Three-Dimensional Electron Microscopy of Individual Biological Objects Part I. Methods

W. Hoppe, H. J. Schramm, M. Sturm, N. Hunsmann, and J. Gaßmann

Max-Planck-Institut für Biochemie, Abteilung für Strukturforschung I, Martinsried bei München, West Germany

(Z. Naturforsch. **31 a**, 645–655 [1976]; received March 30, 1976)

In this paper methods and results of three-dimensional electron microscopy of individual molecules will be presented. Part I describes the general experimental and theoretical methods (microgoniometer, measuring scheme, two-dimensional and three-dimensional reconstruction, determination of the common origin of the projections). Special attention will be given to the image point shapes under different reconstruction conditions.

Three Different Types of Three-dimensional Reconstruction

The idea of three-dimensional reconstruction has been introduced into electron microscopy independently [1] in two papers [2, 3]. In detail both papers were quite different. Applications of a third approach will be presented in this paper.

a) In one paper [2] the idea of 3-dimensional protein crystallography was combined with the concept of phase contrast imaging of protein lattice projections (instead of imaging a heavy atom glass filling the holes in the protein structures). As discussed in detail in another paper [4] the concept of periodicity leads to a virtual elimination of the radiation damage of the extremely radiation sensitive protein molecules and thus to the possibility of imaging native protein molecules at high resolution. Of prime importance is the preparation of the crystals in such a way that the spatial order will be retained, since it is well known from X-ray crystallography that air-dried protein crystals are badly distorted. The concept of stabilizing the protein lattice by replacing water by a vacuum-resistant hydrophylic medium has been introduced and experimentally tested [2]. In recent years other stabilizing agents including water [5] and ice [6] have been used (see also [7, 8]). Recently the first three-dimensional structure determination of a "membrane crystal" has been achieved using glucose as stabilizing medium [9], thus confirming the feasibility of this approach.

b) The other paper [3] had the different aim of imaging stained, especially negatively stained biological macromolecules. It is clear that this restriction means a fundamental limitation in resolution and definition of the result. A heavy atom "glass" is a network of tiny microcrystals. It can cover the inner surfaces of proteins only coarsely. For these reasons the authors' aim was to determine the quaternary structure at "biological resolutions" of 10 – 20 Å. In paper [3] the frame for three-dimensional microscopy of individual molecules by tilting experiments has been described.

It was discovered later that the concept of retrieval of three-dimensional information from projections has already been used in other branches of science (e. g. radiology, radioastronomy) prior to the proposals [2, 3] for electron microscopy. The basic mathematical theory was worked out as early as 1917 in a paper by Radon [10]. As it will be shown later difficulties are not induced by the reconstructions itself, but by interpolations due to limited experimental information. This problem has attracted many research workers. There exists already extended theoretical literature on three-dimensional reconstruction (see e. g. [11–14]).

The tilting experiments were avoided in [3] by relying on the symmetry properties of the object. There are specimens with symmetries (example: helices, Figure 1), which allow the generation of non-measured "tilted" projections by calculation. The three-dimensional calculation is then based on one measured projection and on the calculated projections. Later work on spherical viruses (see [15]) combined a few accidental projections in different orientations of different specimens and calculated projections to construct a three-dimensional model. It is evident that this procedure is not equivalent to three-dimensional microscopy of a single specimen. Even in the case of an ideally symmetrical object the specimen consists of the object, of the supporting foil, and of the negative stain. It is, therefore, asymmetric in

Reprint requests to Prof. Dr. W. Hoppe, Abt. f. Strukturforschung I, Max-Planck-Institut für Biochemie, *D-8033 Martinsried* bei München

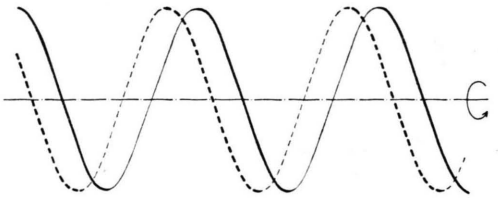

Fig. 1. Schematical diagram of an object with helical symmetry. A rotation of the object corresponds to a linear translation along the rotation axis. "Tilted" projections do not have to be measured but can be taken from one electron microscopical picture. Occupation of the helix by successive identical subunits lead to identity of projections only for a discrete number of equally spaced rotation angles.

character. Moreover, the distribution of the stain inside the periodic cavities will be slightly different. The "calculated" projections can, therefore, not accurately resemble the measurable projections in detail. Nevertheless, the principle of "projection generation by symmetry" proved to be very useful for the determination of the averaged quaternary structure of many biological macromolecules.

It is obvious that this principle cannot be used, if asymmetric objects must be studied. In this case a prediction of projections is in principle impossible.

c) The work described in the present paper deals with systematic studies in three-dimensional microscopy of individual negatively stained biological macromolecules, based on measured projections of the same specimen. The first results have been reported [16, 17].

Three-dimensional reconstruction of an individual specimen can be understood as true three-dimensional electron microscopy.

General Problems Connected with Three-dimensional Electron Microscopy

The results of this paper are based on a set of measured tilted projections. Methods for the solutions of the following problems had to be found:

a) The focusing planes for a molecule to be studied are different in the micrographs of the tilted specimens. Moreover, even the defocusing for different molecules on the same micrographs will vary. These problems could be solved by the use of two-dimensional computerized image reconstruction methods [18, 19].

b) The existing microgoniometers were too inaccurate. Special goniometers had to be developed.

c) The reconstruction needs the knowledge of a common origin in all projections. In macroscopic applications of three-dimensional reconstruction it is easy to define such an origin, as the accuracies of the mechanical movements (tilting axis etc.) are of the same order as the required resolutions. Moreover, "point markers" can be used. In electron microscopy neither the accuracy of goniometer movements ($\geqq 1\,\mu$) nor the definition of the label positions of markers (obscured by the granularity of the support) is sufficient. In [3] there is no problem in origin definition, as a common origin can be defined by symmetries in the three-dimensional body and equally in the projections. A new class of correlation methods [20] has been devised for the solution of this problem.

