

Principles of Trace Structure Analysis in Electron Microscopy

W. Hoppe

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

(Z. Naturforsch. **30 a**, 1188–1199 [1975]; received March 27, 1975)

Structure analysis done with X-rays or with electrons, corresponds to an averaging of the specimen configuration over the exposure time. Such an analysis leads to the structure of the specimen only if the configuration does not change during exposure or if the change is negligible (stable molecule, crystal). Some time ago for instable configurations the method of "trace structure analysis" was proposed. The present paper explains the general principles of this method in detail. The static image of the structure will be divided into a dynamic succession of images along the time axis ("three-dimensional film"). But this division is not really characteristic for the method (it is in fact trivial). The main idea is to work with non-significant three-dimensional elementary structures and to use redundancies of the radiation induced physico-chemical processes for the combination of these elementary structures. Obviously these "redundancy laws" can only be unravelled in detail by trace structural work. It will be shown further that for special classes of these laws correlation analysis and pattern recognition procedures can be applied.

Introduction

Radiation damage is a fundamental problem in the structure analysis of individual aperiodic objects using three-dimensional electron microscopy and related methods¹. Indeed there are types of specimens where the energy caused by inelastic scattering events can be dissipated in the specimen without breaking bonds. However, in the important field of organic molecules all our experience demonstrates that already an extremely small electron dose leads to a large amount of broken bonds. Our knowledge has been considerably improved in the last years through investigations by Glaeser⁶ and by the work of Isaacson, Johnson and Crewe⁷.

A number of solutions to the radiation damage problem have been proposed:

- 1) Flash electron microscopy; 2) Stability by low temperatures; 3) Stability by implanted electron acceptors (e. g. heavy atoms);
 - 4) Restoration by secondary chemical reactions;
 - 5) Repetition of the specimen in periodic structures.
- Breedlove and Trammell⁸ have shown that the extremely fast process of bond breaking would mean that flash microscopy has to be done in an impossibly short time interval. This method may therefore be useful only in the reduction of secondary reactions or in the case of relatively slow primary processes.

Investigations at low temperatures will not influence the primary effect of bond breaking. But

secondary reactions might be prevented and – as Siegel⁹ has shown – most reaction products will remain inside the specimen. Conservation and restoration¹⁰ are interesting possibilities. But only experience can show if by using these methods a real protection of the specimens can be achieved. The only straightforward method to reduce the influence of radiation damage is to use periodic objects. A crystal contains a great number of molecules with identical structure. Simple Fourier transform principles show that the structure determined by diffraction methods is an average of the structures in all unit cells. The destruction of single molecules, hit by electrons, does not influence greatly the average image. In fact, the "periodicity principle" is the only reason that in X-ray structure analysis the radiation damage plays only a minor role. Probably, electron microscopy of tiny crystals will become an important tool in the future. Crystallisation of small crystal fragments might be successful where no normal crystals can be obtained. The phase problem can be solved experimentally (see also¹¹). Imaging of anorganic, radiation insensitive crystal lattice projections has been already achieved at high resolution^{12, 13}. But one of the final aspects of the new electron microscope methods is the study of structure laws of aperiodic objects. For these problems the use of periodicity is obviously useless. One could think of averaging over many identical molecules in an amorphous specimen. But this is dangerous, as a structure depends on the environment, which is – in contrast to the environment in a crystal lattice – different for every molecule.

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



Dieses Werk wurde im Jahr 2013 vom Verlag Zeitschrift für Naturforschung in Zusammenarbeit mit der Max-Planck-Gesellschaft zur Förderung der Wissenschaften e.V. digitalisiert und unter folgender Lizenz veröffentlicht: Creative Commons Namensnennung-Keine Bearbeitung 3.0 Deutschland Lizenz.

Zum 01.01.2015 ist eine Anpassung der Lizenzbedingungen (Entfall der Creative Commons Lizenzbedingung „Keine Bearbeitung“) beabsichtigt, um eine Nachnutzung auch im Rahmen zukünftiger wissenschaftlicher Nutzungsformen zu ermöglichen.

This work has been digitalized and published in 2013 by Verlag Zeitschrift für Naturforschung in cooperation with the Max Planck Society for the Advancement of Science under a Creative Commons Attribution-NoDerivs 3.0 Germany License.

On 01.01.2015 it is planned to change the License Conditions (the removal of the Creative Commons License condition "no derivative works"). This is to allow reuse in the area of future scientific usage.

In this paper a fundamentally new method "trace structure analysis" will be proposed. Its basic ideas have already been explained in a previous paper¹⁴. We had first given the name "trace analysis" to the method. But this name exists already in electron microscopy (of crystals) with another meaning (see e. g.¹⁵). Experiments, making use of the principles in simplified form, have already been published¹⁶. This paper will demonstrate arguments and possible procedures in more detail. Three-dimensional electron microscopy can be done in two different ways:

- 1) mechanical tilting of the specimen (see^{17, 18})
- 2) electronic tilting of the primary beam¹⁹.

It will be shown in the course of this paper that electronic tilting methods have many advantages for "trace structure analysis" of organic molecules.

The Trace Structure Analysis Principle

The characteristic feature of the new method is the use of redundancies by which non-measurable information is replaced. In order to clarify this argument let us recall a situation where the use of redundancies was the only way to solve a fundamental problem. In X-ray crystal structure analysis only the amplitudes of the structure factors, but not their phases can be measured. Therefore part of the information for the calculation of the electron density function is missing. All the methods for phase determination make use of the fact that the electron density function is severely restricted. Electron density is positive everywhere, condenses at the sites of the atoms to peaks with known shape and there exist minimum distances between the peaks. This general knowledge of the structural features of the electron density functions is sufficient to select from the infinite set of possible functions the correct one. Let us note that the redundancies will only be partly used by the existing phase determining procedures. Special structure laws, such as bond distances and bond angles will only be used a posteriori in the check whether a structure is chemically feasible or not.

