

Applications of Mathematics to Structural Chemistry

Jerome Karle

Laboratory for the Structure of Matter, Naval Research Laboratory, Washington, D.C. 20375-5320

Received May 18, 1993*

The main technique for determining molecular structure in the gaseous and crystalline states is diffraction, electron diffraction for gases, and X-ray diffraction for crystals. This article concerns a review of the development of the modern theory for transforming the diffraction data into structural information. Several points are emphasized, namely, the value of introducing mathematical and physical constraints into the analysis, the adjustments that often need to be made in order to apply ideal mathematics to non-ideal experimental data, and how simple concepts that may arise in one technique can be applied with revolutionary effects in a related technique. In the cases considered here the concept of particular interest is nonnegativity. The more complex crystal structures still offer a challenge to the analytical methods. It is pointed out that some of the mathematics that is now decades old continues to offer opportunities, especially with the great progress in computational facilities, to make future advances with structural problems that constitute the present day challenges.

1. INTRODUCTION

Mathematics has played a key role in the unraveling of three-dimensional atomic arrangements in individual molecules and in the crystalline state from information obtained by the use of diffraction techniques. Diffraction concerns the scattering of various types of incident beams, e.g., electrons, X-rays or neutrons by substances in the various states of matter. The structures (atomic arrangements) of many thousands of molecules are now determined each year, mainly in the crystalline state, by X-ray diffraction. Although much of this work is routine, some problems remain. For the most part, they concern the larger structures. Investigators find increasing difficulties as the numbers of non-hydrogen atoms range from 100 to 200 and more. For the very large molecular structures having very large numbers of non-hydrogen atoms, ranging from about 500 atoms to many thousands, heavy-atom techniques have been readily applicable and successfully applied. New types of macromolecular structures are currently inaccessible by the techniques used for smaller molecules. The objective of this article is to recount the philosophical developments in the analysis of diffraction data that have led to the currently applied methods. This is not only of interest in terms of a description of the mathematical techniques involved, it will be seen that the foundation mathematics for crystal structure analysis has unused potential for facilitating the solution of the larger structures.

With greatly enhanced modern computing facilities, the time for using this potential may be approaching. The foundation mathematics for crystal structure analysis is contained within the infinite set of determinants that comprise the necessary and sufficient conditions for a Fourier series to be nonnegative. The first work on nonnegative Fourier series was done by Toeplitz in 1911.¹ Interest in this subject arose from earlier experiences that my wife, Isabella Karle, and I had in the analysis of molecular structure in the vapor state by use of the electron diffraction technique. The developments in the analysis of molecular structure from the point of view of the mathematics involved will be recounted, and some possibilities for the future will be described.

A subject emphasized in this article concerns the role of nonnegativity in structural research in terms of its immediate applications, its special mathematical features, and the use of

additional probabilistic insights. The experimental techniques considered are electron diffraction by gases and X-ray diffraction by crystals.

2. PROBLEMS AND SOLUTIONS IN ELECTRON DIFFRACTION BY GASES

Experiments in electron diffraction by gases involve the perpendicular intersection of a narrow beam of electrons with a narrow jet of gas. The collision generates a radially symmetric scattering pattern which, when recorded on a photographic plate, appears to consist of maxima and minima. In the early days of the technique the intensities of the maxima and minima were estimated visually and their radii were measured by use of homemade instruments such as a light box furnished with calipers. The maxima and minima of interest that represented the interference or "molecular" scattering which could reveal the atomic arrangements were not really present, but were rather a fortunate consequence of the physiological response of the human eye. The molecular scattering was, in fact, superimposed on a steeply falling background intensity, so the recorded total scattering actually consisted of a steeply falling curve with some ripples added. This may be seen in Figure 1, where curve B is a microphotometer trace made while the photographic plate was being spun. Curve A, made without spinning, shows the effect of the graininess of the photographic emulsion on the microphotometer trace. Despite the rather approximate nature of the molecular scattering intensity obtained by the visual estimates of the intensities, by and large rather good results were obtained for the geometries of many molecules, doubtless facilitated by the rather accurate measurements of the positions of the maxima and minima. This was a matter of individual facility, however, and it was bound to vary among the various investigators.

There were therefore strong motivations to make the data collection a more precise process. A major advance occurred with the invention of a mechanical sector by Finbak,² and later independently by P. P. Debye (P. Debye's son),³ that could be placed in the path of the diffracted intensities before they are recorded. The effect of the sector was to flatten the background intensity considerably so that the molecular scattering would stand out much more clearly and be more accurately measurable by the use of a microdensitometer. While a good start on improving the experimental data was

* Abstract published in *Advance ACS Abstracts*, February 15, 1994.

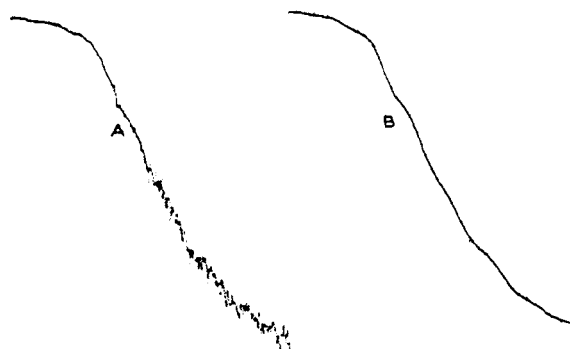


Figure 1. Microphotometer traces of an electron diffraction pattern showing the steeply falling background scattering on which the molecular scattering is superimposed, giving rise to the minor oscillations. Trace A was made without rotation of the photographic plate and trace B with rotation.

being made, a potentially major advance in the mathematical analysis of the data was initiated by a theoretical paper published by Debye.⁴ In it, he provided an interpretive analysis of accurate molecular scattering data that would derive from a molecule composed of point atoms rather than actual atoms having a distributed electron density. Debye showed that the Fourier sine transform of such molecular scattering data would represent the probability of finding various distances in the molecule. This resulted in a Fourier transform that presented not only interatomic distances but also the root mean square amplitudes of vibration between pairs of atoms. The potential indicated by the paper of Debye⁴ impressed Isabella Karle and me very much, and we looked forward to turning the potential into actuality. We pursued this when we arrived at the Naval Research Laboratory.

Actual practice often does not immediately follow from theoretical analyses. Although the mechanical sector and microdensitometer were a great help, there were still many impediments to the application of the theory. Among them were the presence of real atoms rather than point atoms, the limited angular range of diffraction data, and the question of how to draw a background line accurately through the oscillations of the molecular scattering in order to achieve a separation of the molecular scattering intensity from the background intensity. All the problems associated with these three impediments were solved by the application of certain mathematical and physical conditions. The solutions of the problems merit some discussion, not only because it was necessary to solve the problems in order to realize the detailed structural information described in the idealized theory of Debye but also because of the major impact that the solutions would have on the future developments in crystal structure analysis.

In his paper, Debye⁴ recognized that the scattering effects from the electron distributions of real atoms would interfere with the accuracy of the Fourier transform of the molecular intensity data, and, in fact, would introduce negative features into the transform, making it evident that it could not be a probability function. Debye suggested a form for a corrective function that could be applied to an accurate Fourier transform. In practice, such a procedure has never been tried. At best, it would be a complication of uncertain validity, especially when probability distributions for individual interatomic distances were superimposed, as generally occurs for all but the simplest of molecules. In addition, as will be seen, the scattering from the electron densities around atoms actually interferes with the obtaining of an accurate Fourier transform from the molecular intensity, so there would be difficulty in

fulfilling the premise associated with the application of a correction. The problem of transforming the molecular intensity for real atoms to that for point atoms before taking the Fourier transform was solved, and therefore the practical procedure was brought one step closer to achieving an ideal, nonnegative transform. A somewhat simplified description of how this could be done, based on insight concerning the shapes of atomic scattering factors,⁵ follows.