d) Radiation damage causes a change of the specimen during the experiment. This problem has two aspects. First it is necessary to limit the illumination dose to the exposure dose as much as possible. This can be reached by minimal exposure techniques for three-dimensional work similar to the minimal exposure conditions already used for two-dimensional work [21]. The other aspect concerns the correct averaging of the inevitable specimen changes. In two-dimensional work this will be done automatically: A two-dimensional image is a time average of the images of the succeeding specimen structures. In three-dimensional imaging a simultaneous or at least quasi-simultaneous registration of the projections should be achieved, as discussed in [22]. In goniometer work this would mean the replacement of a single tilting cycle by several cycles. As already the execution of a single cycle is experimentally difficult, this problem could not be solved until now. The best approach would be a combination of a three-dimensionally imaging microscope with electronic tilting [22] and electronic image registration. Work along these lines is in progress in our laboratory.

It is obvious that three-dimensional imaging requires a much higher standard of reproducibility than conventional microscopy. It requires the exposure of a complete set of images of the same specimen. An instability of any kind – electronical drift, drift of the specimen etc. – in a single picture requires repetition of the whole experiment (at least in the single cycle scheme). Also more sophisticated requirements have to be met in future, such as ultra high vacuum conditions around the specimen or cooling to liquid nitrogen or helium temperatures.

With respect to these criteria a) – d) (two-dimensional reconstruction, systematic determination of

the common origin of the projections, sufficient number of projections [23], minimal dose conditions) an earlier attempt for a three-dimensional ribosome reconstruction applying the ART technique [24] can be seen as an interesting precursor to true three-dimensional electron microscopy. Another interesting precursor is the "polytropic montage" by Hart [25], which can be understood as a method of backprojection.

Three-dimensional Electron Microscopy of Individual Objects and Crystals

Let us compare the above requirements with the different situation in electron microscopical protein crystallography [2, 8]: Crystals can be reproduced, therefore spoiled exposures of projections can be replaced. Correction of radiation damage by periodicity does not only compensate primary effects, but also secondary effects like contamination, reactions with residual gases in the vacuum etc. Therefore, special precautions are not necessary. Moreover, the disturbing structure of the supporting foil can be discriminated as "amorphous structure" against the "periodic structure" of the crystal.

In general, there is a real danger in reconstruction of individual objects: In contrast to crystal structure analysis there is no direct proof whether details in a result are meaningful or not; the analysis cannot simply be repeated. Another specimen might show deviations caused by a somewhat different preparation. A further test in crystallography is the chemical feasibility of the result – atomic peaks, bond lengths and bond angles. It is evident that tests of this kind could be incorporated into three-dimensional reconstructions of stable specimens only near atomic resolution. The other general way to overcome these difficulties is to develop an internally consistent procedure with individual steps, which have been tested on computer simulations etc. It is further necessary to define the physical parameters (resolution, noise, clutter) in such a way that it can be recognized with some confidence which details are physically relevant. It is a great advantage that the two-dimensional image reconstruction methods, which have been developed in the last years, allow in general the separation of the physical information from that due to accidental electron optical parameters. Let us remember that a 3-dimensional reconstruction is a "unique" piece of evidence like a single two-dimensional micrograph of the specimen. It shows the structure including artefacts induced by preparation and radiation damage. But due to an order of magnitude increase in experimental information – in going from two to three dimensions – the possibilities of structure research by electron microscopy are greatly enhanced. Moreover – and this is perhaps one of the most important points – one does not destroy part of this extra information by additional radiation damage. It has been shown that the radiation dose at the same significance level is approximately equal for two- and three-dimensional imaging [26].

Experimental and Computational Procedures

1) Two-dimensional Image Reconstruction

These methods correct for different defocusing and for the spherical abberation. They have mainly been discussed for enhancement of resolution. In our case the resolution is sufficient and only the influence of different defocusing has to be corrected. If one works at low resolutions the first zero point of the transfer function might be outside of the highest space frequency to be considered. In this case only a simple cut-off is necessary. At higher defocus, phase inversions have to be done.

2) The Microgoniometer

It has already been mentioned that individual specimens should be treated as asymmetric objects. As there is no preferred axis of symmetry the two tilting modes of the usual tilting stages are not necessary. A single tilting mode enables the construction of microgoniometers with better angular accuracies. Figure 2 shows such a single axis goniometer, which has been developed from a commercial double tilting stage (Siemens Elmiskop) with an angular definition of $\leqq 0.1^\circ$.

There is another reason which makes the orientation with respect to a preferred axis in the specimen useless: Experience has shown that substantial preillumination is necessary if one selects, focuses, orients etc. a single molecule. In our experiments we illuminate regions of the supporting foil containing many molecules. The selection of an individual molecule will be done afterwards from the photographic plates. Often the molecules have a preferred orientation on the supporting foil, but in general there is no preferred orientation in the plane of the foil. The orientation of the tilt axis normally changes there-

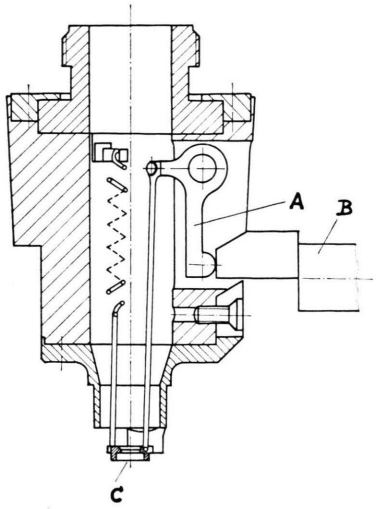

Fig. 2. Single axis goniometer for object tilting developed from a commercial double tilting stage. The angular accuracy is better than 0.1 degrees, the maximum tilting angle is approximately $\pm 50°$. (A = tilting lever, C = specimen holder, B = driving rod.)