The phase problem does not exist in electron microscopic structure analysis. It is one of the great advantages of this method that the phases can be measured by experiment. However, — and this concerns the main idea of our new principle — the structural redundancies can be used in another way. Let us remember in this context that an experiment

in the electron microscope should be regarded as a dynamic process and not — as in crystallography — as a static process. The "reagents" are the specimen, the residual gases in the vacuum (they can be minimized by working under ultra high vacuum conditions) and the electrons, which take part in the inelastic scattering processes of the reaction. A reaction is a time dependent process and it seems obvious that it should be represented as such.

The complete information would correspond to something like a cinematographic film, which shows a series of three-dimensional structures. The first "image" shows the native specimen, the last image shows the specimen changed by the radiation reaction process during exposure. If each of these images could be measured with sufficient accuracy, there would be no further problem. The first image would show the native structure and it would not be very interesting in most cases to study the further images. Let us now regard the features of the "film" in more detail. The atoms appear in this four-dimensional representation as lines (traces), extending in the direction of the time axis. In the case of a stable specimen these traces are parallel to the time axis, but due to radiation damage they will have a more complicated shape. They can be somewhat inclined to the time axis — this means a slow movement of the atom. The most characteristic features are discontinuities in the individual traces, which generally (especially if the specimen is kept at helium temperature) are of the order of atomic distances. These "jumps" correspond to bond breaking reactions in the specimen. Let us now assume that we project the four-dimensional image along the traces. Obviously we get the structure of the molecule damaged by the radiation at the end of the exposure. If we do the projection in the opposite direction, we get the native molecule at the beginning of the experiment. If the molecule is stable, both images are identical and the projection along the traces is a simple projection along the time axis. The important point is that in these projections the significance of every atomic image is much higher than the significance in the last (or first) three-dimensional image in the "film". The reason is simply that not the electrons scattered in the time interval Δt , corresponding to a single image, but the electrons scattered during the time $t = N \cdot \Delta t$ (N = number of the images in the "film") will be used for the definition of the atomic site. This projection

is only straightforward, if the structure can be recognized in every three-dimensional image taken in an interval Δt . A refinement of the projected image is also possible if we apply a cyclic projection process, which tends to sharpen the projected atomic images.

Let us now assume that a single three-dimensional image is not significantly determined and that we cannot recognize the individual atoms. It is evident that in this case the projection described above cannot be made. Let us further assume that we know certain structural implications ("redundancy") of the radiation process. The simplest assumption is a stable specimen: In this case we know that all traces are parallel to the time axis. Obviously the projection process degenerates into a simple projection along the time axis. The projected image is now significant, in spite of the fact that the successive images in the "film" are insignificant. However, stability against radiation is only one such type of a redundancy. The question arises if more general types of redundancies (for radiation instable specimen) can be found, which allow projection along the traces.

Let us compare our situation with that in crystal structure analysis. The phases there are unknown, but only structures with atomic peaks are chemically feasible. The situation is quite similar in our case: The traces are unknown, but the projection must lead again to an atomic peak structure. In both cases the constraints are inherent to the structure. In crystal structure analysis we have constraints in the static structure. In our present case constraints in the dynamic structure must be used. In crystallography the constraints can be of a general nature – positivity in the electron density function, peaked electron density – or they can contain the whole chemical information about possible molecular structures – bond lengths, bond angles etc. Similarly in our case there might also be more general dynamical constraints or detailed constraints. This corresponds to a knowledge of the structural laws of all possible radiation chemical processes. It has already been shown in crystallography that the general constraints can solve the phase problem (direct methods). Are there similar constraints in our case?

In a stable specimen all traces are parallel to each other and parallel to the time axis. We relax these conditions in such a way that we demand identity of the traces (except for translations) for groups of

adjacent atoms and that we allow inclined and curved traces for these atomic groups with accidental jumps. These constraints mean chemically that bond breaking or radiation induced binding is a singular process, which changes the structure of the specimen only in a local region near to the radiation induced event. The internal structures of the "broken parts" will not be changed. But the different space requirements corresponding to a bond breaking (or binding) radiation process will force the broken parts to make a lateral displacement (Figure 1). We note that this principle does not mean preservation of all structure details. It is allowed that a group of atoms has a structure I at a time t_1 and a completely different structure II at a time t_2 . In the time intervals before t_1 and after t_2 the identity principle will be valid. But due to the singular radiation events the overall structure will slowly change.

It is obvious that the redundancy principle mentioned above is to some extent tentative. It is to be proven by the experience in molecular radiation chemistry, which will be gained by this type of analysis. It is evident that more detailed redundancies can be built into the scheme. On the other hand our knowledge of special radiation chemistry for the type of the specimens studied can be used for an a posteriori check of the procedure (similar to the a posteriori check of bond lengths and bond angles in a crystallographic structure). The projection process is not restricted to single images in the "film". Due to its self consistency it can produce any intermediate image. Thus refinement procedures

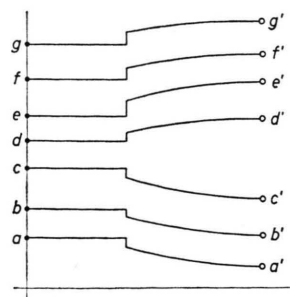


Fig. 1. Trace structure analysis schematically shown with a one-dimensional example. The molecule will be split into two halves by breaking a bond. Before the radiation event the traces are parallel to the time axis. After the event both parts tend to shift in opposite directions in order to fulfil the different space requirement for an intermolecular bond. The changes of the shifts (velocities) depend on external parameters (temperature, size of the molecule, environment etc.).

are feasible by recycling the projection operations using knowledge from radiation chemistry etc.²⁰.

Correlation Analysis as a Mathematical Method for Pattern Analysis

Redundancy principles of the type described above mean mathematically correlations between the subsequent three-dimensional images of the "film". Therefore correlation analysis is a method adapted to the problem.

Let us denote by ϱ_1 a partial structure with N atoms existing (and measured) at the time t_1 . At the time $t_1 + \Delta t$ the structure ϱ_1 will be changed into a structure ϱ_1' . It should be noted that the time parameter is strictly speaking not appropriate. Radiation damage depends on the intensity of the primary beam and on time. If there were complete reciprocity, the dose $I \cdot t$ could be used as a better parameter. But in fact reciprocity can not be expected. Re-arrangement of the complete structure after a radiation event will not depend on the intensity of the primary beam but only on the forces and on the temperature. Therefore a two-parameter representation — I and t — would in fact be adequate. The intensity of the illuminating beam might become an important experimental parameter in trace structure analysis.