The total intensity of scattering is

$$I_t(s) = K \left\{ \sum_{i=1}^n \frac{S_i(s)}{s^4} + \sum_{i=1}^n \frac{f_i(s)^2}{s^4} + 2 \sum_{i < j}^n \frac{f_i(s) f_j(s)}{s^4} A_{ij} \right\} \quad (1)$$

where the constant K is definable in terms of the intensity of the incident beam and a variety of physical constants, n is the number of atoms in the molecule, $S_i(s)$ is the inelastic scattering function for the i th atom, $s = [4\pi \sin(\theta/2)]/\lambda$, θ is the angle between the incident and scattered beams, λ is the wavelength of the electron beam, f_i/s^2 is the atomic scattering factor for electrons for the i th atom, and $A_{ij}(r_{ij}, s)$ is an interference function associated with the i th and j th atoms, where r_{ij} is the distance between the i th and j th atoms. The molecular scattering intensity can be defined in terms of the A_{ij} for all pairs of atoms in the molecule. The incoherent term is relatively small except at small values of s .

The factor f_i may be written $f_i = f x Z_i$, where f is a shape function that changes slowly with increasing atomic number. This suggests that dividing $I_t(s)$ by the background intensity, I_b , where

$$I_b(s) = K \sum_{i=1}^n (S_i(s) + f_i(s)^2)/s^4 \quad (2)$$

may well produce coefficients of the interference terms in the molecular scattering intensity that are essentially constant. The coefficients then become

$$c_{ij} = \{f_i(s) f_j(s) / \sum_{i=1}^n [S_i(s) + f_i(s)^2]\} \quad (3)$$

and the molecular intensity becomes

$$I_m(s) = \sum_{i=1}^n \sum_{j=1}^n c_{ij} A_{ij} \quad i \neq j \quad (4)$$

The use of a mechanical sector to lower the steepness of the background affects all the terms in (1) similarly and therefore does not affect the results in (3) and (4).

A general form for A_{ij} is

$$A_{ij} = \int_0^\infty P_{ij}(\rho) (\sin s\rho/s\rho) d\rho \quad (5)$$

and a common form for A_{ij} is

$$A_{ij} = \exp(-\langle l_{ij}^2 \rangle s^2/2) \sin sr_{ij}/sr_{ij} \quad (6)$$

where $P_{ij}(\rho)$ is the probability that the distance between the i th and j th atoms has a value in the interval ρ and $\rho + d\rho$. If the molecule were rigid, then the interference function A_{ij} would be equal to $\sin sr_{ij}/sr_{ij}$. The exponential term in (6) is appropriate when molecules undergo harmonic vibration. The quantity $\langle l_{ij}^2 \rangle$ is the mean square amplitude of vibration between the i th and j th atoms. Deviations from harmonic vibration require special treatment.

As noted, it is important that the c_{ij} be independent of s . This has been found to be the case. There are circumstances

where complications arise, for example, when atoms of rather different atomic number are present in a molecule and the f_i must be treated as complex numbers. Such circumstances have been adequately handled. For further insights into these matters, refer to the analysis and references presented in ref 6.

We now know that to very good approximation, it should be possible, with the use of the background intensity, to transform the molecular intensities from ones having atomic scattering factors that vary with scattering angle to ones having constant scattering factors. There is still an additional matter that, unless handled properly, can interfere seriously with the possibility of discovering a proper background intensity to be used with the molecular intensity to transform it to one with constant coefficients. This matter is the angular limitation of the experimental data. It becomes unreliable at some maximum s -value, s_{\max} . In the usual experimental circumstances, the absent data would cause negative regions to appear in the Fourier transform of the molecular intensity data. It had been important to remove this cutoff error even when visually estimated molecular intensity data were used in the Fourier transform calculation to make a so-called radial distribution function (RDF). The latter was used to obtain initial information concerning interatomic distances and other geometric features. The problem was solved by Degard⁷ and by Schomaker⁸ in a simple fashion by the multiplication of the molecular intensity data by a function of the type e^{-as^2} , where the value of a would be so chosen that the molecular intensity would be suitably small at s_{\max} and beyond.

If (6) is substituted into (4), the molecular scattering intensity becomes

$$I_m(s) = \sum_{i=1}^n \sum_{j=1}^n c_{ij} \int_0^\infty P_{ij}(\rho) (\sin s\rho/s\rho) d\rho \quad i \neq j \quad (7)$$

It was shown by Debye⁴ that if the coefficients c_{ij} in (7) were independent of the variable s , a significant result could be readily obtained by the application of Fourier transform theory to (7). The Fourier transform theory in terms of $I_m(s)$ is

$$D(r) = (2/\pi)^{1/2} \int_0^\infty s I_m(s) \sin sr ds \quad (8)$$

where, evidently, $D(r)$ is a function of the distance variable r . When c_{ij} are independent of s , integration of (8) gives

$$D(r) = (\pi/2)^{1/2} \sum_{i=1}^n \sum_{j=1}^n c_{ij} P_{ij}(r)/r \quad i \neq j \quad (9)$$

It is seen from (9) that $r(D)r$ is expressible as a weighted sum of probability functions for all the interatomic distances that occur in a molecule. From this information, molecular structure and internal motion readily follow.

The integrand of (8) is now altered by introducing the factor $\exp(-as^2)$. A suitably chosen value of a would ensure that the value of the integration between the limits s_{\max} to infinity would be negligible while preserving as well as possible the information in the range $0 \leq s \leq s_{\max}$. Equation 8 is thus replaced by

$$f(r) = (2/\pi)^{1/2} \int_0^{s_{\max}} s I_m(s) \exp(-as^2) \sin sr ds \quad (10)$$

If molecules undergo harmonic vibrations, the function $rD(r)$

in (9) can be represented by

$$rD(r) = (\pi/2)^{1/2} \sum_{i=1}^n \sum_{j=1}^n c_{ij} (h_{ij}/\pi)^{1/2} \exp[-h_{ij}(r_{ij}-r)^2] \quad i \neq j \quad (11)$$

where $(2h_{ij})^{-1/2} = \langle l_{ij}^2 \rangle^{1/2}$ is the root mean square amplitude of vibration for the pair of atoms denoted by i and j . It has been shown by use of the falting integral in Fourier transform theory^{9,10} that if $rD(r)$ is defined by (11), $r_f(r)$ defined by (10) has the form

$$r_f(r) = (\pi/2)^{1/2} \sum_{i=1}^n \sum_{j=1}^n c_{ij} (h_{ij}/\pi)^{1/2} \exp[-h_{ij}(r_{ij}-r)^2] / (4ah_{ij} + 1) / (4ah_{ij} + 1)^{1/2} \quad i \neq j \quad (12)$$

It is seen from (12) that the effect of introducing $\exp(-as^2)$ is to broaden the peaks representing the individual interatomic distances in the RDF in a known way. The positions of the maxima are preserved. It is a simple matter to find the h_{ij} from the values of the $h_{ij}/(4ah_{ij} + 1)$ obtained from fitting the peaks in the experimental RDF since the value of a used in the calculation of (10) is known.

Special motions that involve deviations from harmonic motion, e.g., internal rotation, require special analysis of the resulting RDF. Debye⁴ used the saddle point method¹¹ to obtain a function that represented the molecular intensity for pairs of atoms affected by free internal rotation. Evaluation of the appropriate integral proceeds by identifying certain extrema called the saddle points and choosing a path of deepest descent, steepest decrease of magnitude of the integrand, through the saddle points. The special cases of free and hindered rotation are accompanied by additional motions that derive from the overall vibration of the molecular frame.¹² Multiple rotors present additional complications. They too have been investigated.¹³

At this point, it is known what to do to obtain a molecular intensity function with coefficients that are independent of s . It is also known how to remove the problem of having to use data whose angular range is limited. One main problem remains. The manner in which it was solved will be seen to have had a significant effect on the course of crystal structure analysis.

