac{0.2692}{x^2} \sin(x\sqrt{3}). \tag{15}
 \end{aligned}$$

This compares well with formulas of Miller & Schmidt (1962; Miller, 1961) for large x . The function $F(x)$, plotted in Fig. 5, is everywhere positive, as the intensity must be. For x above 25, $F(x)$ is essentially $(3 - 2.292 \cos x)/x$. $F(x)$ approaches zero smoothly for $x \rightarrow 0$, the functional form of (14) and (15) becoming invalid for x less than about 0.5. For $h \rightarrow 0$, the integral (14) is just $\int r^2 P_0 dr = 0.7958l^3$, so that $F(x)$ should become proportional to x^3 for $x \rightarrow 0$.

For the purposes of making comparisons to experimental results, it is sometimes convenient to use the moments of the intensities $I(h)$, and consider dimensionless ratios of these moments. The j th moment of the intensity for point collimation would be

$$M_j = \int_0^\infty I(h)h^j dh = C \int_0^\infty F(h)h^{-3} h^j dh,$$

where C is the same for all j . Moments higher than the second do not converge because of terms in x^{-1} in $F(x)$ for large x ; M_j for $j > 0$ also does not converge. The 0th, 1st and 2nd moments are, respectively, $0.240282Cl^2$, $0.45338Cl$, and $1.57654C$. The dimensionless ratio M_0M_2/M_1^2 is thus 1.84288. The other available random theory (Debye *et al.*, 1957)

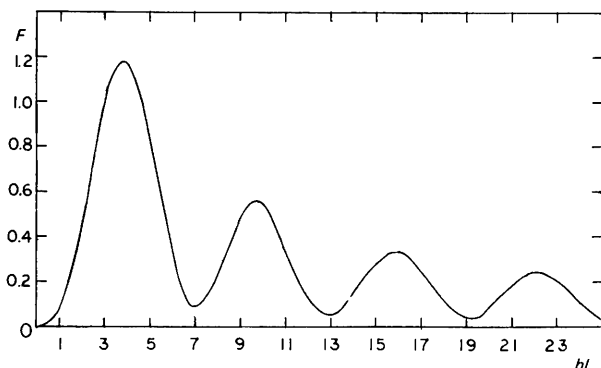


Fig. 5. The scattering intensity for our systems is proportional to $F(h)l/h^3$, where l is the cell (cube) edge length and $h = 4\pi \sin \theta/\lambda$.

predicts an exponential correlation function and a scattering intensity proportional to $(1 + h^2 a^2)^{-2}$. Thus, within a common multiplying constant the 0th, 1st and 2nd moments are, respectively, $\pi/4a$, $1/2a^2$, and $\pi/4a^3$, so that the dimensionless ratio M_0M_2/M_1^2 is $\pi^2/4 = 2.4674$.

It may be noted that, if $J(h)$ is defined as the Fourier transform of γ or P_0 according to (13), the second moment of $J(h)$ is exactly $\pi/2$ times $P_0(0)$, and usually $P_0(0) = 1$. The proof is as follows:

$$\begin{aligned}
 J(h) &= \int_0^\infty P_0(r) \frac{\sin hr}{hr} r^2 dr \\
 &= \int P_0(r) \frac{e^{ih \cdot r}}{4\pi} dr. \tag{16}
 \end{aligned}$$

Inverting the Fourier transform, we have

$$\begin{aligned}
 P_0(r) &= 4\pi(2\pi)^{-3} \int J(h) e^{-ih \cdot r} dh \\
 &= \frac{2}{\pi} \int_0^\infty h^2 dh \frac{\sin hr}{hr} J(h) \tag{17}
 \end{aligned}$$

and one recalls that $\sin x/x$ approaches unity as x approaches zero. The difference between 1.57654 and $\frac{1}{2}\pi = 1.5708$ is a measure of the accuracy of our numerical methods. For the other theory, the correlation function $e^{-r/a}$ gives $J(h) = 2a^3(1 + h^2 a^2)^{-2}$ (Debye *et al.*, 1957).