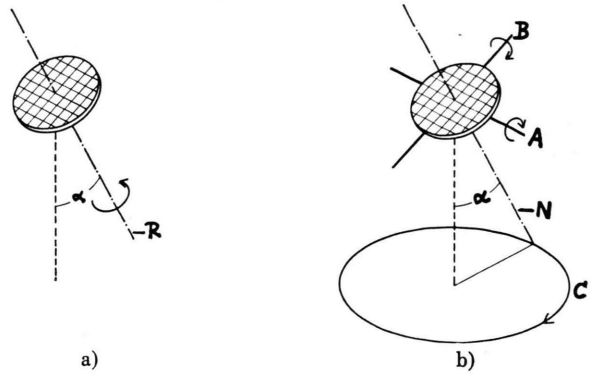

Fig. 3. Principles of cone tilting goniometers. The object lies on a plane inclined with respect to the incident electron beam by the angle α. In Fig. 3a the speciment will be rotated around R. In Fig. 3b a double tilting stage with the two axis A, B must be used with correlation of both tilting angles in such a way, that the normal N to the specimen plane moves conically along C.

fore for every molecule. The number of molecules selected in every tilt series depends on the magnification and on the size of the molecule, it varies normally between $10-100$. This procedure has further the advantage that molecules with exactly the same preparation condition and radiation loading may be compared. The focus conditions for different molecules are compensated by two-dimensional image reconstruction (see above).

It is, however, desirable that the direction of the tilt axis corresponds to one of the translation directions of the table. Then translations along the tilt axis show only slight changes of focus. Focusing and tilting on a specimen region different from the exposed region is thus possible in three-dimensional work. One can also shift simultaneously the illuminating beam and image by electronic means. An important parameter in every microgoniometer is the tilting range. For mechanical reasons we have it restricted in our case to $45° - 50°$, but a larger range would be advisable in order to decrease the blind region. Tilting around a single axis introduces asymmetries of the image point, which can be reduced, if one uses conical tilting [25, 26]. Figure 3 shows two mechanical principles for execution of cone tilting. A goniometer corresponding to the principle of Fig. 3 (rotation of the specimen in an inclined plane) would again be a one-movement instrument.

3) Common Origin Determination of Projections

One can already show by simple backprojection arguments that a common origin must be determined with an accuracy substantially higher than the resolution limit. A shift of a projection means an equivalent shift of the projection lines. Therefore, they do not cross the image point exactly. If the deviation is greater than the resolution limit, an appreciable spread of the image point will occur. The difficulties with mechanical reproducibility and accuracy of label determination in the resolution range of an electron microscope have already been mentioned.

Another obvious possibility to find a common origin would be to determine the center of gravity of the scattering potential in the projections using the well known mechanical law that the projection of a center of gravity corresponds to the center of gravity of the projection. But this determination can only be done if the specimen is finite or if at least the correct cut-outs of the projection from an infinite specimen are known in advance. If the borders of a cut-out are shifted, the center of gravity shifts in the same direction.

Based on the great success of cross correlation methods in our two-dimensional work [27–30] we developed correlation functions of slightly inclined projections for the definition of the origin [20]. It could be shown that it is possible to define a new class of correlation functions [31] with topological features characterized by the projections. The derivation in [20] has been done in a rigorous, but purely

geometrical way without use of any formulae. Corresponding to the fundamental importance of this principle for the three-dimensional reconstruction of individual microscopical objects we repeat this derivation in a more algebraic form.

Let us first recall the principle of the shift determination by cross-correlation. We assume that a two-dimensional function $\varrho(x, y)$ is seen from a shifted origin: $\varrho_2(x, y) = \varrho(x + \Delta x, y + \Delta y)$. The identical but unshifted function $\varrho_1(x, y) = \varrho(x, y)$ should be correlated with the shifted function ϱ_2. The increments Δx, Δy are assumed to be unknown and can be determined, if one calculates the convolution product (cross-correlation function) $c(u, v) = \varrho_1^* \cdot \varrho_2$ between the function ϱ_1^* (which is the enantimorph of ϱ_1) and ϱ_2. We assume further that ϱ consists of N point "atoms" with the weight g_j. Then the cross-correlation function c can be written

$$c(u, v) = \widehat{\varrho_1^* \cdot \varrho_2} = (\sum_{j=1}^{N} g_j^2)\, \delta(u - \Delta x)\, \delta(v - \Delta y) + B(u, v, \Delta x, \Delta y)\,. \quad (1)$$

The function consists of an N-fold point-maximum (cross-correlation maximum) with the weight Σg_j^2 at the position Δx, Δy and of a general background B, which need not to be discussed in the context of this paper. The shape of the peaks in the correlation function can be described by the transform of a function $|f_0|^2$, if the scattering factor of the "atoms" in ϱ is f_0. If ϱ_2 is identical with $\varrho_1(\Delta x = 0, \Delta y = 0)$, the cross-correlation function degenerates into the auto-correlation function (Patterson function in X-ray crystallography) $\widehat{\varrho_1^* \cdot \varrho_1}$ [32].

Figure 4 a represents a view [33] through a plate-like specimen parallel to the tilt axis (which will be assumed parallel to the y-axis). The "atoms" in this section are represented by dots. We now divide the specimen into K such thin sections ϱ_k parallel to x and calculate for every section the corresponding auto-correlation function C_k

$$c_k(u, v) = \widehat{\varrho_k^* \cdot \varrho_k} = (\sum_{j}^{N_k} g_{j,k}^2)\, \delta(u)\, \delta(v) + B_k(u, v)\,. \quad (2)$$

If we add these functions

$$c_s(u, v) = \sum_{k=1} c_k(u, v) = (\sum_{k}^{K} \sum_{j}^{N_k} g_{j,k}^2)\, \delta(u)\, \delta(v) + \sum_{k} B_k(u, v) \quad (3)$$

we get an integral auto-correlation peak, which is identical with the auto-correlation peak of the total

