template around its centre of mass using the quasi-uniformly distributed three Eulerian angles [equation (2.5.7.17) with $\Delta\psi = \delta\theta$]. However, in application to tomography the angular step $\delta\theta$ can be relatively large, resulting in a much smaller number of templates than in two dimensions, the reason being the rather low resolution of typical electron tomograms (not exceeding 50 Å). Next, a brute-force 3D cross-correlation search with all templates is performed (Frangakis *et al.*, 2002). After windowing out 3D subvolumes containing putative complexes, subsequent averaging and classification can be performed.

Cryo-EM is a unique structural technique in its ability to detect conformational variability of large molecular assemblies within one sample that may contain a mixture of complexes in various conformational states. In addition to the expected conformational heterogeneity of the assemblies, due to fluctuations of the structure around the ground state one can expect to capture molecules in different functional states, especially if the binding of a ligand induces a conformational change in the macromolecular assembly. Therefore, a data set of images from an EM experiment must be interpreted as a mixture of projections from similar but not identical structures. The analysis of the extent of the resulting variability requires the calculation of the real-space distribution of 3D variance/covariance in macromolecules reconstructed from a set of their projections. The problem is difficult, as there is no clear relation between the variance in sets of projections that have the same angular direction and the variance of the 3D structure calculated from these projections. Penczek, Chao *et al.* (2006) proposed calculating the variance in the 3D mass distribution of the structure using a statistical bootstrap resampling technique, in which a new set of projections is selected with replacements from the available whole set of N projections. In the new set, some of the original projections will appear more than once, while others will be omitted. This selection process is repeated a number of times and for each new set of projections the corresponding 3D volume is calculated. Next, the voxel-by-voxel bootstrap variance $\sigma_{\rm B}^2$ of the resulting set of volumes is calculated. The target variance is obtained using a relationship between the variance of arithmetic means for sampling with replacements and the sample variance,

$$\sigma^2 = N\sigma_{\rm B}^2. \qquad (2.5.7.29)$$

The estimated structure-variance map can be used for (i) detection of different functional states (for example, those characterized by binding of a ligand) and subsequent classification of the data set into homogeneous groups (Penczek, Frank & Spahn, 2006*a*), (ii) analysis of the significance of small details in 3D reconstructions, (iii) analysis of the significance of details in difference maps, and (iv) docking of known structural domains into EM density maps.

The bootstrap technique also leads to the analysis of conformational modes of macromolecular complexes, and this is due to the fact that the covariance matrix of the structure can be directly calculated from the bootstrap volumes. The covariance matrix obtained this way would be very large. One possibility is to calculate only correlation coefficients between regions of interest that have large variance (Penczek, Chao *et al.*, 2006). Another possibility is to use the iterative Lanczos technique (Parlett, 1980) and calculate eigenvolumes directly from bootstrap volumes without forming the covariance matrix. These eigenvolumes are related to conformational modes of the molecule, as captured by the projection data of the sample (Penczek, Frank & Spahn, 2006*b*). Thus, this direct relation to the actual cryo-EM projection data positively distinguishes this approach from other techniques in which conformations are postulated based on flexible models of the EM map (Ming *et al.*, 2002; Mitra *et al.*, 2005).

2.5.8. Direct phase determination in electron crystallography

By D. L. Dorset

2.5.8.1. Problems with 'traditional' phasing techniques

The concept of using experimental electron-diffraction intensities for quantitative crystal structure analyses has already been presented in Section 2.5.4. Another aspect of quantitative structure analysis, employing high-resolution images, has been presented in Sections 2.5.5 to 2.5.7. That is to say, electron micrographs can be regarded as an independent source of crystallographic phases.

Before direct methods (Chapter 2.2) were developed as the standard technique for structure determination in small-molecule X-ray crystallography, there were two principal approaches to solving the crystallographic phase problem. First, 'trial and error' was used, finding some means to construct a reasonable model for the crystal structure *a priori*, *e.g.* by matching symmetry properties shared by the point group of the molecule or atomic cluster and the unit-cell space group. Secondly, the autocorrelation function of the crystal, known as the Patterson function (Chapter 2.3), was calculated (by the direct Fourier transform of the available intensity data) to locate salient interatomic vectors within the unit cell.

The same techniques had been used for electron-diffraction structure analysis (nowadays known as *electron crystallography*). In fact, advocacy of the first method persists. Because of the perturbations of diffracted intensities by multiple-beam dynamical scattering (Chapter 5.2), it has often been suggested that trial and error be used to construct the scattering model for the unit crystal in order to test its convergence to observed data after simulation of the scattering events through the crystal. This indirect approach assumes that no information about the crystal structure can be obtained directly from observed intensity data. Under more favourable scattering conditions nearer to the kinematical approximation, *i.e.* for experimental data from thin crystals made up of light atoms, trial-and-error modelling, simultaneously minimizing an atom–atom nonbonded potential function with the crystallographic residual, has enjoyed widespread use in electron crystallography, especially for the determination of linear polymer structures (Brisse, 1989; Pérez & Chanzy, 1989).

Interpretation of Patterson maps has also been important for structure analysis in electron crystallography. Applications have been discussed by Vainshtein (1964), Zvyagin (1967) and Dorset (1994*a*). In face of the dynamical scattering effects for electron scattering from heavy-atom crystals realized later (*e.g.* Cowley & Moodie, 1959), attempts had also been made to modify this autocorrelation function by using a power series in $|F_{\mathbf{h}}|$ to sharpen the peaks (Cowley, 1956). (Here $F_{\mathbf{h}} \equiv \Phi_{\mathbf{h}}$, replacing the notation for the kinematical electron-diffraction structure factor employed in Section 2.5.4.) More recently, Vincent and coworkers have selected first-order-Laue-zone data from inorganics to minimize the effect of dynamical scattering on the interpretability of their Patterson maps (Vincent & Exelby, 1991, 1993; Vincent & Midgley, 1994). Vainshtein & Klechkovskaya (1993) have also reported use of the Patterson function to solve the crystal structure of a lead soap from texture electron-diffraction intensity data.

It is apparent that trial-and-error techniques are most appropriate for *ab initio* structure analysis when the underlying crystal structures are reasonably easy to model. The requisite positioning of molecular (or atomic) groups within the unit cell may be facilitated by finding atoms that fit a special symmetry position [see *IT* A (2005)]. Alternatively, it is helpful to know the molecular orientation within the unit cell (*e.g.* provided by the Patterson function) to allow the model to be positioned for a conformational or translational search. [Examples would include

the polymer-structure analyses cited above, as well as the layer-packing analysis of some phospholipids (Dorset, 1987).] While attempts at *ab initio* modelling of three-dimensional crystal structures, by searching an n-dimensional parameter space and seeking a global internal energy minimum, has remained an active research area, most success so far seems to have been realized with the prediction of two-dimensional layers (Scaringe, 1992). In general, for complicated unit cells, determination of a structure by trial and error is very difficult unless adequate constraints can be placed on the search.