We neglect the scattering on H and we regard the scattering power of the light atoms C, N, O as approximately equal. We assume further, that every light atom has scattered n' electrons during the time interval Δt . A number $n = \eta n'$ of these electrons should be collected in one point at the centre of the atom image by the three-dimensional microscope. η is a "utilization factor"; it should be as near as possible to unity. A small utilization factor η leads to an enhancement of radiation damage. If there are heavier atoms in the specimen, the number of image forming electrons will be increased by a factor corresponding to the relation of the scattering powers of the heavy atom to the light atom. The description of the result as a point structure with infinite resolution is certainly a crude approximation. The next approximation would be the collection of the electrons in a "resolution box" defined by the wave optical calculations. Note, that the number of collected electrons n is only the average with a standard deviation \sqrt{n} [see (4)]. Only in dark field image formation can the result be expressed as the number of electrons n in the image point. In the

equivalent bright field case the number of electrons in the image point is much greater than n , but the signal to noise ratio (in dark field case n/\sqrt{n}) is approximately the same²².

We firstly assume, that the change in the structure ϱ_1 to ϱ_1' is a movement by a translation vector $\Delta \mathbf{r}$. This seems to be a very special assumption. It will be discussed later that the specimen can be divided into domains, which in their internal structure are unaffected by radiation damage during the exposure Δt . The atoms 1–5 and 6–9 in Fig. 1 constitute domains. The partial structures ϱ_1 and ϱ_1' can be understood as domain structures. We further assume that the time interval Δt , which has been used for the measurement of ϱ_1 , is equal to the exposure time Δt for the measurement of ϱ_1' . We calculate the cross correlation function (convolution of ϱ_1 and the inverted structure $\tilde{\varrho}_1'$ ²³) following a simple principle which to our knowledge was first used by Buerger²⁴ for the calculation of the Patterson function in crystallography. The structure factors of the two point structures ϱ_1 and $\tilde{\varrho}_1'$ are $(\mathbf{r}, \mathbf{r}^*$ direct space and reciprocal space vectors, u_j relative scattering power of the j -th atom; we set $u_j = 0$ for H, $u_j = 1$ for C, N, O):

$$\mathcal{F}_{\mathbf{r}^*} = \sum_j^N u_j n \exp \{2 \pi i (\mathbf{r}^*, \mathbf{r}_j)\}, \quad (1)$$

$$\tilde{\mathcal{F}}_{\mathbf{r}^*}' = \mathcal{F}_{\mathbf{r}^*}' = \sum_j^N u_j n \exp \{2 \pi i (\mathbf{r}^*, -\mathbf{r}_j - \Delta \mathbf{r})\}. \quad (2)$$

In reciprocal space the convolution is equivalent to the product:

$$\begin{aligned} \mathcal{F}_{\mathbf{r}} \cdot \mathcal{F}_{\mathbf{r}^*}' &= \sum_{j=j'}^N u_j^2 n^2 \exp \{2 \pi i (\mathbf{r}^*, -\Delta \mathbf{r})\} \\ &+ \sum_{j \neq j'}^N \sum_{j'}^N u_j u_{j'} n^2 \exp \{2 \pi i (\mathbf{r}^*, \mathbf{r}_j - \mathbf{r}_{j'} - \Delta \mathbf{r})\}. \end{aligned} \quad (3)$$

Now the structure factor of the product (3) can again be understood as a structure factor of a point structure (the point structure of the convolution of ϱ_1 and $\tilde{\varrho}_1'$). It consists of N^2 points²⁵, which are quasi-randomly²⁶ distributed over a region with the 8-fold volume of ϱ_1 . We neglect all terms in (3) except the first one, which corresponds to an N -fold peak at the site $-\Delta \mathbf{r}$. For ϱ_1 the signal to noise ratio p of every point is given by:

$$p = u_j n / \sigma_p, \quad \sigma_p = \sqrt{u_j n}. \quad (4)$$

The signal to noise ratio P of the convolution peaks in (3) can simply be calculated by the Gaussian error propagation law:

$$P = u_j u_{j'} n^2 / \sigma_P, \quad \sigma_P = n \sqrt{n [u_j u_{j'} (u_{j'} + u_j)]} \quad (5)$$

which can be reduced for pairs of identical atoms to:

$$P = u_j^2 n^2 / u_j n \sqrt{2 u_j n} = p / \sqrt{2}. \quad (6)$$

Of special interest is the signal to noise ratio P_N of the "correlation" peak [first term in (3)]:

$$P_N = n^2 \sum_j u_j^2 / \sigma_{P_N}, \quad \sigma_{P_N} = n \sqrt{\sum_j 2 u_j^3 n}. \quad (7)$$

For equal atoms (7) can be written:

$$P_N = P \cdot N / \sqrt{N} = P \frac{N}{2} / \sqrt{\frac{N}{2}}. \quad (8)$$

Equation (8) is of considerable interest. It states that the signal to noise ratio of the cross correlation peak P_N is substantially increased compared with the signal to noise ratio p of an atomic peak in ϱ_1 or ϱ_1' . In fact, it would be necessary to measure the structure ϱ_1 with $N/2$ -fold radiation dose if the same signal to noise ratio for the atomic peak should be achieved. Therefore also correlation peaks of structures with insignificant atomic peaks can be of significance.

Let us now assume, that the correlation (3) has been calculated for all successive three-dimensional images taken in time intervals Δt . The trace of the correlation peak can thus be determined. We project now the four-dimensional image along this trace (generalized trace projection) and we get a projected three-dimensional image with a significance of the atomic peaks (9), where

- N' denotes the number of successive images,
- the radiation dose is equivalent to one scattered electron per atom in the picture:

$$P_{N'} = p N' / \sqrt{N'}. \quad (9)$$

As already mentioned, this procedure can be done with completely insignificant atomic structures and it leads to the significant atomic structure.