By integrating (17), we obtain

$$\begin{aligned}
 \pi \int_0^\infty P_0(r) dr &= 2 \int_0^\infty h^2 dh J(h) \int_0^\infty dr \frac{\sin hr}{hr} \\
 &= 2 \int_0^\infty h dh J(h) \frac{\pi}{2}
 \end{aligned}$$

so that the first moment of h is actually equal to the integral of the correlation function over all r . This also serves as a check on the accuracy of the numerical techniques we have employed.

V. Discussion

The theory presented in this work is the simplest in a class of models for multiphase systems, in which the correlation function is calculated by consideration of a grid of cells, each filled with a single phase. In this simplest model, all cells are of the same size, and each cell's contents is determined independently of the others', with no correlation between cells. Whether this is a reasonable assumption for a particular system can only be decided by comparison of the predictions of the model with experimental X-ray scattering intensities. Consideration of the moments of the intensity is one way to do this.

For the simplest model, we have discussed only the two-phase case in detail. Of course, if one is interested solely in the surface-to-volume ratio for a two-phase system, one piece of experimental data suffices, without recourse to any theory. Assuming one knows all the quantities appearing in (1) except for the correlation function, the limit of $h^4 I(h)$ for large h gives the surface-to-volume ratio or specific surface, regardless of the model. In the present case, taking the limit involves averaging over oscillating terms; this gives (see equations 1, 14 and 15) $4\pi I_e(h)\eta^2 V(3/l)$ which (Guinier & Fournet, 1955, p. 80) is equal to $2\pi I_e(h)(n_1 - n_2)^2 S$, so $S/V = 6p(1 - p)/l$ in agreement with (10). The advantage of having a model, of course, is the additional information given about the surfaces. If it is verified that the model applies to a given system, considerations like those of § II give an idea of the arrangement of the two phases and how the specific surface is distributed.

For a three-phase system, at least three pieces of experimental data (except for V and η^2) are needed. The simplest model predicts that the shape of the $I(h)$ curve will be the same as for the two-phase system. If indeed this model applies, it is not difficult to predict the three specific surfaces. Let the volume fractions be ϕ_1 , ϕ_2 and ϕ_3 ; and let the grid contain kmn cells of volume l^3 and $3kmn$ boundary faces. The probability of any one boundary face contributing to the surface between phases i and j is $2\phi_i\phi_j$, so the ratio of the i - j surface area to the total volume is $6\phi_i\phi_j/l$. It is irrelevant whether the three-phase system is formed by mixing the three components or by replacing one phase of a two-phase system by a mixture of two phases. Only the final volume fractions count. This is certainly not true if there are correlations between cells. Interestingly, the situation in which the surfaces S_{ij} are strictly in the ratios of the products of volume fractions, $\phi_i\phi_j$, constitutes a degenerate case in the previously presented multiphase theory (Goodisman & Brumberger, 1971).

In comparing the cell theory with the previous one (Debye, Anderson & Brumberger, 1957) we may note that each has certain features which are physically unreal. The Debye-Anderson-Brumberger theory allows for inclusions of any size, with no minimum. On the other hand, it is truly isotropic, whereas the cell theory must be made so by averaging over orientation angles. All the configurations are constructed of cubes, which possibly introduces special problems because of the corners. However, use of a cubic or other space-filling form reduces all the computations to counting of discrete units, a great simplification.

We believe that the great advantage of the cell models lies in their physical visualizability, which has been invoked in the paragraphs above. It is possible to show the specific shapes that contribute to the surface and their occurrence probabilities. The meaning of

randomness is simply the lack of correlations between the contents of different cells. The meaning of randomness is subtle in the Debye-Anderson-Brumberger theory (Goodisman & Brumberger, 1979). In generating the differential equation for the P_{ij} , it is assumed that the chance of finding an end of a line segment at the ij surface is proportional to the ratio of the area of that surface to the total volume, regardless of the location of the other end. How to modify this assumption to introduce nonrandomness is not clear. For a three-phase system, the volume fractions do not determine the surface areas, whereas they do in our uncorrelated theory; is this theory, then, more random?