It is evident from the preceding discussion that the molecular intensity data and the background scattering that comprise the total intensity must be separated. It was not possible to employ the theory for the background scattering for this purpose because of inaccuracies in the theory and also certain limitations in the photographic recording of the total intensity. Nevertheless the problem was readily solved by taking advantage of the fact that the resulting RDF obtained from an appropriate Fourier transform of the molecular intensity had to be a nonnegative function. The smooth background line to be drawn through a microdensitometer trace of the total intensity, that would distinguish the background scattering from the molecular scattering, had to be so drawn that, after transformation of the molecular scattering to constant coefficients, correcting for cutoff of the data, and taking the Fourier transform, the resulting RDF had to be nonnegative. If the first attempt was not entirely successful, as was often the case, successive corrections could be made. The deviations from nonnegativity could be used to determine the nature of the adjustments to the background line. There are additional operations on the molecular intensity that are required because of the nature of the experiment, e.g., the absence of data in the vicinity of the main beam (i.e. near $s = 0$) because of the

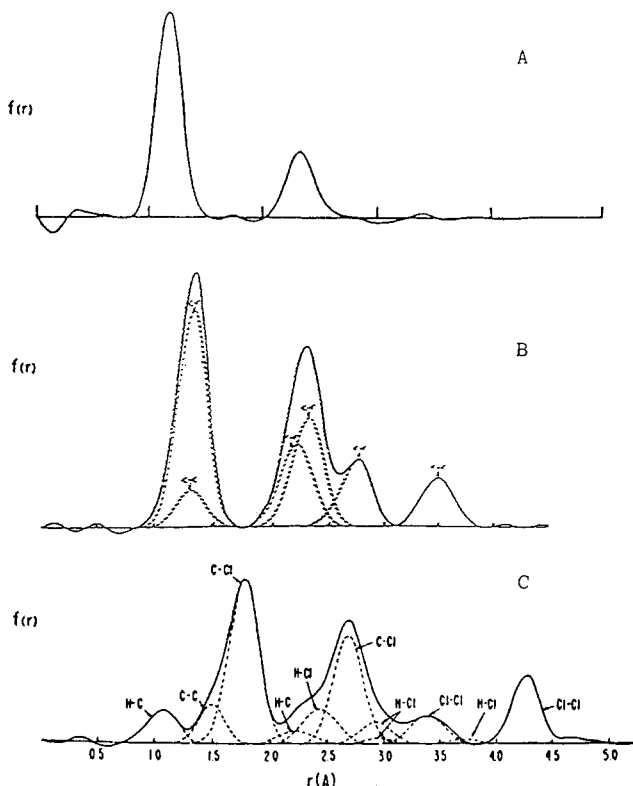


Figure 2. Composite of three different RDFs for CO_2 , F_2CCF_2 , and $\text{ClH}_2\text{CCH}_2\text{Cl}$. The point of interest here is the nonnegativity of the functions, which improved somewhat as the technique for effecting small corrections improved. The solid lines are the computed RDFs, and the dashed lines indicate the individual peaks for each interatomic distance.

need to exclude it from the collected diffraction pattern. This matter and other details are discussed in the review⁶ and the referenced literature therein.

With the success of the nonnegativity criterion in ensuring the determination of an appropriate background line, it was possible to make molecular structure applications that gave good estimates of the root mean square amplitudes of vibration between pairs of atoms as well as rather accurate geometries and interatomic distances.

The RDFs for CO_2 ,¹⁰ C_2F_4 ,¹⁴ and $\text{ClH}_2\text{CCH}_2\text{Cl}$ ¹⁵ exhibited in Figure 2A–C, respectively, show the results of applying the nonnegativity criterion to the drawing of the line representing the background intensity. Improvements in our computing facilities made successive adjustments to the background line easier to carry out, so that, as time passed, the obtaining of essentially nonnegative RDFs was readily achieved. For the more complex structures, it was evidently necessary to decompose those peaks that were composed of overlapping contributions into their individual contributors. The labeled, dashed curves show the individual contributors. Uniqueness in the decomposition arises from the known shapes of the contributors, usually Gaussian except for special types of internal motion, and the known areas under the curves. The widths of the peaks are related to the root mean square amplitudes of vibration between the pairs of atoms, and the maxima afford information concerning equilibrium interatomic distances. In the case of 1,2-dichloroethane, the analysis showed that the gas sample was a mixture of conformers with a ratio of trans to gauche of about 3 to 1. RDFs can also afford information concerning potential barriers hindering internal rotation. Generally, final structural parameters are obtained by computing least-squares fits to the

molecular intensity curve which is the Fourier transform of the RDF.

Many thousands of structures have been investigated by the gas electron diffraction technique, providing information concerning molecular configuration, bond distances and angles, gross internal motion and internal rotation, preferred orientation in conformers, and conjugation and aromaticity. Theoretical investigations appropriate to the analysis of internal rotation have provided the opportunity to estimate the heights of barriers to internal rotation. Gas electron diffraction studies at different temperatures have facilitated the estimation of barrier heights and also the evaluation of thermodynamic quantities. Experiments have been made at very high temperatures in order to make gas electron diffraction studies of low-volatility compounds. Some studies have concerned mixtures in equilibrium and a few free radicals. It has been found worthwhile to combine results from spectroscopic data with those from gas electron diffraction to the benefit of both techniques. Extensive discussions of many of these topics are to be found in the two-volume publication *Stereochemical Applications of Gas-phase Electron Diffraction* edited by I. and M. Hargittai.¹⁶

As a consequence of observing the usefulness of the application of the nonnegativity criterion to the analysis of gas diffraction data, it was considered worthwhile to examine other areas in which this criterion may be useful for obtaining useful information. The electron density distributions of free atoms and those bound in crystal structures are examples of areas of interest. The profound effect that the nonnegativity criterion has had on the course of crystal structure analysis will be discussed next.

3. PROBLEMS AND SOLUTIONS IN X-RAY DIFFRACTION BY CRYSTALS

The diffraction of X-rays by crystals and electrons by gases have certain experimental parallels. In the X-ray diffraction experiment, an X-ray beam replaces the electron beam and a crystal replaces the jet of gas. There are not many more parallels, certainly not in the nature of the diffraction intensities nor the manner in which structures are generally derived from the intensities. The incident X-ray beam appears to be scattered from various imaginary planes cutting through a crystal, similarly to the reflection of light from a mirror, with the exception that the reflection takes place at only a specific angle of incidence whose value depends upon the wavelength of the incident beam. For a fixed beam direction and wavelength, it is necessary to change the crystal orientations in order to collect a large number of data.

The structure of a crystal was shown by Bragg to be representable by a three-dimensional Fourier series appropriate to the three-dimensional periodicity of a crystal. The Fourier series represents the electron density distribution in a crystal which is the structure since the maxima occur at the centers of the atoms. Crystal structure determination from the diffraction data is not straightforward. The problem arises from the fact that the coefficients of the Fourier series are in general complex numbers, whereas only their magnitudes are normally obtained from experiment. The associated angle (phase) is not obtained, and this was the basis for the general impression in the crystallographic community that experimental diffraction data did not suffice for the direct determination of a crystal structure. The coefficients are real for crystals with a center of symmetry. There is still an ambiguity, however, because the phases can have values of either zero or π . There were some early workers before 1940 who realized

that the diffraction data may suffice to determine crystal structures and made some studies to solve the problem. By and large nothing practical was forthcoming and the work seemed to have been ignored. An exception was the work of Patterson^{17,18} who developed a Fourier series that has the magnitudes squared as the coefficients. The maxima of this series represent the interatomic vectors among the atoms in a crystal and are directly calculable from the diffraction data. This function soon played a valuable role in crystal structure determination, particularly when not more than a few heavier atoms were present in a structure. The interatomic vectors associated with the heavier atoms could be readily identified, and therefore the heavy-atom structure could be deduced. Once the heavier atom structure is known, there are methods for determining the rest of the structure, often fairly readily. This so-called heavy-atom method has made an important contribution to structure analysis, and there are aspects of it that still play a valuable role. When no disproportionately heavy atoms are present, the resolution of the interatomic vectors is too poor to identify individual ones for even moderately complex structures. The structure analysis problem was ultimately solved by finding a way to obtain the needed values of the phases directly from use of the measured magnitudes.