A special example of a correlation function will be discussed in Figure 2. Figure 2a is an aromatic molecule. Figure 2b shows the correlation function for the ideal case (equal weight for all atoms). Note, that there are strong secondary maxima in the background (up to a weight 10 compared with the weight 16 of the correlation peak) which arise from the translational symmetry of condensed aromatic

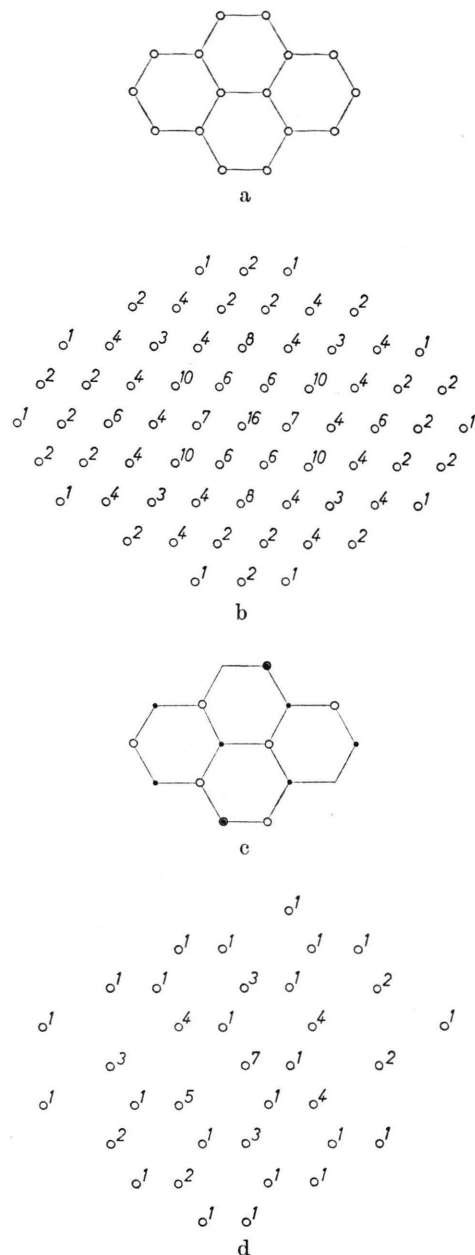


Fig. 2. Correlation functions in the special case of aromatic groups. For molecules of a quasi-random structure the cross correlation function between identical units (equal to the auto-correlation function) has a quasi-random background. In the case of an aromatic molecule (Fig. 2a) the correlation function (Fig. 2b) has strong side maxima (weight of the correlation peak 16, weight of the strongest side peaks 10). This corresponds to a new type of redundancy. The search should be done for a hexagonal lattice of peaks instead for a single peak. In Fig. 2c it has been assumed that due to a short Δt only half of the atoms have scattered (dark field illumination). First image: points, second image: small circles. The correlation function (Fig. 2d) changes the weights, but retains the hexagonal array of Figure 2b.

ring structures. This "aromatic constraint" leads to a new important redundancy. It is obviously possible to look (using an appropriate pattern recognition function) for groups of strong peaks in the hexagonal array of Fig. 2b instead of for a single correlation peak, thus enhancing the statistical significance.

Correlation functions of the type (3) are routinely used in our present method for reconstructing a two-dimensional image from a number of subsequent exposures. In this case the translations $\Delta \mathbf{r}$ are generated by the drift of the specimen.

It should be noted that in the general case not only translations but also rotations of the specimen can occur. Some time ago we have shown that in these cases the rotational and the translational parameters can be separated, and we have used these procedures in a special crystallographic phase determining method (convolution molecule method)^{27, 28}. This means that firstly the rotation should be determined by our rotation dependent correlation procedure (rotation function) which is invariant against translations. With the known rotation parameters the translation can then be determined using Equation (3). In drift correction the rotation is in general negligible.

The ideas of the trace structure analysis lead to a generalization of the image difference method, which has already been experimentally tested (for projections)¹⁶. The image difference technique reveals molecular reactions at the surface of the specimen. The site of the reaction products is determined by measuring the difference between the images after and before the reaction. If these steps can be properly separated, the results are unambiguous. But there are cases where the reaction takes place continuously during illumination. Such a reaction is for example the contamination of the specimen by the residual gases in the vacuum. It is obvious that a projection of the four-dimensional image along the trace, determined by cross correlations of the subsequent images leads to more information with respect to the gas reactions. Besides the traces of the stable atoms, which can be found from the beginning of the exposure until its end, there exist traces which appear and (or) disappear somewhere in the four-dimensional space. It might be mentioned that in these cases only trace structure analysis can determine the true peak height of the difference atoms¹⁶. Simple image difference technique

will give ambiguous results. In this type of analysis the time intervals should be taken in such a way that every image in the "film" has already some significance. Otherwise the beginning or the end of a trace cannot be determined properly.

Following these more introductory remarks we discuss now the case of an unstable specimen. We compare again the structures of the specimen at the arbitrary time t_1 and at the time $t_1 + \Delta t$. Due to bond breaking and bond binding by radiation damage we assume that the specimen has been broken into K domains. The structure of every domain is identical to the structure of the corresponding part of the specimen. At the borders of the domains the structure will, in general, be different. But this is of no importance for the analysis as the identical parts produce the correlation peaks. Note, that the number of domains need not be equal to the number of radiation events. In the case of chain like structures (e.g. proteins, nucleic acids) there might be proportionality between domains and radiation events.