Although Patterson techniques have been very useful in electron crystallography, there are also inherent difficulties in their use, particularly for locating heavy atoms. As will be appreciated from comparison of scattering-factor tables for X-rays [*IT* C (2004), Chapter 6.1] with those for electrons [*IT* C (2004), Chapter 4.3], the relative values of the electron form factors are more compressed with respect to atomic number than are those for X-ray scattering. As discussed in Chapter 2.3, it is desirable that the ratio of summed scattering-factor terms, $r = \sum_{\text{heavy}} Z^2 / \sum_{\text{light}} Z^2$, where Z is the scattering-factor value at $\sin\theta/\lambda = 0$, be near unity. A practical comparison would be the value of r for copper (DL-alaninate) solved from electron-diffraction data by Vainshtein *et al.* (1971). For electron diffraction, $r = 0.47$ compared to the value 2.36 for X-ray diffraction. Orientation of salient structural features, such as chains and rings, would be equally useful for light-atom moieties in electron or X-ray crystallography with Patterson techniques. As structures become more complicated, interpretation of Patterson maps becomes more and more difficult unless an automated search can be carried out against a known structural fragment (Chapter 2.3).

2.5.8.2. Direct phase determination from electron micrographs

The 'direct method' most familiar to the electron microscopist is the high-resolution electron micrograph of a crystalline lattice. Retrieval of an average structure from such a micrograph assumes that the experimental image conforms adequately to the 'weak phase object' approximation, as discussed in Section 2.5.5. If this is so, the use of image-averaging techniques, *e.g.* Fourier filtration or correlational alignment, will allow the unit-cell contents to be visualized after the electron-microscope phase contrast transfer function is deconvoluted from the average image, also discussed in Section 2.5.5. Image analyses can also be extended to three dimensions, as discussed in Section 2.5.6, basically by employing tomographic reconstruction techniques to combine information from the several tilt projections taken from the crystalline object. The potential distribution of the unit cell to the resolution of the imaging experiment can then be used, *via* the Fourier transform, to obtain crystallographic phases for the electron-diffraction amplitudes recorded at the same resolution. This method for phase determination has been the mainstay of protein electron crystallography.

Once a set of phases is obtained from the Fourier transform of the deconvoluted image, they must, however, be referred to an allowed crystallographic origin. For many crystallographic space groups, this choice of origin may coincide with the location of a major symmetry element in the unit cell [see *IT* A (2005)]. Hence, since the Fourier transform of translation is a phase term, if an image shift $[\delta(\mathbf{r} + \mathbf{r}_0)]$ is required to translate the origin of the repeating mass unit $\varphi(\mathbf{r})$ from the arbitrary position in the image to a specific site allowed by the space group,

$$\mathrm{g}(\mathbf{r}) = \varphi(\mathbf{r}) \otimes \delta(\mathbf{r} + \mathbf{r}_0) = \varphi(\mathbf{r} + \mathbf{r}_0),$$

where the operation '$\otimes$' denotes convolution. The Fourier transform of this shifted density function will be

$$G(\mathbf{s}) = F(\mathbf{s}) \exp(2\pi i \mathbf{s} \cdot \mathbf{r}_0) = |F(\mathbf{s})| \exp[i(\phi_s + 2\pi i \mathbf{s} \cdot \mathbf{r}_0)].$$

In addition to the crystallographic phases ϕ_s, it will, therefore, be necessary to find the additional phase-shift term $2\pi i \mathbf{s} \cdot \mathbf{r}_0$ that will access an allowed unit-cell origin. Such origin searches are carried out automatically by some commercial image-averaging computer-software packages.

In addition to applications to thin protein crystals (*e.g.* Henderson *et al.*, 1990; Jap *et al.*, 1991; Kühlbrandt *et al.*, 1994), there are numerous examples of molecular crystals that have been imaged to a resolution of 3–4 Å, many of which have been discussed by Fryer (1993). For π-delocalized compounds, which are the most stable in the electron beam against radiation damage, the best results (2 Å resolution) have been obtained at 500 kV from copper perchlorophthalocyanine epitaxically crystallized onto KCl. As shown by Uyeda *et al.* (1978–1979), the averaged images clearly depict the positions of the heavy Cu and Cl atoms, while the positions of the light atoms in the organic residue are not resolved. (The utility of image-derived phases as a basis set for phase extension will be discussed below.) A number of aromatic polymer crystals have also been imaged to about 3 Å resolution, as reviewed recently (Tsuji, 1989; Dorset, 1994*b*).

Aliphatic molecular crystals are much more difficult to study because of their increased radiation sensitivity. Nevertheless, monolamellar crystals of the paraffin n-$C_{44}H_{90}$ have been imaged to 2.5 Å resolution with a liquid-helium cryomicroscope (Zemlin *et al.*, 1985). Similar images have been obtained at room temperature from polyethylene (Revol & Manley, 1986) and also a number of other aliphatic polymer crystals (Revol, 1991).

As noted by J. M. Cowley and J. C. H. Spence in Section 2.5.1, dynamical scattering can pose a significant barrier to the direct interpretation of high-resolution images from many inorganic materials. Nevertheless, with adequate control of experimental conditions (limiting crystal thickness, use of high-voltage electrons) some progress has been made. Pan & Crozier (1993) have described 2.0 Å images from zeolites in terms of the phase-grating approximation. A three-dimensional structural study has been carried out on an aluminosilicate by Wenk *et al.* (1992) with thin samples that conform to the weak-phase-object approximation at the 800 kV used for the imaging experiment. Heavy and light (*e.g.* oxygen) atoms were located in the micrographs in good agreement with an X-ray crystal structure. Heavy-atom positions from electron microscopic and X-ray structure analyses have also been favourably compared for two heavy-metal oxides (Hovmöller *et al.*, 1984; Li & Hovmöller, 1988).