Introduction of correlations between cells may be used to introduce physically reasonable features. One might expect configurations of high surface-to-volume ratio to occur less often than predicted by random statistics, perhaps ascribing this 'clumping' to the influence of surface tension. In the present model, one would modify the expressions for the P_{ij} (§ I), so as to increase P_{ii} in the case of a crossing between cells. The modifications could also take into account the method of preparation of the sample. Thus, if a third phase (for instance, metal) were introduced into an already-formed two-phase (solid plus void) system, one would not find regions of the third phase totally surrounded by the other solid phase.

Different phases may be treated differently as follows. Having filled some predetermined fraction of the cubes with solid material of phase I, one divides each of the unfilled cubes into smaller cubes and fills some predetermined fraction of these with solid phase II. For supported metal catalysts, phase I represents the support and phase II the metal, which is known to be more highly dispersed; this treatment mimics the method of preparation of the catalyst. The correlation function and scattering are readily calculated using the arguments given in the present paper (Coppa & Goodisman, 1981).

As usual, there is only one simple model and a multitude of ways in which it can be modified. A sufficiently complex model can always be parameterized to fit any experimental scattering curve. Nevertheless, the simplicity with which calculations can be made analytically, plus the detailed physical picture provided, make the pursuit of the cell models, introduced in this work, valuable.

This work was partially supported by a grant from the National Science Foundation.

APPENDIX

Crossing probability for squares

We consider crossing probabilities for r and x values corresponding to regions I-VI, as shown in Fig. 3. For

region I, the probability averaged over u ($u = L - y$) is

$$L^{-1} \left(\int_0^r du p_2 + \int_r^L du p_1 \right) = r/\pi L.$$

Here, the subscripts 1–4 refer to the cases listed in the second paragraph of § III. For region II the probability averaged over u is

$$L^{-1} \left(\int_0^{(r^2-x^2)^{1/2}} du p_3 + \int_{(r^2-x^2)^{1/2}}^r du p_2 + \int_r^L du p_1 \right) \\ = (2\pi L)^{-1} [r + x + x \ln(r/x)]$$

and for the other regions it is (III)

$$L^{-1} \left(\int_0^t du p_4 + \int_t^{(r^2-x^2)^{1/2}} du p_3 + \int_{(r^2-x^2)^{1/2}}^r du p_2 + \int_r^L du p_1 \right) \\ = (2\pi L)^{-1} \left[(L-x) \ln \frac{r}{L-x} + x \ln \frac{r}{x} + L \right]$$

with $t^2 \equiv r^2 - (L-x)^2$; (IV)

$$L^{-1} \left(\int_0^t du p_4 + \int_t^{(r^2-x^2)^{1/2}} du p_3 + \int_{(r^2-x^2)^{1/2}}^L du p_2 \right) \\ = (2\pi L)^{-1} \left[(L-x) \ln \frac{r}{L-x} + x \ln \frac{r}{x} + L \right. \\ \left. + 2L \cos^{-1} \frac{L}{r} - 2(r^2 - L^2)^{1/2} \right];$$

(V)

$$L^{-1} \left(\int_0^t du p_4 + \int_t^L du p_3 \right) \\ = (2\pi L)^{-1} \left[(L-x) \ln \frac{r}{L-x} + \frac{x}{2} \ln \frac{x^2 + L^2}{x^2} \right. \\ \left. + L \cos^{-1} \left(\frac{L}{r} \right) - (r^2 - L^2)^{1/2} + L \cot^{-1} \left(\frac{L}{x} \right) \right];$$

(VI)

$$L^{-1} \int_0^L du p_4 = (2\pi L)^{-1} \left[L \cot^{-1} \left(\frac{L}{L-x} \right) \right. \\ \left. + L \cot^{-1} \left(\frac{L}{x} \right) + \left(\frac{L-x}{2} \right) \ln \frac{L^2 + (L-x)^2}{(L-x)^2} \right. \\ \left. + \frac{x}{2} \ln \left(\frac{L^2 + x^2}{x^2} \right) \right].$$

We now must average over x , r still being fixed.