We calculate now the cross correlation function in a similar way as in (3). The structure factor of the structure ϱ of the specimen after Δt can be written as a sum over the structures of the domains:

$$F_{r^*} = \sum_k^K \sum_j^{N_k} u_{k,j} n \exp \{ 2 \pi i (\mathbf{r}^*, \mathbf{r}_{k,j} + \Delta \mathbf{r}_k) \}. \quad (10)$$

(2) can be regarded as one of the K partial structures in (10). N_k is a variable number, as the number of peaks in the domains need not be constant. We calculate again $F \cdot F'^*$, but we neglect all terms which give single-weighted peaks [the second term in (3)]

$$F_{\varrho} \cdot F_{\varrho'}^* \approx \sum_k^K \sum_j^{N_k} u_j^2 n^2 \exp \{ 2 \pi i (\mathbf{r}^*, -\Delta \mathbf{r}_k) \}. \quad (11)$$

It can be seen that the former single cross correlation peak of a stable specimen is now replaced by a sum of cross correlation peaks, which are displaced by the vectors $-\Delta \mathbf{r}_k$ from the autocorrelation peak of the structure at time t_1 . This "spectral" distribution of the cross correlation peaks in (11) is of considerable interest, as it gives an overall measure of the influence of the radiation damage during the time period Δt ²⁹. It contains, for example, information about the splitting of the specimen into domains. From the relative height of one of the cross correlation peaks (11) (compared with

the autocorrelation peak) the number of atoms in the corresponding domain can be estimated. From this number of atoms the approximate domain volume can be calculated if the average space requirement of an atom is known. Other relationships concern the connection between radiation loading (size of Δt) and domain structure. An interesting parameter is the radiation loading time $\Delta \tau$ which corresponds (on the average) to one elastic scattering event per atom. If we apply the same radiation flux during less time than this "unity loading time" $\Delta \tau$ less radiation events will occur. The domains become larger – and not only an insignificant but also (in dark-field illumination) an incomplete structure will be found. Many atoms will not be hit by an incident electron and it is evident that these atoms cannot show up in the structure. It is evident that time intervals much longer than the unity interval $\Delta \tau$ (several scattered electrons per atom) are dangerous, because the atoms might have moved between the scattering events. The situation is somewhat different, if only a fraction of the scattered electrons can be utilized for the imaging process. Obviously in this case the unity $\Delta \tau$ can be multiplied by the relative factor of all scattered

electrons to the image forming electrons. But the traces are no more "elementary traces", produced by single radiation events.

The correlation procedure (11) gives a good account of the general behaviour of the radiation reaction but it is not very well suited for the practical determination of the traces. It cannot be recognized from (11), which $\Delta \tau_k$ corresponds to which domain. This problem disappears, if we use a method, which we call "cross correlation domain scanning". The procedure will be explained in Fig. 3 for the two-dimensional case. The generalization for three dimensions is evident. Let us assume that, for example, by (11) the average domain size has been determined. We specify now an arbitrary region in q (hatched circles in Fig. 3 a) having the average size of a domain and the (nearly) corresponding region in q' . We set the densities q and q' outside these regions to zero and calculate the cross correlation function. If it happens that the position of this region corresponds roughly to the position of a domain in q' (case A in Fig. 3 a), the cross correlation function contains only the displacement peak of this domain. It is easy to see that the complete information about the displacement vectors

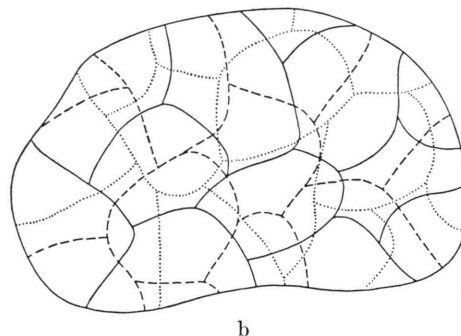
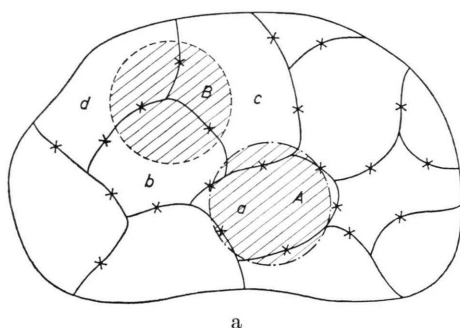


Fig. 3. Trace structure analysis with the "cross correlation domain scanning method" explanation of the principle for two schematic two-dimensional examples.

a) Assumption 1: specimen composed of relatively radiation resistant domains (e.g. aromatic groups), which are separated by radiation sensitive bonds. (A nucleic acid could perhaps be regarded as such a structure.) The boundaries of the domains a, b, c, ... are marked by lines with crosses representing sites of radiation reactions. While taking the images, a, b, c, ... will move in a non-predictable way. In the analysis, subsequent images (time interval Δt) are correlated within a circular region of the approximate size of a domain at the positions A, B, ... In position A we get one cross correlation peak which indicates the movement of a in the time interval Δt . In position B a spectrum of three correlation peaks will occur, stemming from the domains b, c, d. By appropriate scanning of the images the complete information on the domains and their shifts will be gained.

b) Assumption 2: specimen uniformly sensitive to radiation. In this case after each irradiation interval a new domain structure occurs which is characterized by boundary lines with respect to the previously imaged structure. The full, dashed and dotted lines are the boundaries corresponding to three subsequent time intervals Δt . (In Fig. 3 a these lines would coincide.) By cross correlation domain scanning of the successive images all domains can be determined. But the projection along the traces will only be possible within the smaller regions u, v, ... which have not been internally affected by a radiation induced effect during the integral illumination of $3 \Delta t$.

For very radiation sensitive specimens the average "projection length" might become very small. In this case a more complicated analysis (e.g. pattern recognition) becomes necessary. Radiation insensitive "isles" in a radiation sensitive matrix (e.g. aromatic side group in a protein) will, due to their stability, be recognized during this analysis.

$\Delta \mathbf{r}_k$ of all domains can be determined in a systematic way, if we scan ϱ and ϱ' synchronously with this region and use the peaks of the cross correlation function as a criterion. Only if the scanning region coincides with a domain, will a single $\Delta \mathbf{r}_k$ be registered. In intermediate positions (e.g. case B in Fig. 3a) one gets a "spectrum" of the displacement vectors of all domains, which will be (partially) screened out from the specimen by the scanning volume. But this "spectrum" is very much simpler than the "spectrum" of Eq. (11) as it contains only the few immediate neighbours. Moreover, the change of these spectra through scanning can even be used to roughly determine the shapes of the domains and to further determine the border walls within the specimen. This is an important prerequisite for the later projection procedure. Figure 4 shows the result of a displacement analysis for a domain consisting of the aromatic structure in Figure 2a.