2.5.8.3. Probabilistic estimate of phase invariant sums

Conventional direct phasing techniques, as commonly employed in X-ray crystallography (*e.g.* see Chapter 2.2), have also been used for *ab initio* electron-crystallographic analyses. As in X-ray crystallography, probabilistic estimates of a linear combination of phases (Hauptman & Karle, 1953; Hauptman, 1972) are made after normalized structure factors are calculated *via* electron form factors, *i.e.*

$$|E_{\mathbf{h}}^2| = I_{\text{obs}} / \varepsilon \sum_i f_i^2, \text{ where } \langle |E|^2 \rangle = 1.000.$$

(Here, an overall temperature factor can be found from a Wilson plot. Because of multiple scattering, the value of B may be found occasionally to lie close to 0.0 Å^2.) The phase invariant sums

$$\psi = \phi_{\mathbf{h}_1} + \phi_{\mathbf{h}_2} + \phi_{\mathbf{h}_3} + \ldots$$

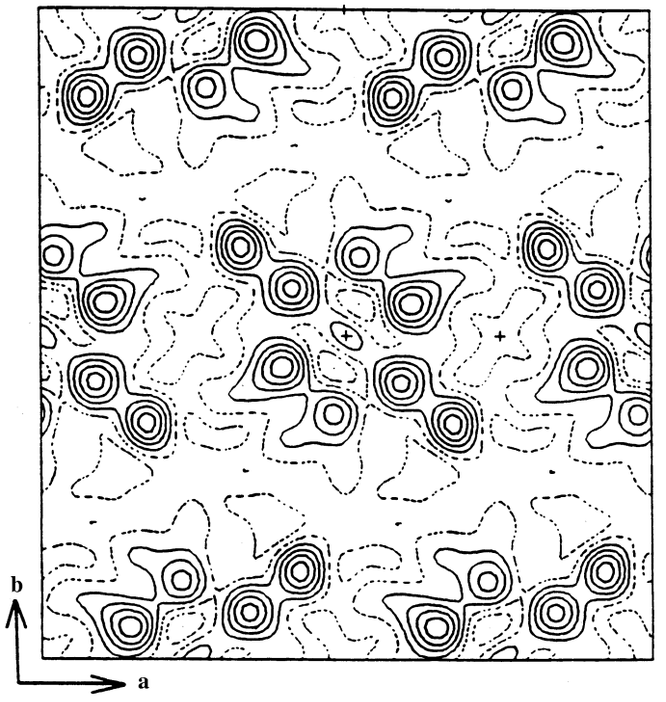

Fig. 2.5.8.1. Potential map for diketopiperazine ([001] projection) after a direct phase determination with texture electron-diffraction intensity data obtained originally by Vainshtein (1955).

can be particularly effective for structure analysis. Of particular importance historically have been the Σ_2-triple invariants where $\mathbf{h}_1 + \mathbf{h}_2 + \mathbf{h}_3 = 0$ and $\mathbf{h}_1 \neq \mathbf{h}_2 \neq \mathbf{h}_3$. The probability of predicting $\psi = 0$ is directly related to the value of

$$A = (2\sigma_3/\sigma_2^{3/2})|E_{\mathbf{h}_1} E_{\mathbf{h}_2} E_{\mathbf{h}_3}|,$$

where $\sigma_{\mathbf{h}} = \sum_{j=1}^{N} Z_j^n$ and Z is the value of the scattering factor at $\sin\theta/\lambda = 0$. Thus, the values of the phases are related to the measured structure factors, just as they are found to be in X-ray crystallography. The normalization described above imposes the point-atom structure (compensating for the fall-off of an approximately Gaussian form factor) often assumed in deriving the joint probability distributions. Especially for van der Waals structures, the constraint of positivity also holds in electron crystallography. (It is also quite useful for charged atoms so long as the reflections are not measured at very low angles.) Other useful phase invariant sums are the Σ_1 triples, where $\mathbf{h}_1 = \mathbf{h}_2 = -1/2\mathbf{h}_3$, and the quartets, where $\mathbf{h}_1 + \mathbf{h}_2 + \mathbf{h}_3 + \mathbf{h}_4 = 0$ and $\mathbf{h}_1 \neq \mathbf{h}_2 \neq \mathbf{h}_3 \neq \mathbf{h}_4$. The prediction of a correct phase for an invariant is related in each case to the normalized structure-factor magnitudes.

The procedure for phase determination, therefore, is identical to the one used in X-ray crystallography (see Chapter 2.2). Using vectorial combinations of Miller indices, one generates triple and quartet invariants from available measured data and ranks them according to parameters such as A, defined above, which, as shown in Chapter 2.2, are arguments of the Cochran formula. The invariants are thus listed in order of their reliability. This, in fact, generates a set of simultaneous equations in crystallographic phase. In order to begin solving these equations, it is permissible to define arbitrarily the phase values of a limited number of reflections (three for a three-dimensional primitive unit cell) for reflections with Miller-index parity $hkl \neq ggg$ and $\sum_i h_i k_i l_i \neq ggg$, where g is an even number. This defines the origin of a unit cell. For noncentrosymmetric unit cells, the condition for defining the origin, which depends on the space group, is somewhat more complicated and an enantiomorph-defining reflection must be added.

In the evaluation of phase-invariant sums above a certain probability threshold, phase values are determined algebraically after origin (and enantiomorph) definition until a large enough set is obtained to permit calculation of an interpretable potential map (*i.e.* where atomic positions can be seen). There may be a few invariant phase sums above this threshold probability value which are incorrectly predicted, leading either to false phase assignments or at least to phase assignments inconsistent with those found from other invariants. A small number of such errors can generally be tolerated. Another problem arises when an insufficient quantity of new phase values is assigned directly from the phase invariants after the origin-defining phases are defined. This difficulty may occur for small data sets, for example. If this is the case, it is possible that a new reflection of proper index parity can be used to define the origin. Alternatively, $\phi_n = a, b, c \ldots$ algebraic unknowns can be used to establish the phase linkage among certain reflections. If the structure is centrosymmetric, and when enough reflections are given at least symbolic phase assignments, 2^n maps are calculated and the correct structure is identified by inspection of the potential maps. When all goes well in this so-called 'symbolic addition' procedure, the symbols are uniquely determined and there is no need to calculate more than a single map. If algebraic values are retained for certain phases because of limited vectorial connections in the data set, then a few maps may need to be generated so that the correct structure can be identified using the chemical knowledge of the investigator. The atomic positions identified can then be used to calculate phases for all observed data (*via* the structure-factor calculation) and the structure can be refined by Fourier (or, sometimes, least-squares) techniques to minimize the crystallographic R factor.

The first actual application of direct phasing techniques to experimental electron-diffraction data, based on symbolic addition procedures, was to two methylene subcell structures (an *n*-paraffin and a phospholipid; Dorset & Hauptman, 1976). Since then, evaluation of phase invariants has led to numerous other structures. For example, early texture electron-diffraction data sets obtained in Moscow (Vainshtein, 1964) were shown to be suitable for direct analysis. The structure of diketopiperazine (Dorset, 1991*a*) was determined from these electron-diffraction data (Vainshtein, 1955) when directly determined phases allowed computation of potential maps such as the one shown in Fig. 2.5.8.1. Bond distances and angles are in good agreement with the X-ray structure, particularly after least-squares refinement (Dorset & McCourt, 1994*a*). In addition, the structures of urea (Dorset, 1991*b*), using data published by Lobachev & Vainshtein (1961), paraelectric thiourea (Dorset, 1991*b*), using data published by Dvoryankin & Vainshtein (1960), and three mineral structures (Dorset, 1992*a*), from data published by Zvyagin (1967), have been determined, all using the original texture (or mosaic single-crystal) diffraction data. The most recent determination based on such texture diffraction data is that of basic copper chloride (Voronova & Vainshtein, 1958; Dorset, 1994*c*).