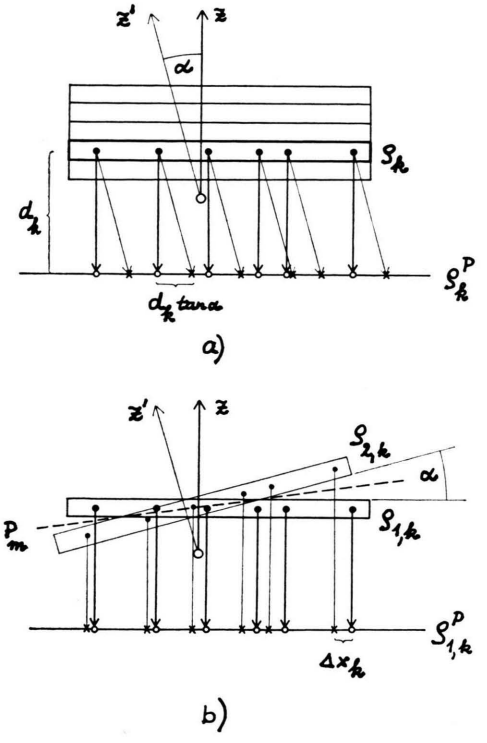

Fig. 4. Construction of autocorrelation for tilted specimen. The projected function ϱ_k^P in Fig. 4 a and 4 b is equal except for a translational shift vector (x_0, y_0).

specimen. It obviously makes no difference, whether the summation to the auto-correlation peak will be done immediately or with precalculations of the subcorrelation function C_k. This is not true for the background, as the cross terms between the sections should be taken into account. But as already mentioned we are not interested in the background in our derivation. The auto-correlation peak does not change in shape or position if we first project the function ϱ_k in Fig. 4 a onto a plane parallel to the x-y-plane (ϱ_k^P) and then calculate the auto-correlation function of this projected function ϱ_k^P.

Let us now regard the projection along the slightly inclined (inclination angle α) Z'-direction ($\varrho_{2,k}^P$) and let us calculate the cross-correlation function between $\varrho_{1,k}^P$ and the inclined projection $\varrho_{2,k}^P$. We can replace this calculation again (except for the background) by a sum over all correlation functions.

The important point is that the function $\varrho_{2,k}^P$ is identical with $\varrho_{1,k}^P$ except for translation along the x-axis (see Fig. 4 b) which depends linearly on the position of the sections inside ϱ_1:

$$\Delta x_k = x_0 + d_k \tan\alpha\,, \quad \Delta y_k = y_0\,. \quad (4)$$

The arbitrary constants x_0, y_0 correspond to the fact that in an experiment the relative shift of both projections $\varrho^{\mathrm{p}}_{1,k}$ and $\varrho^{\mathrm{p}}_{2,k}$ is unknown. The correlation peaks of c_k are shifted proportional to the shifts of the sections [see also (1)]. If the angle between Z and Z' is only small, the individual subcorrelation peaks overlap and form one broadened cross-correlation peak. The location of the maximum value for this broadened peak is obtained as the weighted average of the individual correlation peaks:

$$c_k{}^{\mathrm{p}}(u,v) = \sum_{j}^{N_k} g^2_{j,k}\, \delta(u - \Delta x_k)\, \delta(v - \Delta y_k) + B^{\mathrm{p}}(u,v)\ .$$

The weighting factor is the square of the scattering contribution of the individual "atoms". If one defines the relative shift of $\varrho^{\mathrm{p}}_{1,k}$ and $\varrho^{\mathrm{p}}_{2,k}$ by the maximum of the cross-correlation peak, one does not only define the arbitrary constants x_0, y_0 in (4) for one specific inclination angle α, but defines also a common origin in $\varrho_1{}^{\mathrm{p}}$ for all projections independently from their inclination angles. This origin can be chosen arbitrarily in one of the projections. It is fixed in $\varrho_1{}^{\mathrm{p}}$ to a line containing the center of gravity of the squared scattering contributions.

Note, that the cross-correlation peak is not spread in y-direction, as tilting does not induce shifts of the section planes ϱ_k in this direction. As already mentioned, common origin determination can be done at any inclination angle α. But the determination will obviously fail, if α is too large. The cross-correlation peak broadens and splits and cannot be discriminated from the background. This is no limitation, as successive projections can be used for the correlation. The accuracy of determination increases with the resolution of the images. On the other hand larger angular increments can be tolerated at lower resolutions.

It might be mentioned that the projections used in this derivation are not entirely identical with the projections taken in a tilting series. Corresponding to the geometry in the microscope the projection plane is always perpendicular to the projection direction. In Fig. 4 a we have to assume inclined projection planes otherwise the structures in the section planes would not be identical. The experimentally measured projections have, therefore, to be stretched by $1/\cos\alpha$ in the x-direction. This expansion would not be necessary, if the images have been taken in the geometry of the three-dimensional imaging microscope [22]. The correction can also be avoided, if the origin determination will be done on successive projections. If one chooses as reference plane P_{m} (Fig. 4 b), which bisects the two successive projection planes, no distortion correction is necessary [34]. The generalisation of these projection laws to a n-dimensional body is obvious [35]. The fact, that not the center of gravity of the structure (scattering potential) but the center of gravity of the squared structure determines the origin has the consequence that in the presence of a single heavy atom cluster the origin is near to the center of gravity of the cluster. This has been observed in test calculations on a simulated carbon foil with a cube of 8 Au atoms on top of it.

Note, that the new methods determine also the correct borders of the cut-out, if we assume that somewhat uniform regions of the support occupy the border regions. The width of the projections decreases with $\cos\alpha$. In practical work the correlations will be done twice; the second time after the determination of the correct cut-outs in order to refine the coordinates. Note, that in general the distances of tilt axis to right and left hand borders are not equal (see Figure 5). The displacement of the borders is dependent on the tilting angle and allows the determination of the distance of the center of gravity of the squared structure to the supporting film. It might be of advantage, to place the center of gravity of the squared structure in the center of the reconstruction volume. In many cases the specimen has this optimal position within the reconstruction volume.