The electron density distribution, $\rho(r)$, is expressed in terms of the three-dimensional Fourier series

$$\rho(r) = V^{-1} \sum_{\mathbf{h}=-\infty}^{\infty} F_{\mathbf{h}} \exp(-2\pi i \mathbf{h} \cdot \mathbf{r}) \quad (13)$$

where V is the volume of the unit cell of the crystal. The unit cell is the basic structural unit from which, through three-dimensional periodicity, the crystal is formed. The coefficients

$$F_{\mathbf{h}} = |F_{\mathbf{h}}| \exp(i\phi_{\mathbf{h}}) \quad (14)$$

are the crystal structure factors associated with the planes labeled with the vector \mathbf{h} . The \mathbf{h} have integer components, h , k , and l , whose values are inversely proportional to the intercepts on the x , y , and z axes, respectively, of planes cutting through the crystal. The angle $\phi_{\mathbf{h}}$ is the phase associated with $F_{\mathbf{h}}$, and \mathbf{r} labels the position of any point in the unit cell. $F_{\mathbf{h}}E$ is the amplitude of the scattered wave associated with the plane labeled by \mathbf{h} , where E is the electric field vector of the incident beam. The measured intensities of X-ray scattering are proportional to $|F_{\mathbf{h}}|^2$. If the values of $\phi_{\mathbf{h}}$ were also obtained directly from experiment, structures could be immediately calculated from (13). The seeming absence of this information gave rise to the so-called "phase problem".

The Fourier inversion of (13) followed by the replacement of the integral by the sum of contributions from the N discrete atoms in the unit cell gives, for the Fourier coefficient,

$$|F_{\mathbf{h}}| \exp(i\phi_{\mathbf{h}}) = \sum_{j=1}^N f_{j\mathbf{h}} \exp(2\pi i \mathbf{h} \cdot \mathbf{r}_j) \quad (15)$$

where $f_{j\mathbf{h}}$ represents the amplitude of scattering of the j th atom in the unit cell and \mathbf{r}_j is its position vector.

It is a simple matter to show that a crystal structure is greatly overdetermined by the measured intensities. This does not tell one how to find a solution, but it makes the search rational. A system of simultaneous equations is formed by the definition of the crystal structure factors given by (15) since the values of the scattered intensities are measured for a large number of \mathbf{h} . The unknown quantities in (15) are the phases $\phi_{\mathbf{h}}$ and the atomic positions \mathbf{r}_j . The known quantities are the $|F_{\mathbf{h}}|$ obtained from the measured intensities and the $f_{j\mathbf{h}}$,

which differ little from the theoretically calculated atomic scattering factors for free atoms. Since each equation in (15) involves complex quantities, there are really two equations, one for the real and one for the imaginary part. In order to determine the overdeterminacy, a comparison is made of the number of unknown quantities with the number of independent data available. With the use of Cu $K\alpha$ radiation, the overdeterminacy can be as great as a factor of about 50 for crystals that have a center of symmetry and about 25 for those that do not. In practice, somewhat fewer than the maximum available data are measured, but the overdeterminacy is still quite high.

The phases may be eliminated from (15) by multiplying by the corresponding complex conjugates to obtain

$$|F_{\mathbf{h}}|^2 = \sum_{j=1}^N \sum_{k=1}^N f_{j\mathbf{h}} f_{k\mathbf{h}} \exp[2\pi i \mathbf{h} \cdot (\mathbf{r}_j - \mathbf{r}_k)] \quad (16)$$

The Fourier transform of (16) is known as the Patterson function.^{17,18}

$$P(\mathbf{r}) = \sum_{\mathbf{h}=-\infty}^{\infty} |F_{\mathbf{h}}|^2 \exp(-2\pi i \mathbf{h} \cdot \mathbf{r}) \quad (17)$$

Its properties and usefulness have been previously discussed.

Relationships between phases and magnitudes that anticipated the later developments were the inequalities of Harker and Kasper.¹⁴ They derived a number of inequalities by application of the Schwarz and Cauchy inequalities to the structure factor equations (15) in the presence of crystallographic symmetry. The Harker-Kasper inequalities have provided valuable insights. For example, the simple inequality formulas can provide useful phase information as shown by Kasper, Lucht, and Harker²⁰ in their solution of the structure of decaborane. In addition, work with the inequalities indicated that they may have probabilistic characteristics. Gillis,²¹ for example, noted that the implication of an inequality was often correct even when the magnitudes of the structure factors were too small to permit a definitive conclusion to be drawn. Gillis speculated that the smallness of the structure factor magnitudes may have been due to uncorrected thermal effects. The probabilistic interpretation, however, remained a possible alternative, namely, that although an inequality does not quite determine the value of a phase definitively, it still does so with a high probability that the value is correct. This could be important because it would imply that the inequalities have probabilistic implications that could extend their range of usefulness.

As indicated, the initial motivation to investigate the mathematics of crystal structure determination arose from experiences in the development of an analytical procedure for obtaining accurate radial distribution functions for determining the structures of gaseous molecules by electron diffraction. At about the time this work in gas electron diffraction was proceeding, Herbert Hauptman joined our group at the Naval Research Laboratory, and, in view of the success of the nonnegativity criterion, we decided to explore the possibility that this criterion might be useful in other areas of structural research. We were quite interested in seeing what the consequences of the application of nonnegativity would be for crystal structure analysis since the electron density distribution defined in (13) is constrained not to be negative. This brought in the work of Toeplitz¹ early in this century on nonnegative Fourier series and subsequent developments by others. We discussed the theory in three dimensions and wrote it in a

form that would have particular relevance to crystallographic data.²²

The fundamental result was that the necessary and sufficient condition for the electron density distribution in a crystal to be nonnegative is that an infinite system of determinants involving the crystal structure factors be nonnegative. A typical determinant is²²

$$\begin{vmatrix} F_{000} & F_{-k_1} & F_{-k_2} & \dots & F_{-h} \\ F_{k_1} & F_{000} & F_{k_1-k_2} & \dots & F_{k_1-h} \\ F_{k_2} & F_{k_2-k_1} & F_{000} & \dots & F_{k_2-h} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ F_h & F_{h-k_1} & F_{h-k_2} & \dots & F_{000} \end{vmatrix} \geq 0 \quad (18)$$

The subscripts in the first column start with 0,0,0 but are otherwise arbitrary. The subscripts in the first row are the same but of opposite sign. The subscript of the element in the *i*th row and *j*th column is the sum of the subscripts of the elements of the *i*th row and first column and the first row and *j*th column. The third-order inequality

$$\begin{vmatrix} F_{000} & F_{-k} & F_{-h} \\ F_k & F_{000} & F_{k-h} \\ F_h & F_{h-k} & F_{000} \end{vmatrix} \geq 0 \quad (19)$$

contains a relationship among the structure factors that has played a most important role in direct crystal structure analysis. This may be seen by rewriting (19) in the form²²

$$\left| F_h - \frac{F_k F_{h-k}}{F_{000}} \right| \leq \left| \frac{F_{000}}{F_k} \right|^{1/2} \left| \frac{F_{000}}{F_{h-k}} \right|^{1/2} / F_{000} \quad (20)$$

For structure factors of unusually large magnitude, the right side of (20) becomes quite small and then

$$F_h \sim F_k F_{h-k} / F_{000} \quad (21)$$

One evident conclusion from (21) is

$$\phi_h \sim \phi_k + \phi_{h-k} \quad (22)$$

This states that for large structure factor magnitudes, the value of ϕ_h may be defined in terms of the values of two other phases.