Symbolic addition has also been used to assign phases to selected-area diffraction data. The crystal structure of boric acid (Cowley, 1953) has been redetermined, adding an independent low-temperature analysis (Dorset, 1992*b*). Additionally, a direct structure analysis has been reported for graphite, based on high-voltage intensity data (Ogawa *et al.*, 1994). Two-dimensional data from several polymer structures have also been analysed successfully (Dorset, 1992*c*) as have three-dimensional intensity data (Dorset, 1991*c*,*d*; Dorset & McCourt, 1993).

Phase information from electron micrographs has also been used to aid phase determination by symbolic addition. Examples include the epitaxically oriented paraffins *n*-hexatriacontane (Dorset & Zemlin, 1990), *n*-tritriacontane (Dorset & Zhang, 1991) and a 1:1 solid solution of n-$C_{32}H_{66}$/n-$C_{36}H_{74}$ (Dorset,

1990*a*). Similarly, lamellar electron-diffraction data to *ca* 3 Å resolution from epitaxically oriented phospholipids have been phased by analysis of Σ_1 and Σ_2-triplet invariants (Dorset, 1990*b*, 1991*e*,*f*), in one case combined with values from a 6 Å resolution electron-microscope image (Dorset *et al.*, 1990, 1993). Most recently, such data have been used to determine the layer packing of a phospholipid binary solid solution (Dorset, 1994*d*).

An *ab initio* direct phase analysis was carried out with zonal electron-diffraction data from copper perchlorophthalocyanine. Using intensities from a *ca* 100 Å thick sample collected at 1.2 MeV, the best map from a phase set with symbolic unknowns retrieves the positions of all the heavy atoms, equivalent to the results of the best images (Uyeda *et al.*, 1978–1979). Using these positions to calculate an initial phase set, the positions of the remaining light C, N atoms were found by Fourier refinement so that the final bond distances and angles were in good agreement with those from X-ray structures of similar compounds (Dorset *et al.*, 1991). A similar analysis has been carried out for the perbromo analogue (Dorset *et al.*, 1992). Although dynamical scattering and secondary scattering significantly perturb the observed intensity data, the total molecular structure can be visualized after a Fourier refinement. Most recently, a three-dimensional structure determination was reported for C_{60} buckminsterfullerene based on symbolic addition with results most in accord with a rotationally disordered molecular packing (Dorset & McCourt, 1994*b*).

2.5.8.4. The tangent formula

Given a triple phase relationship

$$\phi_{\mathbf{h}} \simeq \phi_{\mathbf{k}} + \phi_{\mathbf{h}-\mathbf{k}},$$

where $\mathbf{h}$, $\mathbf{k}$ and $\mathbf{h}-\mathbf{k}$ form a vector sum, it is often possible to find a more reliable estimate of $\phi_{\mathbf{h}}$ when all the possible vectorial contributions to it within the observed data set $\mathbf{k}_{\mathbf{r}}$ are considered as an average, *viz:*

$$\phi_{\mathbf{h}} \simeq \langle \phi_{\mathbf{k}} + \phi_{\mathbf{h}-\mathbf{k}} \rangle_{\mathbf{k}_{\mathbf{r}}}.$$

For actual phase determination, this can be formalized as follows. After calculating normalized structure-factor magnitudes $|E_{\mathbf{h}}|$ from the observed $|F_{\mathbf{h}}|$ to generate all possible phase triples within a reasonably high $A_{\mathbf{h}}$ threshold, new phase values can be estimated after origin definition by use of the tangent formula (Karle & Hauptman, 1956):

$$\tan \phi_{\mathbf{h}} = \frac{\sum_{\mathbf{k}_{\mathbf{r}}} W_{\mathbf{h}} |E_{\mathbf{k}}||E_{\mathbf{h}-\mathbf{k}}| \sin(\phi_{\mathbf{k}} + \phi_{\mathbf{h}-\mathbf{k}})}{\sum_{\mathbf{k}_{\mathbf{r}}} W_{\mathbf{h}} |E_{\mathbf{k}}||E_{\mathbf{h}-\mathbf{k}}| \cos(\phi_{\mathbf{k}} + \phi_{\mathbf{h}-\mathbf{k}})}.$$

The reliability of the phase estimate depends on the variance $V(\phi_{\mathbf{h}})$, which is directly related to the magnitude of $\alpha_{\mathbf{h}}$, *i.e.*

$$\alpha_{\mathbf{h}}^2 = \left[\sum_{\mathbf{k}_{\mathbf{r}}} A_{\mathbf{h},\mathbf{k}} \cos(\phi_{\mathbf{k}} + \phi_{\mathbf{h}-\mathbf{k}})\right]^2 + \left[\sum_{\mathbf{k}_{\mathbf{r}}} A_{\mathbf{h},\mathbf{k}} \sin(\phi_{\mathbf{k}} + \phi_{\mathbf{h}-\mathbf{k}})\right]^2;$$

$A_{\mathbf{h},\mathbf{k}}$ is identical to the A value defined in the previous section. In the initial stages of phase determination $\alpha_{\mathbf{h}}$ is replaced by an expectation value α_{E} until enough phases are available to permit its calculation.

The phase solutions indicated by the tangent formula can thus be ranked according to the phase variance and the determination of phases can be made symbolically from the most probable triple-product relationships. This procedure is equivalent to the one described above for the evaluation of phase-invariant sums by symbolic addition. This procedure may allow determination of a large enough basis phase set to produce an interpretable map.

An alternative procedure is to use an automatic version of the tangent formula in a multisolution process. This procedure is described in Chapter 2.2. After origin definition, enough algebraic unknowns are defined (two values if centrosymmetric and four values, cycling through phase quadrants, if noncentrosymmetric) to access as many of the unknown phases as possible. These are used to generate a number of trial phase sets and the likelihood of identifying the correct solution is based on the use of some figure of merit.