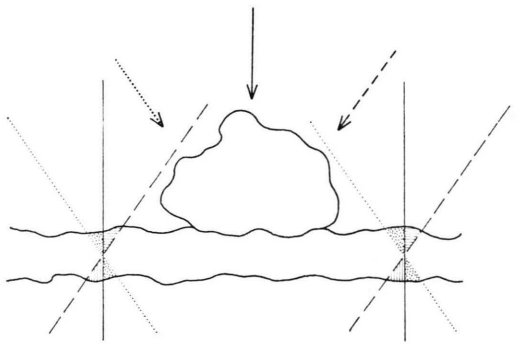

Fig. 5. Different cut-offs of the supporting foil in tilting experiments. These border effects influence the reconstruction of the object from the measured, different projection regions. The shape of the projected body differs from projection to projection.

The cross-correlation function has the same internal structure as the auto-correlation function. If

the molecule has a predominant structure (as in the case of fatty acid synthetase) the characteristic shape of the corresponding molecular structure also appears in the cross-correlation function.

4) Radiation Damage in Stained Specimens

Early work on negatively or possitively stained biological specimens was based on conventional electron micrographs, taken after a considerable pre-illumination for selection, focusing (with an intense illuminating beam for inspection on the fluorescence screen) etc. Not until recently has it been recognized that negatively stained specimens can also be considerably influenced by radiation. After a certain pre-illumination the structure of the specimen stabilizes at an equilibrium stage. Pre-illumination -free micrographs (taken under "minimal exposure conditions" [21]) show "sharper" images. It is probable, that radiation induced recrystallisation procedures take place [36], which – in combination with the detoriation of the organic matrix – lead to re-arrangement of the stain. The radiation sensitivity depends on the stain and probably also on the specimen. The dose dependence of sensitivity has never been quantitatively determined. This can be done by image difference techniques and correlation analysis in a similar way as shown on "stable" specimens at high resolution some time ago [26]. For example practical experience shows that a considerable improvement in image quality of uranyl stained tobacco mosaic virus can be achieved, if the dose is reduced to at least 0.1 C/cm^2 = 63 el/Å^2 [37]. This dose is considerably higher than the minimal dose necessary for negligible radiation damage of a protein structure which will be of the order of 10^{-3} C/cm^2 [38]. The dose 0.1 C/cm^2 already allows in principle the recognition of significant details in the 6 – 10 Å range. At 6 Å (10 Å) resolution the standard deviations of the bright field intensity I_0 stemming from electron fluctuations are $\approx 2\%$ (1.3%). It follows from the error propagation law that the contrasts (defined as $c = \Delta I/I_0$) of 20% – 10% have standard deviations $\Delta c/c \approx \sqrt{2}\,\Delta I/c$ of 15% – 30% (9% to 18%). It should be stressed that – for image relevance – bright field and dark field methods are equivalent [39]. With a frequently used dose of 1 C/cm^2 (corresponding to a density of $S = 0.4$ at $M = 200.000$, assuming an electron sensitivity of the emulsion of 0.5 el/μm^2 for $S = 1$) the figures for standard deviations are 4.70% – 9.5% (2.8% to 5.7%).

Due to the already mentioned equal significance level in two- and three-dimensional imaging a three-dimensional image shows a specimen with the same artefacts as a two-dimensional image taken with the same dose. But the three-dimensional information deductible from the image is greater by one or two orders of magnitude. From this point of view two-dimensional imaging is in fact a waste of information.

Evaluation Methods for Three-dimensional Reconstruction

If one could measure projections at arbitrarily small angular intervals in the tilting range of $\pm 90^\circ$ the reconstruction would be straight forward. Following general Fourier transform principles a two-dimensional projection corresponds to a plane passing through the origin of the three-dimensional Fourier space (reciprocal space). We assume that the Ewald sphere in this region can be approximated by a plane. This approximation is valid at medium resolutions and conventional accelerating voltages. As shown elsewhere [18], in high resolution electron microscopy these planes have to be replaced by sections of the reciprocal space along with the Ewald sphere. Image reconstruction of the complex image function is then necessary. This plane in Fourier space is perpendicular to the projection direction (Figure 6). The bundle of planes covers the reciprocal space continuously. It is easy to see that the density of Fourier coefficients near the tilting axis is higher than the density near the resolution limit. A Fourier synthesis of this three-dimensional density distribution of Fourier coefficients leads to a distorted image due to the higher density of coefficients near the tilting axis. This can be corrected by weighting in Fourier space with the distance d^* from the tilting axis.

These Fourier operations can be translated into real space. The real space equivalent of a plane through the origin of reciprocal space is a three-dimensional body. The density of this body is constant along the projection direction. Perpendicularly to this direction the density corresponds to the projection of the object. According to the linearity of Fourier transformation the three-dimensional object can be reconstructed by the super-

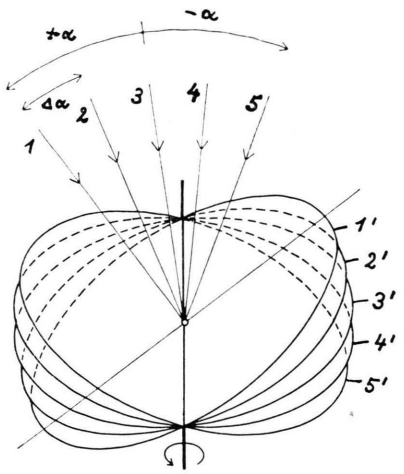

Fig. 6a. Single axis tilting with axis in (or parallel to) the object plane. The indicated projection directions (1−5) in object space (direct space) and the correlated Fourier planes in reciprocal space (1′−5′) are drawn. The radius of the Fourier planes is given by the resolution in projection. Fourier coefficients have a higher density around the tilt axis. Only a limited Fourier region can be scanned within a restricted tilting range. It determines the shape of the image point. The angular increment of successive tilts $\Delta\alpha$ determines the theoretical resolution limit for minimal clutter.

