17.X-7 *DIRECT METHODS OUTSIDE TRADITIONAL FIELD* *By Fan Hai-fu, Institute of Physics, Academia Sinica, Beijing, China.*

After 40 years of development, direct methods are entering into new fields of application. New procedures are under examination in the Institute of Physics in Beijing.
1. In protein crystallography, the combination of direct methods with isomorphous replacement or anomalous scattering data may play an important role in the near future. The method proposed in our group [Fan (1965) Acta Phys. Sin. 21, 1114; Fan, Han, Qian & Yao (1984) Acta Cryst. A 40, 489; Fan & Gu (1985) Acta Cryst. A41, 280.] has been tested with experimental protein data in the one-wavelength anomalous scattering case yielding a thousand of initial phases with average error of about 40 degrees.
2. In high resolution electron microscopy, direct methods may become a powerful tool of image processing. It has been proved by simulation that direct methods are useful in image deconvolution and resolution enhancement [Fan, Zhong, Zheng & Li (1985) Acta Cryst. A41, 163; Han, Fan & Li (1986) Acta Cryst. A42, 353; Liu, Fan & Zheng (1986) IUCr CSM Meeting, Beijing, China.].
3. Modulated structures are important in both solid state physics and structural chemistry. However there was no straightforward way to solve their structure. Recently a direct method has been proposed and used successfully to solve the phase problem of an incommensurate structure.This implies that direct methods will no longer be limited in solving 3-dimensional periodic structures. Hopefully, the method can be extended to solve the phase problem in the determination of quasicrystal structures.

17.X-8 FUTURE METHODOLOGY, SOME PROBABLES AND SOME POSSIBLES. By Jerome Karle, Laboratory for the Structure of Matter, Naval Research Laboratory, Washington, D. C. 20375, U.S.A.

In the area of macromolecular structure determination, there are several developments which indicate the potential for enhanced analytical capability. In addition to the development of tunable, high-intensity sources (synchrotron radiation) and the continuing improvement of computing facilities, the results of theoretical studies of the isomorphous replacement and the anomalous dispersion techniques in recent years imply the potential for enhanced speed and facility in performing structure determinations. An algebraic analysis of multiple-wavelength anomalous dispersion, for example, has resulted in a set of simultaneous equations that are both exact and linear (J. Karle, Int. J. Quantum Chem. Symp., 1980, 7, 357-367). The exact algebraic analysis is valid for any number of anomalous scatterers and any variety of types of anomalous scatterer. The unknown quantities, which do not vary with wavelength, are composed of intensities, phase differences, or a combination of the two. The part that varies with wavelength occurs only as coefficients of the unknowns. The unknown phases and intensities are those that would be obtained from individual types of atoms as if each type were present in isolation from the rest. Knowledge of the intensities for the structure formed by a particular type of atom can facilitate the determination of the structure formed by this particular type of atom. Once the structure is known for any of the types of atom present, the entire structure can be readily determined. The determination of the structure of anomalous scatterers may not always be successful. In those cases, the structures of nonanomalously scattering atoms may, perhaps be obtainable from known values for triplet phase invariants of the type $\phi_h+\phi_k+\phi_{-h-k}$. Values for triplet phase invariants may be obtained from the exact algebraic analysis (J. Karle, Acta Cryst., 1984, A40, 526-531). Evaluations of triplet phase invariants have also been made from use of probability theory (H. Hauptman, Acta Cryst., 1982, A38, 289-294, 632-641; C. Giacovazzo, Acta Cryst., 1983, A39, 585-592; S. Fortier, N. J. Moore and M. E, Fraser, Acta Cryst., 1985, A41, 571-577) and alternative algebraic analyses that make use of special mathematical and physical properties of the isomorphous replacement and the anomalous dispersion phenomena (J. Karle, Acta Cryst., 1983, A39, 800-805; J. Karle, Acta Cryst., 1984 A40, 4-11, 366-373, 374-379; J. Karle, Acta Cryst., 1985, A41, 182-189). It has been further shown that by use of algebraic analysis essentially unique values can be obtained with one-wavelength anomalous dispersion data for the 2-phase invariant (J. Karle, Acta Cryst., 1985, A41, 387-394). This analysis was limited to the case of one predominant type of anomalous scatterer in a one-wavelength experiment. Another investigation has shown that with use of values obtained for 2-phase invariants in one-wavelength anomalous dispersion experiments and single isomorphous replacement experiments, knowledge of the structures of the anomalously scattering atoms or heavy atoms could afford a large number of phase values for initiation of a structure determination of a native or nonanomalously scattering structure (J. Karle, Acta Cryst., 1986, A42, 246-253). This could also provide a possible strategy for use of triplet phase invariants.