Multisolution approaches employing the tangent formula include *MULTAN* (Germain *et al.*, 1971), *QTAN* (Langs & DeTitta, 1975) and *RANTAN* (Yao, 1981). *RANTAN* is a version of *MULTAN* that allows for a larger initial random phase set (with suitable control of weights in the tangent formula). *QTAN* utilizes the $\alpha_{\mathbf{h}_{\mathrm{est}}}$ definition, where

$$\alpha_{\mathbf{h}_{\mathrm{est}}} = \left\{\sum_{\mathbf{k}} A_{\mathbf{h},\mathbf{k}}^2 + 2 \sum_{\mathbf{k} \neq} \sum_{\mathbf{k}'} A_{\mathbf{h},\mathbf{k}} A'_{\mathbf{h},\mathbf{k}} \frac{I_1(A_{\mathbf{h},\mathbf{k}}) I_1(A'_{\mathbf{h},\mathbf{k}})}{I_0(A_{\mathbf{h},\mathbf{k}}) I_0(A'_{\mathbf{h},\mathbf{k}})}\right\}^{1/2},$$

for evaluating the phase variance. (Here I_0, I_1 are modified Bessel functions.) After multiple solutions are generated, it is desirable to locate the structurally most relevant phase sets by some figure of merit. There are many that have been suggested (Chapter 2.2). The most useful figure of merit in *QTAN* has been the NQEST (De Titta *et al.*, 1975) estimate of negative quartet invariants (see Chapter 2.2). More recently, this has been superseded by the minimal function (Hauptman, 1993):

$$R(\phi) = \frac{\sum_{\mathbf{h},\mathbf{k}} A_{\mathbf{h},\mathbf{k}} (\cos \phi_{\mathbf{h},\mathbf{k}} - t_{\mathbf{h},\mathbf{k}})^2}{\sum_{\mathbf{h},\mathbf{k}} A_{\mathbf{h},\mathbf{k}}},$$

where $t_{\mathbf{h},\mathbf{k}} = I_1(A_{\mathbf{h},\mathbf{k}})/I_0(A_{\mathbf{h},\mathbf{k}})$ and $\phi_{\mathbf{h},\mathbf{k}} = \phi_{\mathbf{h}} + \phi_{\mathbf{k}} + \phi_{-\mathbf{h}-\mathbf{k}}$.

In the first application (Dorset *et al.*, 1979) of multisolution phasing to electron-diffraction data (using the program *QTAN*), *n*-beam dynamical structure factors generated for cytosine and disodium 4-oxypyrimidine-2-sulfinate were used to assess the effect of increasing crystal thickness and electron accelerating voltage on the success of the structure determination. At 100 kV samples at least 80 Å thick were usable for data collection and at 1000 kV this sample thickness limit could be pushed to 300 Å – or, perhaps, 610 Å if a partial structure were accepted for later Fourier refinement. NQEST identified the best phase solutions. Later *QTAN* was used to evaluate the effect of elastic crystal bend on the structure analysis of cytosine (Moss & Dorset, 1982).

In actual experimental applications, two forms of thiourea were investigated with *QTAN* (Dorset, 1992*d*), using published three-dimensional electron-diffraction intensities (Dvoryankin & Vainshtein, 1960, 1962). Analysis of the centrosymmetric paraelectric structure yielded results equivalent to those found earlier by symbolic addition (Dorset, 1991*b*). Analysis of the noncentrosymmetric ferroelectric form was also quite successful (Dorset, 1992*d*). In both cases, the correct structure was found at the lowest value of NQEST. Re-analysis of the diketopiperazine structure with *QTAN* also found the correct structure (Dorset & McCourt, 1994*a*) within the four lowest values of NQEST, but not the one at the lowest value. The effectiveness of this figure of merit became more questionable when *QTAN* was used to solve the noncentrosymmetric crystal structure of a polymer (Dorset, McCourt, Kopp *et al.*, 1994). The solution could not be found readily when NQEST was used but was easily identified when the minimal function $R(\phi)$ was employed instead.

MULTAN has been used to phase simulated data from copper perchlorophthalocyanine (Fan *et al.*, 1985). Within the 2 Å

resolution of the electron-microscope image, if one seeks phases for diffraction data in reciprocal-space regions where the objective lens phase contrast transfer function $|C(s)| \leq 0.2$, the method proves to be successful. The method is also quite effective for phase extension from 2 Å to 1 Å diffraction resolution, where the low-angle data serve as a large initial phase set for the tangent formula. However, no useful results were found from an *ab initio* phase determination carried out solely with the electron-diffraction structure-factor magnitudes. Similar results were obtained when *RANTAN* was used to phase experimental data from this compound (Fan *et al.*, 1991), *i.e.* the multisolution approach worked well for phase extension but not for *ab initio* phase determination. Additional tests were subsequently carried out with *QTAN* on an experimental *hk*0 electron-diffraction data set collected at 1200 kV (Dorset, McCourt, Fryer *et al.*, 1994). Again, *ab initio* phase determination is not possible by this technique. However, if a basis set was constructed from the Fourier transform of a 2.4 Å image, a correct solution could be found, but not at the lowest value of NQEST. This figure of merit was useful, however, when the basis set was taken from the symbolic addition determination mentioned in the previous section.

2.5.8.5. Density modification

Another method of phase determination, which is best suited to refining or extending a partial phase set, is the Hoppe–Gassmann density modification procedure (Hoppe & Gassmann, 1968; Gassmann & Zechmeister, 1972; Gassmann, 1976). The procedure is very simple but also very computer-intensive. Starting with a small set of (phased) $F_{\mathbf{h}}$, an initial potential map $\varphi(\mathbf{r})$ is calculated by Fourier transformation. This map is then modified by some real-space function, which restricts peak sizes to a maximum value and removes all negative density regions. The modified map $\varphi'(\mathbf{r})$ is then Fourier-transformed to produce a set of phased structure factors. Phase values are accepted *via* another modification function in reciprocal space, *e.g.* $E_{\rm calc}/E_{\rm obs} \geq p$, where p is a threshold quantity. The new set is then transformed to obtain a new $\varphi(\mathbf{r})$ and the phase refinement continues iteratively until the phase solution converges (judged by lower crystallographic R values).

The application of density modification procedures to electron-crystallographic problems was assessed by Ishizuka *et al.* (1982), who used simulated data from copper perchlorophthalocyanine within the resolution of the electron-microscope image. The method was useful for finding phase values in reciprocal-space regions where the transfer function $|C(s)| \leq 0.2$. As a technique for phase extension, density modification was acceptable for test cases where the resolution was extended from 1.67 to 1.0 Å, or 2.01 to 1.21 Å, but it was not very satisfactory for a resolution enhancement from 2.5 to 1.67 Å. There appear to have been no tests of this method yet with experimental data. However, the philosophy of this technique will be met again below in the description of the maximum entropy and likelihood procedure.