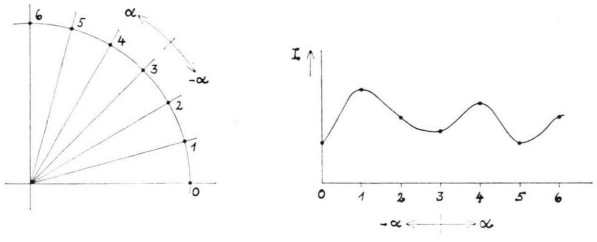

Fig. 6b. In the Cormack-Smith reconstruction scheme the measured projection values at constant radius (1−5) (Fig. 6a) are interpolated by an angular Fourier series approximation to "generate" the missing projections between the measured projections.

position of all "projection bodies". This operation is obviously "back projection". This principle has already been applied in "tomography" a long time ago. But again the result corresponds to a distorted three-dimensional image, as the superposition of the "projection bodies" corresponds to a Fourier inversion of the incorrectly weighted Fourier transform. Since a multiplication in reciprocal space corresponds to a convolution in real space, one must convolute the Fourier transform with the weighting function d^* for correcting the back-projected density.

For experimental reasons the angular increments cannot be made arbitrarily small. Therefore gaps between the Fourier planes will occur. Note, that this is only an experimental, but not a fundamental limitation: With electronic image registration and appropriate goniometers (or in a three-dimensionally imaging microscope [22]) one might better approximate the ideal situation in future. It can be shown that difficulties due to the gaps can be resolved for a finite size specimen.

There already exists a great number of propositions for the "best" reconstruction method (i. e. the best method for dealing with the restricted experimental information). The principles of these methods can be divided into three classes (see a review of algorithms by Gordon [40]):

1) Interpolation and extrapolation to derive non-measured values.
2) Weighting down of badly interpolated or extrapolated information.
3) Use of additional constraints (apart from the constraint of finite object size).

The influence of the experimental errors for every method should be considered. This is obviously especially important, if experimental information (and therefore also noise) will be enhanced. Such an enhancement will already occur in an ideal reconstruction if the d^*-weighting function is applied.

a) The Whittaker-Shannon interpolation scheme with the generalization of the Harris-type super-resolution equations [41] takes into account the finite size constraint in a very fundamental way [12, 13]. In [12] also the problems arising from the unlimited foil have been discussed (see also Figure 5). But the practical application of this method is difficult for several reasons. Another interpolation method has been devised by Cormack [11], which will be discussed later in detail.

b) The interpolation schemes gradually become extrapolation schemes for increased gaps and for "blind" regions in the reciprocal space due to restricted tilting. Noise amplification will occur under these circumstances. Restriction of the reciprocal space (reduction of the resolution) is a possible solution [12, 13]. Another way is to weight (down) the badly defined coefficients. It is already known from two-dimensional image reconstruction (e.g. zone correction plates, Maréchal reconstruction [44]) that zero regions in the Fourier

space need not distort the image point. But they produce additional clutter in the background. In fact, in contrast to the interpolation methods, only the background, not the resolution is dependent on the tilting angle increment. Typical examples of "weighting down to zero" are provided by real space methods (backprojection, ART etc.).

c) Non-linear constraints (e.g. lower and upper limit of the density) can be introduced into the real space methods of the ART-type [14]. They can also be included in reciprocal space methods, if one adopts the ideas of "phase correction" in crystallography [42, 43]. The latter concept allows also the inclusion of the discrete atom constraint [4, 45].

The Cormack Type Reconstruction Scheme

The reconstruction described in this paper have been done with a reconstruction scheme programmed by Smith et al. [46, 47]. This scheme was originally proposed by Cormack [11] and discussed and extended later by Smith et al. [46] and Zeitler [48]. It is an interpolation scheme, therefore the resolution depends on the angular increment. For clutter-free interpolation a relation similar to the relation (3) in [12] has been deduced [48] [Equation (19)]. An advantage of the scheme is that it needs relatively little computer time, because it replaces Fourier transformation by another orthogonal function system which uses immediately the coefficients of the measured projections. A disadvantage is that the image point is dependent in shape on the position within the reconstructed region [49], in a similar way as it has been shown for the related reconstruction by functional expansion [50].

Differently from the Fourier space methods the interpolation will not be done in the reciprocal space. The method "generates" the missing projections by Fourier series interpolation between the measured projections (Figure 6).

It is obvious from simple geometrical arguments that this approximation must lead to errors. As Fig. 7 shows, the simple angular interpolation generates two displaced image points with lower weights. In the reconstruction the backprojection lines cannot cross in the object point. The size of the error depends on the position of the object point within the reconstruction space. Figure 8 shows characteristic image points calculated with the so-called basis set of coefficients (Smith [46]) for

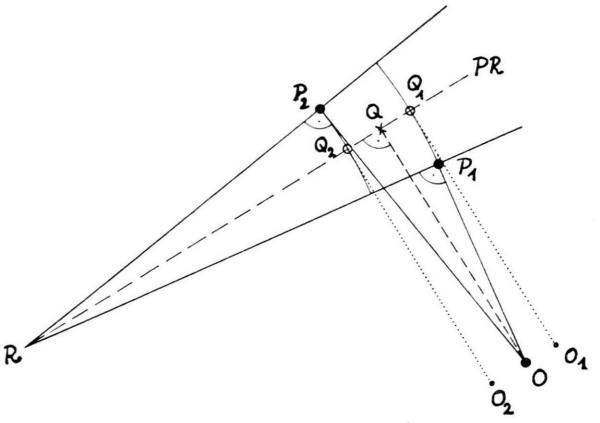

Fig. 7. Approximation of an intermediate, non-measured projection by angular approximation. The object point O occurs in two measured projections as projection points P_1 and P_2. Radial approximation leads to projection points Q_1 and Q_2 in an intermediate non-measured projection Pr. The actual projection point of object O would be Q. Backprojection of Q_1 and Q_2 leads to a displacement of O to O_1 and O_2. This splitting depends on the distance RQ.