The integration over x takes us into different regions of I–VI, depending on the value of r . If $r > L\sqrt{2}$, we are always in region VI. Averaging the probability of

crossing the top over x from 0 to $L/2$, we obtain $\frac{1}{4}$. This is expected, since the line segment must cross if it is longer than the square diagonal. If $L\sqrt{2} \geq r > L\sqrt{5}/4$, we pass through regions VI and V; if $L\sqrt{5}/4 \geq r > L$, we pass through V and IV. Both cases give the same averaged probability

$$\pi^{-1} \left[\frac{1}{2} + \frac{r^2}{4L^2} + \cos^{-1} \frac{L}{r} - \frac{(r^2 - L^2)^{1/2}}{L} \right].$$

Similarly, we get the same result for $L/2 < r \leq L$ (considering regions II and III) and for $0 < r \leq L/2$ (considering II and I):

$$(\pi L^2)^{-1} (rL - r^2/4).$$

We may now write the probability of *no* crossing for the square lattice:

$$P_0 = 1 - \frac{4}{\pi} \frac{r}{L} + \frac{r^2}{\pi L^2}, \quad 0 \leq r < L \\ = 1 - \frac{2}{\pi} - \frac{r^2}{\pi L^2} - \frac{4}{\pi} \cos^{-1} \left(\frac{L}{r} \right) \\ + \frac{4}{\pi} \frac{(r^2 - L^2)^{1/2}}{L}, \quad L \leq r < L\sqrt{2} \\ = 0, \quad r \geq L\sqrt{2}.$$

P_0 is continuous and has a continuous first derivative.

References

- COPPA, N. & GOODISMAN, J. (1981). *J. Appl. Cryst.* Submitted.
 DEBYE, P., ANDERSON, H. R. & BRUMBERGER, H. (1957). *J. Appl. Phys.* **28**, 679–683.
 GOODISMAN, J. (1980). *J. Appl. Cryst.* **13**, 132–134.
 GOODISMAN, J. & BRUMBERGER, H. (1971). *J. Appl. Cryst.* **4**, 347–351.
 GOODISMAN, J. & BRUMBERGER, H. (1979). *J. Appl. Cryst.* **12**, 398–399.
 GUINIER, A. & FOURNET, G. (1955). *Small-Angle Scattering of X-rays*. New York: John Wiley.
 MÉRING, J. & TCHOUBAR, D. (1968). *J. Appl. Cryst.* **1**, 153–165.
 MILLER, A. (1961). PhD Thesis. Univ. of Missouri.
 MILLER, A. & SCHMIDT, P. W. (1962). *J. Math. Phys. (NY)*, **3**, 92–96.
 POWELL, M. J. (1979). *Phys. Rev. B*, **20**, 4194–4198.
 ROACH, S. A. (1968). *The Theory of Random Clumping*. London: Methuen & Co.
 SASVÁRI, K. (1976). In *Contact Catalysis*, Vol. 2, edited by Z. G. SZABO. Amsterdam: Elsevier Scientific.
 SATTERFIELD, C. N. (1970). *Mass Transfer in Heterogeneous Catalysis*. Cambridge, Mass.: MIT Press.
 SCHMIDT, P. W. (1965). *J. Math. Phys. (NY)*, **6**, 424–431.
 SHANTE, V. K. S. & KIRKPATRICK, S. (1971). *Adv. Phys.* **20**, 325–351.

SOLYMOSSI, F. (1976). In *Contact Catalysis*, Vol. 2, edited by Z. G. SZABO. Amsterdam: Elsevier Scientific.

SOMORJAI, G. A., POWELL, R. E., MONTGOMERY, P. W. & JURA, G. (1967). In *Small-Angle X-ray Scattering*, edited by H. BRUMBERGER. New York: Gordon & Breach.