2.5.8.6. Convolution techniques

One of the first relationships ever derived for phase determination is the Sayre (1952) equation:

$$F_{\mathbf{h}} = \frac{\theta}{V} \sum_{\mathbf{k}} F_{\mathbf{k}} F_{\mathbf{h}-\mathbf{k}},$$

which is a simple convolution of phased structure factors multiplied by a function of the atomic scattering factors. For structures with nonoverlapping atoms, consisting of one atomic species, it is an exact expression. Although the convolution term resembles part of the tangent formula above, no statistical averaging is implied (Sayre, 1980). In X-ray crystallography this relationship has not been used very often, despite its accuracy. Part of the reason for this is that it requires relatively high resolution data for it to be useful. It can also fail for structures comprised of different atomic species.

Since, relative to X-ray scattering factors, electron scattering factors span a narrower range of magnitudes at $\sin\theta/\lambda = 0$, it might be thought that the Sayre equation would be particularly useful in electron crystallography. In fact, Liu *et al.* (1988) were able to extend phases for simulated data from copper perchlorophthalocyanine starting at the image resolution of 2 Å and reaching the 1 Å limit of an electron-diffraction data set. This analysis has been improved with a 2.4 Å basis set obtained from the Fourier transform of an electron micrograph of this material at 500 kV and extended to the 1.0 Å limit of a 1200 kV electron-diffraction pattern (Dorset *et al.*, 1995). Using the partial phase sets for zonal diffraction data from several polymers by symbolic addition (see above), the Sayre equation has been useful for extending into the whole *hk*0 set, often with great accuracy. The size of the basis set is critical but the connectivity to access all reflections is more so. Fan and co-workers have had considerable success with the analysis of incommensurately modulated structures. The average structure (basis set) is found by high-resolution electron microscopy and the 'superlattice' reflections, corresponding to the incommensurate modulation, are assigned phases in hyperspace by the Sayre convolution. Examples include a high T_c superconductor (Mo *et al.*, 1992) and the mineral ankangite (Xiang *et al.*, 1990). Phases of regular inorganic crystals have also been extended from the electron micrograph to the electron-diffraction resolution by this technique (Hu *et al.*, 1992).

In an investigation of how direct methods might be used for phase extension in protein electron crystallography, low-resolution phases from two proteins, bacteriorhodopsin (Henderson *et al.*, 1986) and halorhodopsin (Havelka *et al.*, 1993) were extended to higher resolution with the Sayre equation (Dorset *et al.*, 1995). For the noncentrosymmetric bacteriorhodopsin *hk*0 projection a 10 Å basis set was used, whereas a 15 Å set was accepted for the centrosymmetric halorhodopsin projection. In both cases, extensions to 6 Å resolution were reasonably successful. For bacteriorhodopsin, for which data were available to 3.5 Å, problems with the extension were encountered near 5 Å, corresponding to a minimum in a plot of average intensity *versus* resolution. Suggestions were made on how a multisolution procedure might be successful beyond this point.

2.5.8.7. Maximum entropy and likelihood

Maximum entropy has been applied to electron crystallography in several ways. In the sense that images are optimized, the entropy term

$$S = -\sum_i P_i \ln P_i,$$

where $P_i = p_i / \sum_i p_i$ and p_i is a pixel density, has been evaluated for various test electron-microscope images. For crystals, the true projected potential distribution function is thought to have the maximum value of S. If the phase contrast transfer function used to obtain a micrograph is unknown, test images (*i.e.* trial potential maps) can be calculated for different values of $\Delta f_{\rm trial}$. The value that corresponds to the maximum entropy would be near the true defocus. In this way, the actual objective lens transfer function can be found for a single image (Li, 1991) in addition to the other techniques suggested by this group.

Another use of the maximum-entropy concept is to guide the progress of a direct phase determination (Bricogne & Gilmore, 1990; Gilmore *et al.*, 1990). Suppose that there is a small set H of known phases $\phi_{\mathbf{h}\in H}$ (corresponding either to origin

definition, or the Fourier transform of an electron micrograph, or both) with associated unitary structure-factor amplitudes $|U_{\mathbf{h}\in H}|$. [The unitary structure factor is defined as $|U_{\mathbf{h}}| = |E_{\mathbf{h}}|/(N)^{1/2}$.] As usual, the task is to expand into the unknown phase set K to solve the crystal structure. From Bayes' theorem, the procedure is based on an operation where $p(\text{map}|\text{data}) \propto p(\text{map})p(\text{data}|\text{map})$. This means that the probability of successfully deriving a potential map, given diffraction data, is estimated. This so-called posterior probability is approximately proportional to the product of the probability of generating the map (known as the prior) and the probability of generating the data, given the map (known as the likelihood). The latter probability consults the observed data and can be used as a figure of merit.

Beginning with the basis set H, a trial map is generated from the limited number of phased structure factors. As discussed above, the map can be immediately improved by removing all negative density. The map can be improved further if its entropy is maximized using the equation given above for S. This produces the so-called maximum-entropy prior $q^{\text{ME}}(X)$.

So far, it has been assumed that all $|U_{\mathbf{h}\in K}| = 0$. If large reflections from the K set are now added and their phase values are permuted, then a number of new maps can be generated and their entropies can be maximized as before. This creates a phasing 'tree' with many possible solutions; individual branch points can have further reflections added *via* permutations to produce further sub-branches, and so on. Obviously, some figure of merit is needed to 'prune' the tree, *i.e.* to find likely paths to a solution.

The desired figure of merit is the likelihood $L(H)$. First a quantity

$$\Lambda_{\mathbf{h}} = 2NR \exp[-N(r^2 + R^2)] I_o(2NrR),$$

where $r = |^{\text{ME}}U_{\mathbf{h}}|$ (the calculated unitary structure factors) and $R = |^{o}U_{\mathbf{h}}|$ (the observed unitary structure factors), is defined. From this one can calculate

$$L(H) = \sum_{\mathbf{h}\notin H} \ln \Lambda_{\mathbf{h}}.$$

The null hypothesis $L(H_o)$ can also be calculated from the above when $r = 0$, so that the likelihood gain

$$LLg = L(H) - L(H_o)$$

ranks the nodes of the phasing tree in order of the best solutions.

Applications have been made to experimental electron-crystallographic data. A small-molecule structure starting with phases from an electron micrograph and extending to electron-diffraction resolution has been reported (Dong *et al.*, 1992). Other experimental electron-diffraction data sets used in other direct phasing approaches (see above) also have been assigned phases by this technique (Gilmore, Shankland & Bricogne, 1993). These include intensities from diketopiperazine and basic copper chloride. An application of this procedure in protein structure analysis has been published by Gilmore *et al.* (1992) and Gilmore, Shankland & Fryer (1993). Starting with 15 Å phases, it was possible to extend phases for bacteriorhodopsin to the limits of the electron-diffraction pattern, apparently with greater accuracy than possible with the Sayre equation (see above).