clutter-free reconstrution. The calculations show that the resolution in a plane perpendicular to the tilt axis is virtually independent of the resolution in the projection, and only dependent on the angular increment between projections, provided that the resolution in the projection is better than this resolution. The image point resolution in the direction of the tilt axis is equal to the resolution in the projection. The resolution element is therefore anisotropic, in the case of Fig. 8 a it is for example a disc with a diameter of ~ 20 Å and a thickness

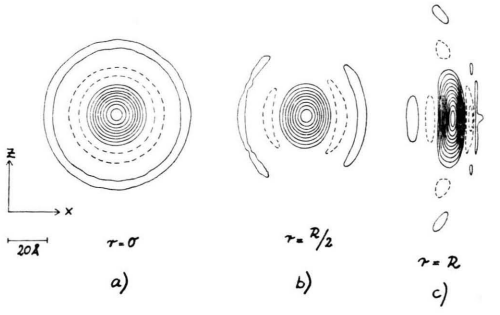

Fig. 8. Image points calculated with the basic sets of coefficients for clutter-free reconstruction. The shape and size of the image points is dependent on the position of the image point in the reconstruction area, the resolution due to tilting and the resolution in the measured projection. Generally, an anisotropic image point will be obtained in three-dimensions due to the different resolutions in projection and tilting increments $\Delta\alpha$ (see also Figure 4). Complete tilting range $\alpha = \pm 90$ degrees.

of ~ 6 Å. Figure 8 b shows the image point at $R/2$, demonstrating a slight distortion which can be tolerated. However, near R excessive distortions will occur, Figure 8 c.

Further computer simulations have been done for projection sets with blind regions. In order to adjust the basically equi-angular Cormack scheme to this experimental situation the missing projections have been introduced as zero-projections. Under these conditions already in the ideal case of arbitrarily small angular increments the image point in the plane perpendicularly to the tilt axis becomes elongated and additional strong side minima occur. Figure 9 shows image points under

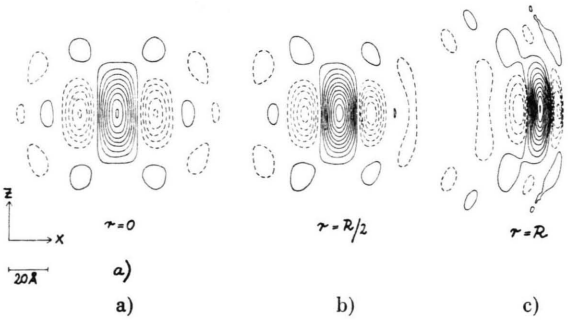

Fig. 9. Image points resulting from a restricted tilting range ($\alpha = \pm 45$ degrees) obtainable in tilting stages used for our experiments. For detailed discussion of the image points see text.

different conditions calculated from projections in the $\pm 45°$ tilting range. This tilting range corresponds to our present microgoniometer.

Let us discuss the image point in Fig. 9 a in more detail. Image points with restricted negative regions around the central maximum are well-known in optics. These restricted regions appear if the low frequency regions in the Fourier transform are weighted down (see e.g. [51]). A characteristic example is electron microscopical imaging under Scherzer focus conditions. The image points have a ring-shaped negative region. In Figure 9 b the asymmetric, restricted negative regions are concentrated along the X-axis. Corresponding to general Fourier transform principles the section parallel to the X-axis through the center of the image point corresponds to the projection of the image point spectrum onto x (see Figure 10). The low frequency weighting leads to the strong negative region in Figure 10 b. Imaging of slowly varying

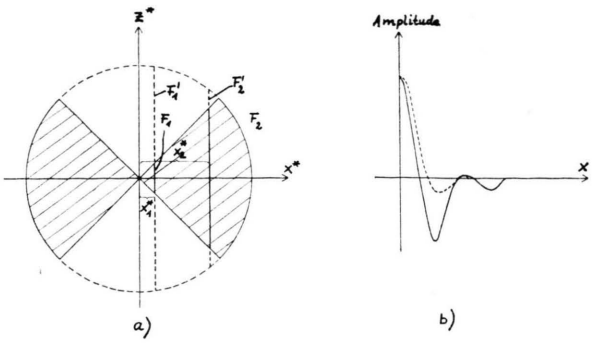

Fig. 10. Low frequency weighting due to restricted tilting range. The missing low frequency contributions lead to the negative regions in the image points (see Figure 9). Let us assume point atoms (constant weights in the Fourier plane x^*-z^*). For the calculation of the x-section through the image point the Fourier distribution has to be projected onto x^*. For a complete tilting range F_1' and F_2' are the projection values at x_1^* and x_2^*. The Fourier inversion leads to the known image point of a circular aperture (dashed curve in Figure 10 b). F_1' and F_2' has to be replaced by F_1 and F_2 for the restricted tilting angle. This weighting produces in Fourier inversion the strong negative side maximum. Note, that the rippling after the negative side maximum is negligible. Point structures (atomic structures) with minimal distances larger than twice this distance (in this direction) will not be affected by this image point shape.

object functions with image points of these types are error free. The image density is roughly proportional to the difference between the integrals of the positive and negative contributions in the restricted region. However, for sharp cut-offs and for pronounced object features near the resolution limit, image artefacts might occur (for example Fresnel seams near the cut-offs). In practice these artefacts will be partly averaged out, if several specimens with randomly oriented tilting axes are reconstructed. However, at restricted tilting conditions it would be advantageous to work with a reconstruction resolution substantially higher than the desired resolution. This enhanced resolution leads also to a reduction of image point spread in the z-direction (Figure 9 a).

The unwanted side maxima could be completely removed, if one replaces the single tilt axis scheme by conical tilting (Figure 3). Furthermore, the elongation of the image point will be diminished. An elongation ratio of about 2:1 can be realized with accessible tilting angles. These features are a strong argument for the replacement of single axis tilting by conical tilting in future work on individual three-dimensional electron microscopy.

Supported by the Deutsche Forschungsgemeinschaft (the biological work especially by the Sonderforschungsbereich 51, Medizinische Molekularbiologie und Biochemie in München) and the Fonds der Chemischen Industrie. We should like to express our sincere thanks.