The analysis discussed above is based on the tentative hypothesis that radiation damage is characterized by bond breakings and lateral shifts of the fragments (domains). It is easy to see that other variants of trace structure analysis can be worked out if radiation induced changes of the specimen follow other general laws. Let us, for example, assume that at Helium-temperature fragments do not displace laterally ("frozen structure"). Mainly local distortions appear (similar to radiation induced lat-

tice distortions). A similar analysis leads to a discrimination of non-distorted and distorted regions in the specimen. Almost nothing is known today about the molecular radiation chemistry, which is vitally important for these approaches. The other interesting aspect of trace structure analysis is that it also leads to an analysis of these laws.

Some Remarks Concerning the Physics Related to Trace Structure Analysis

It is evident that the mathematical and topological relationships discussed here only give the formal frame for the trace structure analysis. For a discussion of the physical facts, the number L of elastic scattering events per irreversible-bond-altering-event, is of fundamental importance. This parameter determines the number M of atoms in a domain. The earlier mentioned work on cross-sections of radiation damage can already give some information. This number L is essentially independent from dose rate (in the region between 10^{-6} and 10^2 Coul/cm²) and from the acceleration voltage to at least 1 MeV^{6, 7, 32, 33}. But L (and therefore M) is highly dependent upon the specimen. For the nucleic acid base adenine ($C_5N_5H_5$) (the most stable biological molecule studied by Isaacson, Johnson and Crewe⁷) one radiation damaging event in the molecule occurs for every 29 elastic scattering events ($L=29$). If one defines the domain by the size of the molecule ($M=10$), ~ 3 elastic scattering effects occur per atom. For thymine – the most sensitive base – L is reduced to 7. Similar figures can be found for other aromatic molecules. It should be noted that these figures concern only the native molecule. It is quite probable that other aromatic nucleins will withstand much higher doses. During longer exposures the molecule will be modified in a sequence of reactions leading to intermediate products until a radiation insensitive end product results. Isaacson, Johnson and Crewe⁷ have found that this end product has a different electron loss spectrum from a carbon foil. For trace structure analysis not only the number L of the native molecule but also the individual values of the intermediate products are important. It may be assumed that the later products show increasing stability. This can be estimated from general properties of the band model in solid state physics: The unstable conformations will be "erased". If such a general behaviour can be verified by experiment it might be useful to start trace

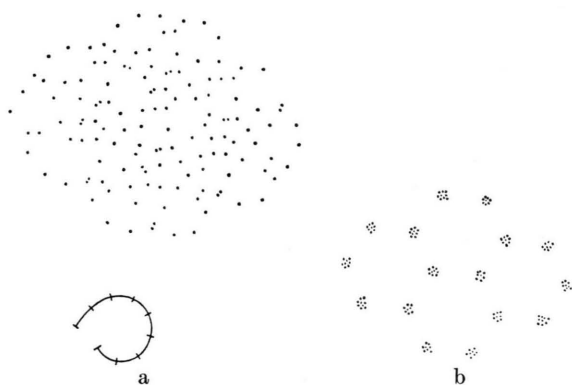


Fig. 4. Trace structure analysis, schematically shown for the aromatic group of Fig. 2a as a domain. It is assumed that no bond-breaking occurs within the group, but that the group will be forced to move (for simplicity only translations) due to environmental effects. a) Picture after an exposure time $8\Delta\tau$: the image is smeared by the movement and has a low resolution. b) Trace of the molecule, determined by correlation analysis of 8 images at $\Delta\tau$ exposure. c) Image of the structure in a) after trace structures analysis.

structure analysis not from the native specimen but – in a reversed sense – from the specimen at the end of the exposure. Thus the least critical steps towards the native specimen can already be done with some preliminary information with regard to the whole system.

For aliphatic molecules the situation is more critical. Here the cross-section for radiation damage seems to be of the same order of magnitude as the elastic cross-section. For example it has been shown that several proteins loose about half their mass after a dose corresponding to about one atom lost for every three elastic events ($L=3$)³⁴. It should be mentioned that it is difficult to draw conclusions with respect to the number of broken bonds from this figure. If the lost atoms leave the specimen as single atoms and if there are no intermediate bonds, there would be a 1:1 correlation. The appearance of low molecular products with several atoms would reduce the number of broken bonds, whereas bond binding inside the specimen would increase the number of radiation damage events³⁵. There is certainly no doubt that aliphatic molecules are more radiation sensitive than aromatic molecules. The domains might become so small that correlation analysis with sufficient significance could be impossible.

But there might be ways to solve the difficulties even in these cases. Biogenic macro-molecules – proteins and nucleic acids – are composed of a limited number of molecular groups (20 amino acids, 4 bases). It would therefore be possible to use more sophisticated correlation methods, which are focused to the determination of these groups rather than to the determination of single atoms. The problem is very similar to the reading of a written text by a computer. Pattern recognition methods are in development for these problems.

It might be mentioned that protein crystallography is faced (for an entirely different reason) with the same problem: The resolution is in general too small for the recognition of single atoms. Atomic groups must therefore be recognized in the density function. Today this "pattern recognition" will simply be done by fitting atomic models of the group into the unsharp Fourier synthesis (resolutions between 2 and 3 Å). If reaction chains for identical side chains in proteins are identical (or at least similar) even a pattern recognition based on the reaction scheme might be possible. The simplest

application would be, for example, the discrimination of aromatic and non-aromatic side chains in a protein by their different radiation sensitivity.

Another interesting possibility is the already mentioned radiation protection of organic molecules by heavy atoms as electron acceptors¹⁰. This proposal originates from theory and experiments in radiation chemistry. For other reasons electron microscopists have developed methods of implanting heavy atoms into a specimen (positive and negative staining). Combined with trace structure analysis not only the smaller radiation damage becomes important but also advantages in the execution of trace structure analysis. Every heavy atom can be regarded as a "centre" in the specimen, which can easily be recognized in all steps.