17.X-9 THE ROLE OF STATISTICS IN STRUCTURE REFINEMENT. E. Prince, Institute for Materials Science and Engineering, National Bureau of Standards, Gaithersburg, MD 20899, U. S. A.

Mathematical statistics provides some powerful tools for the extraction of information from experimental data. These tools are, however, frequently misused for purposes for which they were never intended, and for which they are ill suited, with the result that there is much confusion when a result is stated in statistical terms. In the context of crystallography, if structure factors, amplitude and phase, were known throughout reciprocal space, all possible information about the crystal would be known, but the only quantites that can be measured, with inevitably limited precision, are intensities at a finite number of lattice points within a bounded region of reciprocal space. The problem, then, is to estimate as precisely as possible the values of all structure factors given the finite set of measured intensities, or, mathematically, to find the means and variances of a set of conditional probability density functions (pdfs) of structure factors, given a set of intensities. Because crystals are composed of atoms, each structure factor can be expressed in terms of a model that is a function of a finite set of atom parameters, $F(\mathbf{h}) = M(\mathbf{h},\mathbf{x})$, so that if the elements of $\mathbf{x}$ can be estimated, the values of all $F(\mathbf{h})$ can be estimated.
An estimator is a function of observations that, under some conditions, approximates an unknown parameter of a population pdf. If the expected value of the estimate is equal to the parameter, the estimator is unbiased. If the expected values of a set of observations, $\mathbf{y}$, are linear functions of the parameters, so that $\langle\mathbf{y}\rangle = \mathbf{Ax}$, where $\mathbf{A}$ is some matrix, and $\mathbf{W}$ is a positive definite weight matrix, usually, but not necessarily, diagonal, then the least squares estimate, $\hat{\mathbf{x}} = (\mathbf{A}^T\mathbf{WA})^{-1}\mathbf{A}^T\mathbf{Wy}$ is an unbiased estimate of $\mathbf{x}$. If, in addition, $\mathbf{W} = \mathbf{V}^{-1}$, where $\mathbf{V}$ is the variance-covariance matrix for the elements of

y, then the variances of the parameter estimates are minimum. Only for this choice of weights are the diagonal elements of $(A^TWA)^{-1}$ unbiased estimates of the variances of the parameter estimates. A nonlinear model like the structure factor formula may be replaced by a linear approximation. The estimates are then unbiased only to the extent that the linear approximation is valid, but biases can be reduced to arbitrarily small values by sufficiently precise observations. The "observations" may be raw data (as in the Rietveld method), net integrated intensities, or simple functions of the net integrated intensities, provided that the condition $\langle \underline{y}_i \rangle = M_i(\underline{x})$ is maintained. If the function is nonlinear, such as the extraction of a square root, some care must be taken to ensure this condition, but that care can be rewarded by improved precision in the parameter estimates.

If the model is correct, the properly weighted sum of squared residuals should equal $\underline{n} - \underline{p}$, where $\underline{n}$ is the number of observations and $\underline{p}$ is the number of parameters. Values larger than this are indicators of lack of fit. The common practice of assuming that these values are due to an incorrect scaling of weights that have the correct relative values is questionable, and inferences drawn from it should be viewed with caution. The positive square root of an estimated variance is an estimated standard deviation, or e. s. d. It is an indicator only of precision, which sets a lower limit to the uncertainty in the correspondence between estimated parameters and nature's values when the model is exactly correct. Statistical analysis can tell whether the model plausibly explains the observations. It cannot rule out the existence of systematic effects that bias the parameter estimates without contributing to lack of fit, nor can it rule out the existence of an entirely different model that would explain the observations as well or better. It therefore tells nothing about the actual accuracy of a measurement.