2.5.8.8. Influence of multiple scattering on direct electron-crystallographic structure analysis

The aim of electron-crystallographic data collection is to minimize the effect of dynamical scattering, so that the unit-cell potential distribution or its Fourier transform is represented significantly in the recorded signal. It would be a mistake, however, to presume that these data ever conform strictly to the kinematical approximation, for there is always some deviation from this ideal scattering condition that can affect the structure analysis. Despite this fact, some direct phasing procedures have been particularly 'robust', even when multiple scattering perturbations to the data are quite obvious (*e.g.* as evidenced by large crystallographic residuals).

The most effective direct phasing procedures seem to be those based on the Σ_2 triple invariants. These phase relationships will not only include the symbolic addition procedure, as it is normally carried out, but also the tangent formula and the Sayre equation (since it is well known that this convolution can be used to derive the functional form of the three-phase invariant). The strict ordering of $|E_{\mathbf{h}}|$ magnitudes is, therefore, not critically important so long as there are no major changes from large to small values (or *vice versa*). This was demonstrated in direct phase determinations of simulated n-beam dynamical diffraction data from a sulfur-containing polymer (Dorset & McCourt, 1992). Nevertheless, there is a point where measured data cannot be used. For example, intensities from *ca* 100 Å-thick epitaxically oriented copper perchlorophthalocyanine crystals become less and less representative of the unit-cell transform at lower electron-beam energies (Tivol *et al.*, 1993) and, accordingly, the success of the phase determination is compromised (Dorset, McCourt, Fryer *et al.*, 1994). The similarity between the Sayre convolution and the interactions of structure-factor terms in, *e.g.*, the multislice formulation of n-beam dynamical scattering was noted by Moodie (1965). It is interesting to note that dynamical scattering interactions observed by direct excitation of Σ_2 and Σ_1 triples in convergent-beam diffraction experiments can actually be exploited to determine crystallographic phases to very high precision (Spence & Zuo, 1992, pp. 56–63).

While the evaluation of positive quartet invariant sums (see Chapter 2.2) seems to be almost as favourable in the electron diffraction case as is the evaluation of Σ_2 triples, negative quartet invariants seem to be particularly sensitive to dynamical diffraction. If dynamical scattering can be modelled crudely by a convolutional smearing of the diffraction intensities, then the lowest structure-factor amplitudes, and hence the estimates of lowest $|E_{\mathbf{h}}|$ values, will be the ones most compromised. Since the negative quartet relationships require an accurate prediction of small 'cross-term' $|E_{\mathbf{h}}|$ values, multiple scattering can, therefore, limit the efficacy of this invariant for phase determination. In initial work, negative quartets have been mostly employed in the NQEST figure of merit, and analyses (Dorset, McCourt, Fryer *et al.*, 1994; Dorset & McCourt, 1994*a*) have shown how the degradation of weak kinematical $|E_{\mathbf{h}}|$ terms effectively reduced its effectiveness for locating correct structure solutions *via* the tangent formula, even though the tangent formula itself (based on triple phase estimates) was quite effective for phase determination. Substitution of the minimal function $R(\phi)$ for NQEST seems to have overcome this difficulty. [It should be pointed out, though, that only the Σ_2-triple contribution to $R(\phi)$ is considered.]

Structure refinement is another area where the effects of dynamical scattering are also problematic. For example, in the analysis of the paraelectric thiourea structure (Dorset, 1991*b*) from published texture diffraction data (Dvoryankin & Vainshtein, 1960), it was virtually impossible to find a chemically reasonable structure geometry by Fourier refinement, even though the direct phase determination itself was quite successful. The best structure was found only when higher-angle intensities (*i.e.* those least affected by dynamical scattering) were used to generate the potential map. Later analyses on heavy-atom-containing organics (Dorset *et al.*, 1992) found that the lowest kinematical R-factor value did not correspond to the chemically correct structure geometry. This observation was also made in the

least-squares refinement of diketopiperazine (Dorset & McCourt, 1994*a*). It is obvious that, if a global minimum is sought for the crystallographic residual, then dynamical structure factors, rather than kinematical values, should be compared to the observed values (Dorset *et al.*, 1992). Ways of integrating such calculations into the refinement process have been suggested (Sha *et al.*, 1993). Otherwise one must constrain the refinement to chemically reasonable bonding geometry in a search for a local *R*-factor minimum.

Corrections for such deviations from the kinematical approximation are complicated by the presence of other possible data perturbations, especially if microareas are being sampled, *e.g.* in typical selected-area diffraction experiments. Significant complications can arise from the diffraction incoherence observed from elastically deformed crystals (Cowley, 1961) as well as secondary scattering (Cowley *et al.*, 1951). These complications were also considered for the larger (*e.g.* millimetre diameter) areas sampled in an electron-diffraction camera when recording texture diffraction patterns (Turner & Cowley, 1969), but, because of the crystallite distributions, it is sometimes found that the two-beam dynamical approximation is useful (accounting for a number of successful structure analyses carried out in Moscow).

MT is grateful to Professor J. C. H. Spence for critical reading of the manuscript of Section 2.5.3, for correcting the English and for helpful comments, and to Professor M. Terauchi and Dr K. Tsuda of Tohoku University and Dr K. Saitoh of Nagoya University for useful discussions.

References

Adiga, P. S., Malladi, R., Baxter, W. & Glaeser, R. M. (2004). *A binary segmentation approach for boxing ribosome particles in cryo EM micrographs*. *J. Struct. Biol.* **145**, 142–151.

Agrawal, R. K., Penczek, P., Grassucci, R. A., Li, Y., Leith, A., Nierhaus, K. H. & Frank, J. (1996). *Direct visualization of A-, P-, and E-site transfer RNAs in the Escherichia coli ribosome*. *Science*, **271**, 1000–1002.

Avilov, A. S. (1979). *Electrical measurement of reflection intensities on electron diffraction from mosaic single crystals. Sov. Phys. Crystallogr.* **24**, 103–104.

Avilov, A. S., Kuligin, A. K., Pietsch, U., Spence, J. C. H., Tsirelson, V. G. & Zuo, J. M. (1999). *Scanning system for high-energy electron diffractometry. J. Appl. Cryst.* **32**, 1033–1038.

Avilov, A. S., Parmon, V. S., Semiletov, S. A. & Sirota, M. I. (1984). *Calculation of reflected intensities in multiple-beam diffraction of fast electrons by polycrystalline specimens. Sov. Phys. Crystallogr.* **29**, 5–7.

Baldwin, P. R. & Penczek, P. A. (2007). *The transform class in SPARX and EMAN2*. *J. Struct. Biol.* **157**, 250–261.

Basokur, A. T. (1998). *Digital filter design using the hyperbolic tangent functions*. *J. Balkan Geophys. Soc.* **1**, 14–18.