Experimental Requirements

One of the first points to be discussed in a scattering experiment is the intensity of the primary beam and the integral measuring time. In spite of the great amount of information to be gained by trace structure analysis the total exposure time is very short. If e.g. 100 elastic electrons should be collected for every light atom, the radiation dose at 100 kV would be 20 Coul/cm². This is somewhat more than the exposure of a conventional micrograph. The second point concerns the formation of the image. Experimentally every three-dimensional image will be divided into a number Q (say $Q \sim 30$) of projections or of related functions (if the curvature of the Ewald sphere cannot be neglected). Two principles – mechanical tilting of the specimen or electronic tilting of the primary beam – can be used. We must now bear in mind that the three-dimensional image is not significant. For example if we use for a three-dimensional image the unity loading time Δt only one elastic scattering event will occur for every atom. This indicates that the significance in every two-dimensional projection is even less (for $Q = 30$ projections this means one elastic event for 30 atoms). The exact position of the tilting axis has to be determined a posteriori by correlation functions between successive projections³⁶, as the tilting axis cannot be maintained with the high accuracy required. But very probably the low significance in the projection will make the correlations very difficult if not impossible.

There is no question that the best instrument for trace structure analysis would be a microscope with

electronic tilting of the primary beam¹⁹. Also with such an instrument the tilt axis³⁶ cannot be geometrically defined with the necessary precision. But the deviations will be reproducible, if the corresponding magnetic or electric fields are sufficiently stable. It is therefore possible firstly to project in the computer over several equivalent two-dimensional projections of time-successive three-dimensional images and then to do a single correlation for this set of three-dimensional images. For the tilt-correlation of projections the radiation damage corresponding to the simple superposition of three-dimensional images is unimportant³⁷. The electronic tilting has the further advantage that the repeated mechanical movements of the specimen (which in the case of a mechanical microgoniometer must occur) will be avoided. These remarks should not be understood in the sense that trace structure analysis with mechanical microgoniometers is impossible. The difficulty of tilt axis definition (caused by irreproducibility of mechanical movements on atomic scale) which makes the above described solution of superposition of equivalent projections impossible, could be overcome if two adjacent areas of the specimen would be analyzed in parallel, for example one with a time interval $\Delta\tau$ and the other with a time interval $\Delta t \gg \Delta\tau \approx 1/100 \dots 1/10 \Delta t$. Every projection of the second area would be thus defined with much better significance and could be used for the tilt correlations. The necessary electronic displacements of the (restricted) area of illumination and the simultaneous displacement of the images are already possible in some commercial microscopes.

There are some fundamental differences between three-dimensional image formation in dark field and in bright field (phase contrast). They derive from the fact that one is working with minimal doses. These differences would largely disappear, if one could work with higher doses i.e. better statistics. For the discussion of the differences occurring we use the simplest three-dimensional reconstruction scheme – backprojection. We assume that the specimen (or the primary beam) will be tilted around a single tilt axis. In this case the three-dimensional problem can be reduced to a two-dimensional one: The three-dimensional body can be regarded as a succession of discs perpendicular to the tilt axis (thickness of a disc \sim resolution in direction of the tilt axis). The image derived from the projections of a point in such a disc is a starlike figure. In the

limiting case of infinitesimal tilt angle increments this can be regarded as a continuous function. This function is not a point function in the usual sense³⁸, it radially decreases with $1/r$ ³⁹. But this is here of minor importance. We discuss firstly the dark field case. If the measurement has been done with good statistics, the shape of the starlike figure will be reproduced by the scattered electrons. But let us now assume that elastic scattering has taken place in a very small time interval during the tilting cycle. Then only one projection line will contain scattering information for this point. The other projection lines will be zero. The starlike image point degenerates into a straight line with an orientation, which depends on the accidental tilting angle at the time of elastic event. The straight line may be located according to the probability where the atom scatters in the projection.

Let us now discuss the same situation in bright field. In this case the bright field peak contains many electrons also in the limiting one elastic event case. Nevertheless the significance is not better than in the dark field case, because there is the additional statistical noise in the background. In a three-dimensional bright field image formation these electrons will be collected during the whole tilt cycle and we therefore get a peak (and not a line) as the representation of the atom in the reconstruction. Of course, also in this case the noisy bright field will lead to a picture with restricted significance. We note that in phase contrast a three-dimensional atomic image is similar in its structural features to a two-dimensional image in conventional microscopy, especially if the star-function has been converted to a project function of an appropriate form. The degeneration to a line will not occur. It would be worthwhile to consider this difference in the bright field and the dark field case in more detail and to treat it with probability methods to find its influence on the analysis of three-dimensional electron microscope image formation.

The relation between the equivalent number of bright field electrons and dark field electrons depends on the contrast in the bright field image. The intensity in the image plane is [see Eq. (3) in ⁴⁰]:

$$I = A_0^2 + A_0(\varrho + \varrho^*) + \varrho\varrho^*. \quad (12)$$

In the case of a single atom, ϱ is the amplitude function of the atomic image. We assume that the imaging will be done in an ideal microscope, equipped with a Zernike-plate, and that ϱ is real. We express

further the peak value ϱ_p as a fraction ε of the primary beam image $\varrho_p = A_0 \varepsilon$. The second and third term in Eq. (12) give the relative number v of bright field electrons to dark field electrons as roughly $2\varepsilon/\varepsilon^2 = 2/\varepsilon$ (different shapes of the atomic images in bright and dark field neglected, same aperture, central stop for the primary beam in the dark field case).

ε depends on the scattering cross section and on the imaging conditions. Let us discuss a special example: In ⁴¹ (compare Fig. 4a) the image of a C-atom has been calculated for the ideal conditions mentioned above ⁴² (no spherical aberration, no chromatic aberration, 100 keV-electrons, aperture $\alpha = 1.3 \times 10^{-2}$). ε is of the order 0.05. 40 bright field electrons correspond therefore to one dark field electron. Due to the restricted aperture only 35% ⁴³ of the elastically scattered electrons will be utilized. If one opens up the aperture to $\alpha = 2.6 \cdot 10^{-2}$ ($r^* = 0.7 \text{ \AA}^{-1}$), about 70% ⁴⁴ of the elastically scattered electrons will be used for the image. ε is then roughly 0.1 and 20 electrons correspond to one dark field electron ⁴⁵. We see that the factor $v = 2/\varepsilon$ is of the correct size for the elimination of "line degeneracy" in phase contrast. It is evident, that the factor v will change with the accelerating voltage. Higher voltage will mean lower cross sections and therefore an increase of the factor v . Until now we have not discussed radiation damage by knock on effects. They will hardly occur for light atoms (except for H) at medium voltages. The probability of a knock-on event is low ($\sim 1/3000$ ⁴⁶) but at higher voltages additional radiation damage might occur due to the kinetic energy of the knocked-on atoms. In trace structure analysis knock-on events will show up in a different way than events in the electron shell.