17.X-10
THE EFFECTS OF WEIGHTING SCHEMES ON ESTIMATED STANDARD DEVIATIONS AND ON ACCURACY.

By B.E. Robertson, Department of Physics and Astronomy, University of Regina, Regina, Sask., S4S 0A2, Canada.

Weights used in crystallographic least squares should be the reciprocals of the variances of the observations, but in a real experiment the variances of the observations are not known, other than their contribution from the Poisson counting statistics. Other errors which do not contribute to the difference between the estimate and the best value of a quantity used to describe a model (ie, they are not systematic errors) should also be used to determine the variances and the least-squares weights. How this should be done will depend on one's knowledge of the nature of other errors, and on the goals of the experiment. This is done by; (i) adding a contribution to the variance derived from the extent to which the sample variance of the intensities of the standard reflections exceeds the experimental variance (ii) replacing the experimentally determined variance with the sample variance obtained from the consistency of observations which should be identical according to the model, or adding to the experimentally determined variance a quantity such that the average modified experimental variance and sample variances are equal, or (iii) adding quantities to the variance such that the average value of Δ^2/σ^2 approaches (n-m)/m (where $\Delta = ||F_o| - |F_c||$ or $|I_o - I_c|$, σ^2 is the relevant variance, n is the number of observations and m is the number of least-square variables) for any groups of observations which may be averaged. All of these procedures may introduce systematic error into the weighting scheme. They normally lead to enhanced precision. However, if the values of the parameters which are to be estimated by the experiment are influenced by any of the systematic errors which are present, then the incorporation of systematic error into the least-squares weights is a "feed-back" process which may enhance or diminish the influence of systematic error. Furthermore, the enhanced precision obtained with modified weights may not imply enhanced accuracy.

In order to determine the effects of weight modification on the parameters of interest in routine structure determinations, we assume that an independent measure of the accuracy of the estimates of parameters from a refinement is their consistency; ie, more accurate refinements will lead to smaller sample variances among molecular parameters that are assumed to be exactly equal. This includes (I) chemically equivalent bonds which are not constrained to be equivalent by symmetry, (II) bonds in molecules in structures with more than one molecule per asymmetric unit, (III) multiple determinations of the same crystal structure. Consistency so determined has been used to investigate such questions as (a) the nature of the feedback from the contribution of systematic errors to the weights on accuracy, (b) the relation between accuracy and precision for various weighting schemes, (c) the effects of including weak reflections, (d) the merits of refinement on $|F|$ or I, etc.

17.X-11 THE BAYESIAN VIEWPOINT OF STATISTICAL PROCEDURES IN CRYSTALLOGRAPHY. by H. D. Flack, Laboratoire de Cristallographie, University of Geneva, 24 quai Ernest Ansermet, CH-1211 Genève 4, Switzerland.

From a Bayesian viewpoint, probability is subjective. A probability density function provides a measure of the crystallographer's degree of belief in the value of a random variable, a hypothesis, the estimate of physical parameters or a physical model used to interpret experimental results. Objectivity is regarded as being illusory - experimental observations are always interpreted through a model and one always has some prior idea of the numerical values entering into the model. Consider examples such as: weighting schemes to take account of systematic errors in the data or model, multiplication of e.s.d.'s by the goodness of fit value, averaging of symmetry-equivalent reflections, use of restraints (soft constraints, pseudo-observations), robust-resistant refinement. Such procedures are without any theoretical foundation in Statistics when viewed with the classical, Frequentist, notion of probability. On the other hand the Bayesian viewpoint does provide a clear framework within which to elaborate and criticize the above-mentionned procedures.

The Bayesian Three-Stage Model is a particularly fruitful