Bendersky, L. A. (1985). *Quasicrystal with one-dimensional translational symmetry and a tenfold rotation axis*. *Phys. Rev. Lett.* **55**, 1461–1467.

Bendersky, L. A. (1986). *Decagonal phase. J. Phys. Colloq.* **47**, C3, 457–464.

Bendersky, L. A. & Kaufman, M. J. (1986). *Determination of the point group of the icosahedral phase in an Al–Mn–Si alloy using convergent-beam electron diffraction*. *Philos. Mag.* **B53**, L75–L80.

Bethe, H. A. (1928). *Theorie der Beugung von Elektronen an Kristallen. Ann. Phys. (Leipzig)*, **87**, 55–129.

Biemond, J., Lagendijk, R. L. & Mersereau, R. M. (1990). *Iterative methods for image deblurring*. *Proc. IEEE*, **78**, 856–883.

Blackman, M. (1939). *On the intensities of electron diffraction rings. Proc. R. Soc. London Ser. A*, **173**, 68–82.

Boisset, N., Penczek, P., Pochon, F., Frank, J. & Lamy, J. (1993). *Three-dimensional architecture of human alpha 2-macroglobulin transformed with methylamine*. *J. Mol. Biol.* **232**, 522–529.

Boisset, N., Penczek, P., Taveau, J. C., Lamy, J. & Frank, J. (1995). *Three-dimensional reconstruction of Androctonus australis hemocyanin labeled with a monoclonal Fab fragment. J. Struct. Biol.* **115**, 16–29.

Böttcher, B., Wynne, S. A. & Crowther, R. A. (1997). *Determination of the fold of the core protein of hepatitis B virus by electron cryomicroscopy*. *Nature (London)*, **386**, 88–91.

Bracewell, R. N. (1956). *Strip integration in radio astronomy. Austr. J. Phys*. **9**, 198–217.

Bracewell, R. N. & Riddle, A. C. (1967) *Inversion of fan-beam scans in radio astronomy. Astrophys. J.* **150**, 427–434.

Bricogne, G. & Gilmore, C. J. (1990). *A multisolution method of phase determination by combined maximization of entropy and likelihood. I. Theory, algorithms and strategy. Acta Cryst*. A**46**, 284–297.

Brisse, F. (1989). *Electron diffraction of synthetic polymers: the model compound approach to polymer structure. J. Electron Microsc. Tech*. **11**, 272–279.

Buxton, B., Eades, J. A., Steeds, J. W. & Rackham, G. M. (1976). *The symmetry of electron diffraction zone axis patterns. Philos. Trans. R. Soc. London Ser. A*, **181**, 171–193.

Carazo, J. M. (1992). *The fidelity of 3D reconstruction from incomplete data and the use of restoration methods*. In *Electron Tomography*, edited by J. Frank, pp. 117–166. New York: Plenum.

Carazo, J. M. & Carrascosa, J. L. (1986). *Information recovery in missing angular data cases: an approach by the convex projections method in three dimensions*. *J. Microsc.* **145**, 23–43.

Carazo, J. M. & Carrascosa, J. L. (1987). *Restoration of direct Fourier three-dimensional reconstructions of crystalline specimens by the method of convex projections*. *J. Microsc.* **145**, 159–177.

Cherns, D., Kiely, C. J. & Preston, A. R. (1988). *Electron diffraction studies of strain in epitaxial bicrystals and multilayers*. *Ultramicroscopy*, **24**, 355–370.

Cherns, D. & Preston, A. R. (1986). *Convergent beam diffraction studies of crystal defects*. *Proc. XI Int. Congr. Electron Microsc.*, Kyoto, Japan, p. 721.

Clark, J. J., Palmer, M. R. & Lawrence, P. D. (1985). *A transformation method for the reconstruction of functions from nonuniformly spaced samples*. *IEEE Trans. Acoust. Speech Signal Process*. **33**, 1151–1165.

Cochran, W., Crick, F. H. C. & Vand, V. (1952). *The structure of synthetic polypeptides. 1. The transform of atoms on a helix. Acta Cryst*. **5**, 581–586.

Conway, J. F., Cheng, N., Zlotnick, A., Wingfield, P. T., Stahl, S. J. & Steven, A. C. (1997). *Visualization of a 4-helix bundle in the hepatitis B virus capsid by cryo-electron microscopy*. *Nature (London)*, **386**, 91–94.

Cowley, J. M. (1953). *Structure analysis of single crystals by electron diffraction. II. Disordered boric acid structure. Acta Cryst*. **6**, 522–529.

Cowley, J. M. (1956). *A modified Patterson function. Acta Cryst*. **9**, 397–398.

Cowley, J. M. (1961). *Diffraction intensities from bent crystals. Acta Cryst*. **14**, 920–927.

Cowley, J. M. (1981). *Diffraction Physics*, 2nd ed. Amsterdam: North-Holland.

Cowley, J. M. (1995). *Diffraction Physics*, 3rd ed. Amsterdam: North-Holland.

Cowley, J. M. & Au, A. Y. (1978). *Image signals and detector configurations for STEM*. In *Scanning Electron Microscopy*, Vol. 1, pp. 53–60. AFM O'Hare: SEM Inc.

Cowley, J. M. & Moodie, A. F. (1957). *The scattering of electrons by atoms and crystals. I. A new theoretical approach. Acta Cryst*. **10**, 609–619.

Cowley, J. M. & Moodie, A. F. (1959). *The scattering of electrons by atoms and crystals. III. Single-crystal diffraction patterns. Acta Cryst*. **12**, 360–367.

Cowley, J. M. & Moodie, A. F. (1960). *Fourier images. IV. The phase grating. Proc. Phys. Soc. London*, **76**, 378–384.

Cowley, J. M., Moodie, A. F., Miyake, S., Takagi, S. & Fujimoto, F. (1961). *The extinction rules for reflections in symmetrical electron diffraction spot patterns. Acta Cryst*. **14**, 87–88.

Cowley, J. M., Rees, A. L. G. & Spink, J. A. (1951). *Secondary elastic scattering in electron diffraction. Proc. Phys. Soc. London Sect. A*, **64**, 609–619.

Cramér, H. (1954). *Mathematical Methods of Statistics*. Princeton University Press.

Crowther, R. A. & Amos, L. A. (1971). *Harmonic analysis of electron microscope images with rotational symmetry. J. Mol. Biol*. **60**, 123–130.

Crowther, R. A. & Amos, L. A. (1972). *Three-dimensional image reconstructions of some small spherical viruses*. *Cold Spring Harbor Symp. Quant. Biol.* **36**, 489–494.

Crowther, R. A., Amos, L. A., Finch, J. T., DeRosier, D. J. & Klug, A. (1970). *Three dimensional reconstruction of spherical viruses by*