STOKES, A. R. & WILSON, A. J. C. (1942). *Proc. Cambridge Philos. Soc.* **38**, 313–322.

THOMAS, J. M. & THOMAS, W. J. (1967). *Introduction to the Principles of Heterogeneous Catalysis*. New York: Academic Press.

Acta Cryst. (1981). **A37**, 180–183

A Fast Algorithm for Determining Phase Values for Symbols from Weighted Symbol Relations in Direct Methods

BY PAUL T. BEURSKENS AND PETER A. J. PRICK

Crystallography Laboratory, Toernooiveld, 6525 ED Nijmegen, The Netherlands

(Received 30 January 1980; accepted 25 September 1980)

Abstract

The use of symbols to express phases in direct methods leads to relations between symbols (symbol relations). The relative probability of a symbol relation is defined in such a way that it has additive properties. The algorithm given in this paper is developed for the deconvolution of the symbols and is applicable to symmetric phases (0 or π) and to anti-symmetric phases ($+\frac{1}{2}\pi$ or $-\frac{1}{2}\pi$); weighted symbol relations are used as input and figures of merit are calculated for all permutations of phases for the symbols. The algorithm is especially useful when a large number of symbols is used.

Introduction

Nowadays, the majority of structures solved by direct methods are probably solved by using the multisolution program *MULTAN* (Main, Lessinger, Woolfson, Germain & Declercq, 1977). However, the use of symbols to express unknown phases remains a powerful tool as well. Several methods to process and decode symbolic phases have been developed (Karle & Karle, 1966; Beurskens, 1964; Schenk, 1971; and others).

It is the authors' view that by using more than just a few symbols, causing a large number of reflections to take part in the initiation of the calculations, one avoids the use of single or weak phase relationships in the initial – often crucial – stages of a phase generation procedure. This requires a fast and convenient algorithm for disentangling the large bulk of symbol information. In this paper we describe a suitable algorithm (*SYMAN*) for the deconvolution of symbol relations into symmetric phases ($0, \pi$) or anti-symmetric phases ($+\frac{1}{2}\pi, -\frac{1}{2}\pi$).

Notation

$\varphi_{\mathbf{h}}$ is the (unknown) phase of a reflection \mathbf{h}
 a, b, c, \dots are the symbols that represent unknown phases
 x is a linear combination of symbols (e.g. $a - 2b$)
 n is a numerical phase, or the numerical part of a symbolic phase (e.g. $\varphi_{\mathbf{h}} = a - 2b + \pi$ can be written as $\varphi_{\mathbf{h}} = x + n$)
 α is the weight associated with the use of the sigma-2 or tangent formula for the determination of the phase $\varphi_{\mathbf{h}}$
 $\alpha = a \exp(i\varphi_{\mathbf{h}}) = 2\sigma_3 \sigma_2^{-3/2} |E_{\mathbf{h}}| \sum_k |E_{\mathbf{h}-\mathbf{k}} E_{\mathbf{k}}|$

$$\times \exp i(\varphi_{\mathbf{h}-\mathbf{k}} + \varphi_{\mathbf{k}}), \quad (1)$$

where the summation is restricted to terms in which $(\varphi_{\mathbf{h}-\mathbf{k}} + \varphi_{\mathbf{k}})$ is expressed by the same symbol x , $\sigma_m = \sum_{j=1}^N Z_j^m$ for N atoms per unit cell

Symbol relations

Suppose the phase $\varphi_{\mathbf{h}}$ has been calculated as $x_1 + n_1$, with weight α_1 , and, independently, as $x_2 + n_2$, with weight α_2 . This gives the symbol relation

$$x_1 + n_1 - x_2 - n_2 = 0 \pmod{2\pi}. \quad (2)$$

The variance of this result (acentric), or the probability that this relation is correct (centric), is a function of α_1 and α_2 .

In view of computer time and programming convenience it is generally desirable to replace variances and probabilities by weights which have additive