It can be readily shown²² that all determinants (18) can be written in the form

$$|F_h - \delta| \leq r \quad (23)$$

As the order of the determinants increases, there is a tendency for *r* to decrease in size, making the determination of ϕ_h rather definitive.

Formula 22 has found wide application beyond the range of usefulness of (20). This is because of the probabilistic characteristics of the inequalities²³ which imply that the most likely value of ϕ_h is that of $\phi_k + \phi_{h-k}$, and the probability decreases the farther the value of ϕ_h deviates from that of $\phi_k + \phi_{h-k}$. Therefore, even when the radius, *r*, of the bounding circle is large, the most likely value of ϕ_h is $\phi_k + \phi_{h-k}$.

The structure factors in (18) can be replaced by quasi-normalized structure factors, \mathcal{E} that represent point atoms (to an approximation when atoms of unequal atomic numbers are present) rather than atoms with electron distributions.

$$\mathcal{E}_h = F_h / \left(\sum_{j=1}^N f_j^2 \right)^{1/2} \quad (24)$$

There is a rough parallel here with the method for making point atoms in electron diffraction by gases. Structure factors

representing point atoms are the type of quantity normally used in phase-determining procedures. Instead of (21), we have

$$\mathcal{E}_h \sim \mathcal{E}_k \mathcal{E}_{h-k} / \mathcal{E}_{000} \quad (25)$$

For centrosymmetric crystals, we have

$$s\mathcal{E}_h \sim s\mathcal{E}_k \mathcal{E}_{h-k} \quad (26)$$

where *s* means "sign of", a plus sign implying the phase is equal to zero and a minus sign that it is equal to π . A one-term tangent formula also follows from (25).

$$\tan \phi_h \simeq \frac{|\mathcal{E}_k \mathcal{E}_{h-k}| \sin(\phi_k + \phi_{h-k})}{|\mathcal{E}_k \mathcal{E}_{h-k}| \cos(\phi_k + \phi_{h-k})} \quad (27)$$

The tangent formula composed of the sum of terms over *k* both in the numerator and the denominator is another formula that has played a major role in the practical applications of the theory for structure determination.

After the set of determinantal inequalities (18) were obtained on the basis of the nonnegativity of the electron density distribution in a crystal, it was of interest to investigate their relationship to the inequalities derived by Harker and Kasper¹⁹ from use of the Schwarz and Cauchy inequalities. It was shown²² that when the appropriate symmetry was introduced into the third-order determinantal inequality by means of certain relationships among the structure factors, the Harker-Kasper inequalities could be derived. Examination of the derivation of the Harker-Kasper inequalities shows, as would be expected, that the nonnegativity of the electron density distribution is a requirement for their validity.

The variety of phase determining formulas contained within the determinantal inequalities (18) have their counterpart in probability theory; i.e. similar formulas can be derived with the use of probability theory. Their virtue is that measures of reliability can be attached to them, and the judicious use of such measures was an important feature in bridging the gap between mathematical theory and practical application.

It had been pointed out that the determinants have inherent probabilistic characteristics²³ that can, in fact, be directly read out from the form (23). This was not, however, how the first probability formulas were derived. The pursuit of such formulas was motivated by the expectation that the usefulness of the formulas from the inequalities could be extended because of the great overdeterminancy of the structure problem, the expectation that the points within the bounding circle of radius *r* and center δ would not be uniformly probable, and the experience of Gillis²¹ with structure factor magnitudes that were too small to elicit definitive conclusions from the inequalities but which, when made larger artificially, led to correct conclusions.

The first probabilistic investigation of the phase problem made use of the joint probability distribution which was applied to crystals that have a center of symmetry. The results were in the form of probabilistic formulas and measures for application to the phase problem.²⁴ The probabilistic formulas were essentially equivalent to those extracted from the inequality theory. It was in this investigation²⁴ also that the theory of invariants and semi-invariants was introduced for the purpose of solving the problem concerning how many and what types of phases to specify in order to fix the origin in a crystal and what values are permitted. This theory was ultimately extended to include all 230 crystallographic space groups.²⁵⁻²⁸ The practical aspects of solving crystals that lack a center of symmetry were developed later on, and it was not until 1964 that the structure of the first crystal lacking a

center of symmetry was solved²⁹ by the "direct method" for obtaining phase information by direct use of the measured intensities of scattering.

The main formulas for phase determination are now listed. They will suffice to characterize the nature of the phase determining procedures. There are additional formulas that play a variety of helpful roles.

Centrosymmetric Crystals. The Σ_2 formula is²⁴

$$sE_h \approx s \sum_{\mathbf{k}_r} E_{\mathbf{k}} E_{\mathbf{h}-\mathbf{k}} \quad (28)$$

where s means the sign of and \mathbf{k}_r represents restricted values of \mathbf{k} for which the corresponding $|E_{\mathbf{k}}|$ and $|E_{\mathbf{h}-\mathbf{k}}|$ values are large. A plus sign refers to a phase of zero and a minus sign to a phase of π , the only two values possible for a centrosymmetric crystal when an origin in the crystal is properly chosen. The quantities, E , are normalized structure factors which arise as appropriate quantities to use with probability theory and are the same as the quasi-normalized structure factors, \mathcal{E} , except for a reweighting³⁰ of certain subsets of E . The treatment of the intensity data to obtain normalized structure factors³¹ arises from the work of Wilson,^{32,33} the earliest application of probability methods to crystal structure analysis. Formula 28 is the probability equivalent of the set of inequalities (20) as \mathbf{k} varies over the set \mathbf{k}_r . The appropriate probability function, $P_+(\mathbf{h})$, which represents the probability that the sign of $E_{\mathbf{h}}$ be positive, was given in the monograph of ref 24. It is conveniently applied in the form derived by use of the central limit theorem by Woolfson³⁴ and Cochran and Woolfson.³⁵

$$P_+(\mathbf{h}) \approx 1/2 + 1/2 \tanh(\sigma_3 \sigma_2^{-3/2} |E_{\mathbf{h}}| \sum_{\mathbf{k}} E_{\mathbf{k}} E_{\mathbf{h}-\mathbf{k}}) \quad (29)$$

where

$$\sigma_n = \sum_{j=1}^N Z_j^n \quad (30)$$

and Z_j is the atomic number of the j th atom in the unit cell containing N atoms.