[1] With respect to the independence of both papers: In 1967 during the preparation concerning the programme for the EMBO Workshop on Protein Crystal Structure Analysis (Hirschegg 1968) the organizers (M. F. Perutz and W. Hoppe) realized that work on similar lines was going on in their laboratories. Both groups gave a report in this Workshop.
[2] W. Hoppe, R. Langer, G. Knesch, and Ch. Poppe, Naturwiss. **55**, 333 [1968].
[3] D. J. DeRosier and A. Klug, Nature London **217**, 130 [1968].
[4] W. Hoppe, Ber. Bunsenges. phys. Chem. **74**, 1090 [1970].
[5] D. F. Parsons, Science **186**, 407 [1974].
[6] R. M. Glaeser, J. Ultrastruct. Res. **36**, 466 [1971].
[7] U. F. Thomanek, F. Parak, and R. L. Mössbauer, H. Formanek, P. Schwager, and W. Hoppe, Acta Cryst. **A 29**, 263 [1973].
[8] R. Langer, Ch. Poppe, H. J. Schramm, and W. Hoppe, J. Mol. Biol. **93**, 159 [1975].
[9] R. Henderson and P. N. T. Unwin, Nature London **257**, 28 [1975].
[10] J. Radon, Ber. Saechs. Akad. Wiss. Leipzig, Math. Phys. Kl. **69**, 262 [1917].
[11] A. M. Cormack, J. Appl. Phys. **35**, 2908 [1964].
[12] W. Hoppe, Optik **29**, 617 [1969].
[13] A. Klug and R. A. Crowther, Proc. Roy. Soc. London **A 317**, 319 [1970].
[14] R. Gordon and G. T. Herman, Proc. Soc. Photo-Opt. Instr. Eng. **47**, 2 [1975].
[15] R. A. Crowther and A. Klug, Ann. Rev. Biochem. **44**, 161 [1975].
[16] W. Hoppe, Proceedings of the Eighth Intern. Congr. on Electron Microscopy, Canberra, **I**, 308 [1974].
[17] W. Hoppe, J. Gaßmann, N. Hunsmann, H. J. Schramm, and M. Sturm, Hoppe-Seyler's Z. Physiol. Chem. **355**, 1483 [1974].
[18] W. Hoppe, Acta Cryst. **A 26**, 414 [1970].
[19] W. Hoppe, Proceedings on the Eighth Intern. Congr. on Electron Microscopy, Canberra, **I**, S. 240.
[20] W. Hoppe, Naturwiss. **61**, 534 [1974].
[21] R. C. Williams and H. W. Fisher, J. Mol. Biol. **52**, 121 [1970].
[22] W. Hoppe, Z. Naturforsch. **27 a**, 919 [1972].
[23] R. A. Crowther and A. Klug, J. Theor. Biol. **32**, 199 [1971].
[24] R. Bender, S. H. Bellmann, and R. Gordon, J. Theor. Biol. **29**, 483 [1970].
[25] R. Hart, Science **159**, 1464 [1968].
[26] W. Hoppe, P. Bussler, A. Feltynowski, N. Hunsmann, and A. Hirt, Image Processing and Computer-Aided Design in Electron Optics, Ed. Hawkes, Acad. Press Inc. London, Ltd. **1973**, p. 91 ff. — R. Hegerl and W. Hoppe, Optik, in preparation.
[27] W. Hoppe, R. Langer, J. Frank, and A. Feltynowski, Naturwiss. **56**, 267 [1969].
[28] J. Frank, Optik **30**, 171 [1969].
[29] R. Langer, J. Frank, A. Feltynowski, and W. Hoppe, Ber. Bunsenges. phys. Chem. **74**, 1120 [1970].
[30] J. Frank, P. Bussler, R. Langer, and W. Hoppe, Ber. Bunsenges. phys. Chem. **74**, 1105 [1970].
[31] We have not done a systematic search of the literature and we are quite prepared to learn that what we have found is an already known mathematical result.
[32] Concerning the properties of correlation functions — especially of point functions — see the text books of X-ray crystallography.
[33] This projection has been defined in order to facilitate the derivation. It does not correspond to the projections created by imaging in the microscope (projection direction parallel to z).
[34] N. Hunsmann, Thesis, T.U. München 1975.
[35] but of no practical value for our problems, except perhaps in the four-dimensional case of trace structure analysis.
[36] P. N. T. Unwin, Z. Naturforsch. **29 a**, 158 [1974].
[37] R. G. Hart and J. M. Yoshiyama, J. Ultrastr. Res. **51**, 40 [1975].
[38] M. S. Isaacson and A. V. Crewe, Ann. Rev. Biophys. Bioeng. **4**, 165 [1975].
[39] O. Scherzer, Ber. Bunsenges. phys. Chem. **74**, 1154 [1970].
[40] R. Gordon, IEEE Trans. Nucl. Science, NS **21**, 78 [1974].
[41] J. L. Harris, J. Opt. Soc. Ann. **54**, 931 [1964].
[42] W. Hoppe and J. Gaßmann, Acta Cryst. **B 24**, 97 [1968].
[43] J. Gaßmann and K. Zechmeister, Acta Cryst. **A 28**, 270 [1972].
[44] A. Maréchal and P. Croce, C. R. Acad. Sci. Paris **237**, 607 [1953].
[45] W. Hoppe, Phil. Trans. Roy. Soc. London **261**, 71 [1971].
[46] P. R. Smith, T. M. Peters and R. H. T. Bates, J. Phys. A: Math. Nucl. Gen. **6**, 361 [1973].
[47] We thank Dr. P. R. Smith for a copy of his computer programs and his assistance in adapting the programs.
[48] E. Zeitler, Optik **39**, 396 [1974].
[49] The Fourier calculation based on Whittaker-Shannon linear equations leads to image points, which are independent of the position.
[50] R. A. Crowther, L. A. Amos, and A. Klug, Proc. Fifth Europ. Congr. Electr. Micr. **1972**, 593.
[51] J. Waser and V. Shoemaker, Rev. Mod. Phys. **25**, 671 [1953].