Noncentrosymmetric Crystals. The sum of angles and tangent formulas are, respectively,

$$\phi_{\mathbf{h}} \approx \langle \phi_{\mathbf{k}} + \phi_{\mathbf{h}-\mathbf{k}} \rangle_{\mathbf{k}_r} \quad (31)$$

$$\tan \phi_{\mathbf{h}} \approx \frac{\sum_{\mathbf{k}} |E_{\mathbf{k}} E_{\mathbf{h}-\mathbf{k}}| \sin(\phi_{\mathbf{k}} + \phi_{\mathbf{h}-\mathbf{k}})}{\sum_{\mathbf{k}} |E_{\mathbf{k}} E_{\mathbf{h}-\mathbf{k}}| \cos(\phi_{\mathbf{k}} + \phi_{\mathbf{h}-\mathbf{k}})} \quad (32)$$

Formulas 31 and 32 are comparable to formulas 22 and 27, respectively, and result from combining a number of individual terms as \mathbf{k} varies over some chosen set. An appropriate measure of the reliability of (31) and (32) is a variance, V ,³⁶ given by

$$V = \frac{\pi^2}{3} + [I_0(\alpha)]^{-1} \sum_{n=1}^{\infty} (I_{2n}(\alpha)/n^2) - 4[I_0(\alpha)]^{-1} \sum_{n=0}^{\infty} [I_{2n+1}(\alpha)/(2n+1)^2] \quad (33)$$

where I_n is a Bessel function of imaginary argument³⁷

$$\alpha = \{ [\sum_{\mathbf{k}} \kappa(\mathbf{h}, \mathbf{k}) \cos(\phi_{\mathbf{k}} + \phi_{\mathbf{h}-\mathbf{k}})]^2 + [\sum_{\mathbf{k}} \kappa(\mathbf{h}, \mathbf{k}) \sin(\phi_{\mathbf{k}} + \phi_{\mathbf{h}-\mathbf{k}})]^2 \}^{1/2} \quad (34)$$

$$\kappa(\mathbf{h}, \mathbf{k}) = 2\sigma_3 \sigma_2^{-3/2} |E_{\mathbf{k}} E_{\mathbf{h}-\mathbf{k}}| \quad (35)$$

Expression 33 gives the variance of $\phi_{\mathbf{h}}$ as determined from a given set of $\phi_{\mathbf{k}} + \phi_{\mathbf{h}-\mathbf{k}}$. This variance formula has its origin in a probability distribution (in somewhat different notation) of Cochran³⁸ for $\phi_{\mathbf{h}}$, given a particular $\phi_{\mathbf{h}} + \phi_{\mathbf{h}-\mathbf{k}}$ and the accompanying $|E|$ values. The tangent formula 32 can be derived in many ways. It has arisen, for example, in theoretical investigations of noncentrosymmetric space groups by use of the joint probability distribution.³⁹ The average in (31) is to be taken in the context of maximum clustering, i.e. a minimum deviation of the contributions of individual addition pairs, $\phi_{\mathbf{k}} + \phi_{\mathbf{h}-\mathbf{k}}$ from the average value. All ϕ are kept in a range $-\pi < \phi \leq +\pi$ and maximum clustering requires the addition of 0, $+2\pi$, or -2π to each addition pair. A practical method for effecting appropriate clustering has been described.³⁶

With formulas for phase determination in hand, it was possible to consider the development of a procedure that would have broad practical applications to both centrosymmetric and noncentrosymmetric crystals. The first such procedure was the symbolic addition procedure^{29,36,40} that arose mainly from the efforts of Isabella Karle to combine the probabilistic mathematics of phase determination with experimental data in an appropriate fashion, in order to overcome problems, subtle and otherwise, that could impede successful application. The main, currently used phase determining procedures are variations of the archetypal symbolic addition procedure. Several problems arose in phase determination for noncentrosymmetric crystals concerning, for example, the assignment and handling of symbols, the use of the probability measure (33), the number of possible values to assign to symbols that represent phases whose values range continuously from $-\pi$ to $+\pi$, the combined use of (31) and (32) for phase determination, the proper use of the tangent formula for the processes of phase refinement and phase extension, the development of techniques for specifying an enantiomorph or axis directions or both, and special considerations such as the avoidance of certain troublesome triplet phase invariants involving one- and two-dimensional centric reflections.

It is apparent on examining (28), (31), and (32) that it is necessary to know the values of some phases before additional ones can be evaluated. There are several sources of such information, from certain phase specifications associated with establishing an origin in the crystal,²⁸ the assignment of symbols to some phases for later evaluation, and the use on occasion of auxiliary formulas, such as Σ_1 and Σ_3 , that define individual phases in terms of structure factor magnitudes alone.⁴¹ The number and types of phases to be used for specifying the origin in a crystal has been determined by use of the theory of invariants and semi-invariants that was developed for this purpose. Depending upon the type of space group involved, the number can vary from none at all to as many as 3. Suitable tables²⁸ are available for carrying out this task.

The phase-determining procedure is a stepwise one with few contributors to (28) or (31) at the start. Use at the start of phases associated with the largest possible values of the normalized structure factor magnitudes, $|E|$, will ensure that the probability measures, (29) and (33), will be as large as possible. The large overdeterminacy of the problem helps to

ensure that initial probabilities will be large enough to proceed in a stepwise fashion to build-up a sufficiently reliable set of phase values to effect a solution to the structure problem. Because the nature of phase determination is inherently probabilistic and contingent in a stepwise and interdependent fashion, the problem of establishing optimal procedures based on experimental data was not at all straightforward. There are a very large number of paths through a phase determination. Among many of them are pitfalls in which there arise, for example, temptations to take a path in which the interconnections between phase evaluations flow easily at the expense somewhat of the probability measures. Such paths are more likely to lead to missteps and cumulative errors that could damage or defeat a phase determination than ones that are based on the highest values of the probability measures.

There is also an ambiguousness inherent in procedures for phase determination which is controlled by the use of symbols. The symbols can assume more than one value. For centrosymmetric crystals, they can have only two phase values, zero or π . For noncentrosymmetric crystals, experience has shown that whereas phase values for the general reflections can have any value in the range from $-\pi$ to $+\pi$, it is usually sufficient to use only four possible values for the symbols spaced $\pi/2$ apart. One of the virtues of using symbols is that, as the phase determination develops, relationships develop among the symbols reducing the number to be assigned values. Here again, one must proceed with caution so that reliable relationships are distinguished from those that are not. With centrosymmetric crystals, the entire phase determination is carried through before the remaining symbols are given alternative numerical values and tested to see which set yields a Fourier series that makes good chemical sense and reproduces the measured intensities well. For noncentrosymmetric crystals the symbolic addition procedure initially makes use only of (31). After about 100 phases have been evaluated, the remaining symbols are given alternative numerical values after which (32) is applied to further extend the phase set. Again, the correct phase set is one whose Fourier series makes good chemical sense and yields a structure that is in fine agreement with the measured intensities. When a satisfactory result is not obtained, it is appropriate to try an alternative path through the phase determination.

As the application of direct phase determination began to increase during the 1960s and structure determination became more and more a part of research programs, there began to be developed at the end of the 1960s "program packages"—software for determining structures from X-ray diffraction data. Among the ones that are widely used, alternative numerical values have been used instead of symbolic phases in the case of noncentrosymmetric crystals, although there are some programs that retain the use of symbols for such types of crystals. For those programs that use alternative numerical values for phases, large numbers of alternative phase sets are generated by use of the tangent formula (32) and the selection of the most likely solutions is dependent upon the use of an elaborate set of probability formulas, auxiliary formulas, and acceptance criteria. Other computational techniques have also evolved. For example, random sets of phases have been used as starting sets to be refined by application of the tangent formula 32. By considering large numbers of alternative starting sets, it is often practicable to obtain a correct answer, although a large amount of computing is involved. There are also special programs for specific purposes such as the development of a structural fragment into a complete structure.

The computer programs are quite successful with centrosymmetric crystals and also do fairly well with noncentrosymmetric crystals having up to 100 independent (non-hydrogen) atoms to be placed in the unit cell. On occasion an answer will not be forthcoming from the use of a program package. In that case, crystallographers may pursue the problem with special techniques and the application of insights and acumen.

There have been many attempts over the years to enhance the power and facility of procedures for structure analysis by including a variety of additional phase-determining formulas. Some of the simpler formulas that arose from the first probabilistic investigation have been of some use. Others have been moderately useful. It is legitimate to question, however, whether they provide a significant step forward for the more complex problems. One type of approach intended to enhance phase determination has concerned formulas for evaluating phase invariants of higher order than the third, derived from probabilistic investigations. We recall that phase determination is principally associated with triple phase invariants. The higher invariants are referred to as quartets, quintets, etc. I believe that the limitations of formulas for individual higher order invariants derive in part from the fact that the model of a crystal, inherent in the probabilistic analyses, differs significantly from actual crystals, some more than others. The values of higher order invariants in real crystals may well be less reliably represented than those of triplet phase invariants by probabilistic formulas for which the crystal model is not a good representative of the actual crystal. If that is true, it may not be possible to develop useful probabilistic formulas for individual higher order phase invariants. Perhaps a useful alternative would be to develop further the practical aspects of mathematical relationships that involve the joint interactions of phase invariants of various orders. Such interactions are, in effect, found in the higher order determinants that provided the conditions that a Fourier series be nonnegative.²²

There is the possibility of developing useful *ab initio* procedures for phase determination based on higher order determinants of the type (18). A determinant of order n would contain all orders of phase invariants up to and including order n . Determinants that contain quasi-normalized structure factors (24) that are predominantly of larger magnitude have special properties. One property is that a determinant with elements of generally larger magnitude is more likely to have a value close to zero than a determinant whose elements have values of smaller magnitude. In addition, determinants having elements of generally larger magnitude approach zero as the order increases long before the rank of the determinant is reached. Use could perhaps be made of such properties to create mathematical systems for phase determination that could be handled by newer theoretical and computational developments.⁴²

Some efforts to use the higher order determinants have already been made. They include the work of Tsoucaris⁴³ that led to the "maximum determinant rule"⁴³⁻⁴⁵ and the generalization of the tangent formula (32).⁴⁶ Several algorithms have been proposed to take advantage of the maximum determinant rule, for example, by Tsoucaris,⁴³ by de Graaff and Vermin,⁴⁷ and by Gelder, de Graaff, and Schenk.⁴⁸

The joint probability distribution for the elements in determinants of any order as well as for elements replaced by triplet phase invariants has been obtained heuristically.⁴⁵ The joint probability distribution for the elements of fourth-order determinants has been obtained rigorously.⁴⁴ There may be

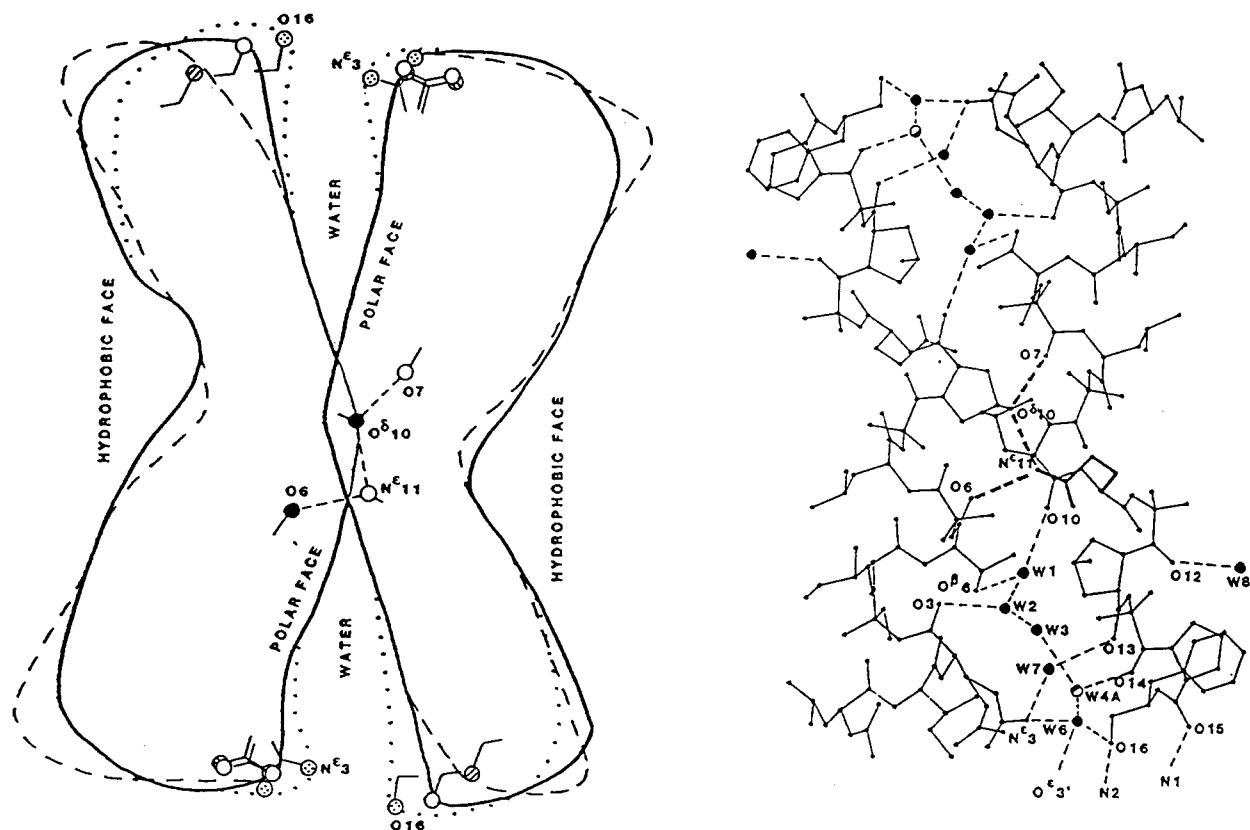


Figure 3. Packing of a pair of bent helical molecules of [Leu¹]zervamicin is shown schematically (left) and in the form of an atomic model derived from an X-ray crystal structure determination (right). The latter corresponds to one of the crystal forms represented by one of the three superimposed envelopes of [Leu¹]zervamicin (solid, dotted and dashed lines) in the schema. The polar faces of the pair of helices associate in an antiparallel fashion and thereby form a discontinuous water channel. The channel is closed in the middle by interpeptide hydrogen bonds involving backbone carbonyls O(6) and O(7) and side chain moieties O⁴(Hyp¹⁰) and N⁴(Gln¹¹). The heavier dashed lines in the center of the atomic model show the interpeptide hydrogen bonds. In a more recent study, about 20% of the water channels were found to be open as a consequence of a rotation of the side chain of Gln¹¹ and the subsequent formation of intrapeptide hydrogen bonds.

worthwhile opportunities involving these probability distributions yet to be exploited.

A problem in numerical analysis that merits further study is the global minimum problem. A useful solution to this problem could find a valuable role in facilitating the application of much of the analytical mathematics mentioned so far. For one thing, a successful program for obtaining global minima could be quite helpful in future applications of formulas for phase determination or direct determination of atomic positions. A new approach to the global minimum problem has been described.⁴² It concerns the use of least-squares minimization. The objective is to find suitable modified minimization functions that preserve the correct global minima but possess a relatively small number of false minima. A suitable procedure involves the alteration of the minimization function in various ways that still preserve the global minima but change the shape of the functions in such a way that the false minima change. Under such circumstances, the least-squares calculation may permit the system to reach a global minimum. Success can be measured in terms of the number of trial starting values for the variables required before reaching a global minimum, and the time required for each trial.

Many mathematical systems could benefit from such a development. In structure determination these would include the higher order determinants that have already been discussed, a number of algebraic systems, some of which have probabilistic origins, and the original symbolic addition procedure because much larger numbers of symbols could be used and readily evaluated. The published work to date⁴² may already have the capacity to facilitate the use of some of the above analytical tools.

Because knowledge of atomic arrangements facilitates the progress of research in many scientific disciplines, structural research has played a valuable role in scientific progress. Many areas of chemistry include structure determination as a part of the research regime. Applications include the structural analysis of substances such as natural products, reaction intermediates, photorearrangement products, and the final products of synthetic or reaction processes. Such information relates structure to function, clarifies synthetic pathways and reaction mechanisms, establishes the conformation of flexible molecules, and provides the geometric parameters for calculations in theoretical chemistry.

A rather interesting example of structural research that has the potential to provide structural insights, on the atomic scale, for the gating mechanism associated with the transport of ions across cell membranes, has been carried out recently.⁴⁹ Peptides can penetrate the lipid bilayers that form cell membranes and form channels through which ions may flow. Gating mechanisms within the channels have been postulated because of the observation that ions flow in response to a potential difference across a membrane and do not do so otherwise. Certain peptides are known to form channels that have gating mechanisms in cell membranes. Structural research on such peptides has concerned not only the three-dimensional atomic arrangement but also the nature of their packing in the crystalline state. This research has revealed structural features that could provide a gating mechanism and, in addition, an assembly of helical peptides that could form a polar channel in cell membranes. In a study of Leu-zervamicin initiated in 1990, the manner in which an ion channel could be formed by a membrane-active antibiotic has

been determined for the first time.⁴⁷ The structural details showed how bent α -helices can assemble to form a polar channel and the possible site and action of a gating mechanism at atomic resolution, Figure 3.

Ion transport across cell membranes is a fundamental process of living organisms. It is one of a myriad of subjects for the application of structural research that are of interest to industry, academia, and other research institutions.

In this historical outline of some of the main contributions of mathematics to structural research in the areas of electron diffraction by gases and X-ray diffraction by crystals, it is seen that these areas have benefited considerably by the application of mathematics. The same is true of other areas of structural research. It is also so that improvements in the mathematical techniques bode well for the future. In view of the broad implications of structural research to many fields of science, it may be said that structural research is an example of a discipline in which those who enjoy the application of mathematics to scientific problems can also enjoy the broad fruitfulness of the results.

The theory described covers a period that started over 40 years ago. It offers, however, many new and untried opportunities. The great developments in computing should facilitate the pursuit of such opportunities.

ACKNOWLEDGMENT

The work of the author referenced in this article has been in the main supported by the Office of Naval Research.

REFERENCES AND NOTES

- (1) Toeplitz, O. *Rend. Circ. Mat. Palermo* **1911**, 191.
- (2) Finbak, C. *Avh. Nor. Vidensk.-Acad. Oslo, Mat.-Naturvidensk. Kl.* **1937**, No. 13.
- (3) Debye, P. P. *Phys. Z.* **1939**, 40, 66, 404.
- (4) Debye, P. *J. Chem. Phys.* **1941**, 9, 55.
- (5) Karle, J.; Karle, I. L. *J. Chem. Phys.* **1947**, 15, 764.
- (6) Karle, J. In *Determination of Organic Structures by Physical Methods*; Nachod, F. C., Zuckerman, J. J., Eds.; Academic Press: New York, 1973; Vol. 5, pp 1-74.
- (7) Degard, C. *Bull. Soc. R. Sci. Liege* **1937**, 12, 383.
- (8) Schomaker, V. Meeting of the American Chemical Society, Baltimore, MD, 1939.
- (9) Karle, J.; Karle, I. L. *Am. Mineral.* **1948**, 33, 767.
- (10) Karle, I. L.; Karle, J. *J. Chem. Phys.* **1949**, 17, 1052.
- (11) Debye, P. *Math. Ann.* **1909**, 67, 535.
- (12) Karle, J. *J. Chem. Phys.* **1954**, 22, 1246.
- (13) Karle, J. *J. Chem. Phys.* **1973**, 59, 3659.
- (14) Karle, I. L.; Karle, J. *J. Chem. Phys.* **1950**, 18, 963.
- (15) Ainsworth, J.; Karle, J. *J. Chem. Phys.* **1952**, 20, 425.
- (16) *Stereochemical Applications of Gas-phase Electron Diffraction*, Parts A and B; Hargittai, I., Hargittai, M., Eds., VCH Verlagsgesellschaft: Weinheim, 1988.
- (17) Patterson, A. L. *Phys. Rev.* **1934**, 46, 372.
- (18) Patterson, A. L. *Z. Kristallogr.* **1935**, 90, 517.
- (19) Harker, D.; Kasper, J. S. *Acta Crystallogr.* **1948**, 1, 70.
- (20) Kasper, J. S.; Lucht, C. M.; Harker, D. *Acta Crystallogr.* **1950**, 3, 436.
- (21) Gillis, J. *Acta Crystallogr.* **1948**, 1, 174.
- (22) Karle, J.; Hauptman, H. *Acta Crystallogr.* **1950**, 3, 181.
- (23) Karle, J. *Acta Crystallogr.* **1971**, B27, 2063.
- (24) Hauptman, H.; Karle, J. *American Crystallographic Association Monograph*, No. 3; Polycrystal Book Service: Dayton, OH, 1953.
- (25) Hauptman, H.; Karle, J. *Acta Crystallogr.* **1956**, 9, 45.
- (26) Hauptman, H.; Karle, J. *Acta Crystallogr.* **1959**, 12, 93.
- (27) Karle, J.; Hauptman, H. *Acta Crystallogr.* **1961**, 14, 217.
- (28) Karle, J. *Comput. Math. Appl.* **1988**, 16, 397.
- (29) Karle, I. L.; Karle, J. *Acta Crystallogr.* **1964**, 17, 835.
- (30) Stewart, J. M.; Karle, J. *Acta Crystallogr.* **1976**, A32, 1005.
- (31) Karle, J.; Hauptman, H.; Christ, C. L. *Acta Crystallogr.* **1958**, 11, 757.
- (32) Wilson, A. J. C. *Nature (London)* **1942**, 150, 151.
- (33) Wilson, A. J. C. *Acta Crystallogr.* **1949**, 2, 318.
- (34) Woolfson, M. M. *Acta Crystallogr.* **1954**, 7, 61.
- (35) Cochran, W.; Woolfson, M. M. *Acta Crystallogr.* **1955**, 8, 1.
- (36) Karle, J.; Karle, I. L. *Acta Crystallogr.* **1966**, 21, 849.
- (37) Watson, G. N. *Theory of Bessel Functions*; Cambridge University Press: Cambridge, U.K., 1945.
- (38) Cochran, W. *Acta Crystallogr.* **1955**, 8, 473.
- (39) Karle, J.; Hauptman, H. *Acta Crystallogr.* **1956**, 9, 635.
- (40) Karle, I. L.; Karle, J. *Acta Crystallogr.* **1963**, 16, 969.
- (41) Karle, J. In *Advances in Chemical Physics*; Prigogine, I., Rice, S. A., Eds.; Interscience: London, 1969; Vol. XVI, pp 131-222.
- (42) Karle, J. *Proc. Natl. Acad. Sci. U.S.A.* **1991**, 88, 10099.
- (43) Tsoucaris, G. *Acta Crystallogr.* **1970**, A26, 492.
- (44) Heinerman, J. J. L.; Kroon, J.; Krabbendam, H. *Acta Crystallogr.* **1979**, A35, 105.
- (45) Karle, J. *Proc. Natl. Acad. Sci. U.S.A.* **1978**, 75, 2545.
- (46) Karle, J. *Acta Crystallogr.* **1971**, B27, 2063.
- (47) de Graaff, R. A. G.; Vermin, W. J. *Acta Crystallogr.* **1982**, A38, 464.
- (48) de Gelder, R.; de Graaff, R. A. G.; Schenk, H. *Acta Crystallogr.* **1993**, A49, 287.
- (49) Karle, I. L.; Flippen-Anderson, J. L.; Agarwalla, S.; Balaram, P. *Proc. Natl. Acad. Sci. U.S.A.* **1991**, 88, 5307.