Computer Requirements

Trace structure analysis means replacement of a single three-dimensional image reconstruction by N image reconstructions. It is therefore necessary to minimize the computer time for a single reconstruction. According to our present experience, a three-dimensional reconstruction will take approximately 3 minutes (reference computer IBM 360/91), if electronic tilting is used. If mechanical tilting must be used, there is an additional factor of about 3 for the time consuming operations to determine the parameters of the transfer function and for the cor-

relations between projections, which have to be done for all images instead of for a group of images. If N is of the order of 100, the total computer time would be of the order of 5 hours (respectively 15 hours). This estimate does not include the pattern recognition programs of trace structure analysis. It can therefore be seen that even with the existing computers the problems can be handled. But let us analyse the situation further: Most of the operations are matrix operations in a general sense, mainly Fourier operations and related procedures (for correlations). For this type of calculations the existing single processor computers are very uneconomical. The important point is that many calculations can be done simultaneously in vector type calculations, as they do not depend on the results of concurrent parallel computations. This is the reason that one of the simplest multichannel analogue computers for Fourier inversion – the light diffractometer – is by orders of magnitude faster than the most advanced present electronic computer. It is obvious that by using digital multi-parallel computers the computing times can be considerably reduced. In fact, development of such computers is already in progress, very much stimulated by the success of integrated micro-computers which are fabricated in mass production. It can therefore be expected that the computational difficulties will be overcome in near future and that perhaps dedicated computers or processors can be used for the special operations in trace structure analysis.

The last point concerns the more general pattern recognition programs, which are necessary for the interpretation. These programs in an intelligent form make use of all information in the "cinematographic three-dimensional film" of trace structure analysis. It is obvious that these procedures cannot be done by hand (similar to the present pattern recognition in protein crystallography). Evaluating programs have to be written. On the other hand, this type of logical program will by then be used in many branches of science (documentation, language translation etc.) and it is unlikely that fundamental difficulties might occur (especially if real atomic resolution has been achieved experimentally).

This research has been supported by the Deutsche Forschungsgemeinschaft, Sonderforschungsbereich 51, Medizinische Molekularbiologie und Biochemie in München. We should like to express our sincere thanks.

- ¹ e.g. holography in the diffraction plane or ptychography²⁻⁴. These methods are in the initial stage of development and will not be discussed here. Reconstruction of electron microscope images to improve resolution (see e.g.⁵) may also be understood as image plane holography.
- ² W. Hoppe, *Acta Cryst. A* **25**, 495 and 508 [1969]; W. Hoppe and G. Strube, *Acta Cryst. A* **25**, 502 [1969].
- ³ R. Hegerl and W. Hoppe, *Ber. Bunsenges. Physikal. Chem.* **74** (Nr. 11), 1148 [1970].
- ⁴ R. Hegerl and W. Hoppe, *Proc. Fifth Europ. Congr. on Electr. Micr.* 1972, Manchester, p. 628.
- ⁵ W. Hoppe, *Proc. Fifth Europ. Congr. on Electr. Micr.* 1972, Manchester, p. 612.
- ⁶ R. M. Glaeser, *J. Ultrastruct. Res.* **36**, 466 [1971].
- ⁷ M. Isaacson, D. Johnson, and A. V. Crewe, *Radiat. Res.* **55**, 205 [1973].
- ⁸ J. R. Breedlove jr. and G. T. Trammell, *Science* **170**, 1310 [1970].
- ⁹ G. Siegel, *Microscopie Electronique* 1970, Grenoble, Vol. II, 221.
- ¹⁰ D. Schulte-Frohlinde, *Ber. Bunsenges. Physikal. Chem.* **74**, 1196 [1970].
- ¹¹ W. Hoppe, R. Langer, G. Knesch, and Ch. Poppe, *Naturwiss.* **55**, 333 [1968].
- ¹² J. G. Allpress, *J. Solid State Chem.* **2**, 78 [1970].
- ¹³ S. Iijima, *J. Appl. Phys.* **42**, 5891 [1971]; S. Iijima, J. M. Cowley, and G. Donnay, *Tscherm. Min. Petr. Mitt.* **20**, 216 [1973]; P. R. Buseck and S. Iijima, *Amer. Min.* **59**, 1 [1974].
- ¹⁴ W. Hoppe, *Naturwiss.* **61**, 239 [1974].
- ¹⁵ P. B. Hirsch, A. Howie, R. B. Nicholson, D. W. Pashley, and M. J. Whelan, *Electron Microscopy of Thin Crystals*, Butterworths, London 1965, p. 311 ff.
- ¹⁶ A. Feltynowski and W. Hoppe, *Eighth Intern. Congr. on Electr. Micr.* 1974, Canberra, Vol. I, 206.
- ¹⁷ D. J. De Rosier and A. Klug, *Nature London* **217**, 130 [1968].
- ¹⁸ 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 1973, p. 91 ff. (see Fig. 1 c, d, e).
- ¹⁹ W. Hoppe, *Z. Naturforsch.* **27a**, 919 [1972]; see also ¹⁸, Fig. 1 a, b.
- ²⁰ The real space refinement of protein crystal structures, introduced by Diamond²¹ is an example in crystallography.
- ²¹ R. Diamond, *Acta Cryst. A* **27**, 436 [1971].
- ²² O. Scherzer, *Ber. Bunsenges. Physikal. Chem.* **74**, 1154 [1970].
- ²³ Inversion: Replacement of all atomic vectors by their negative values, $\tilde{q}_1' = q_1' (-r - \Delta r)$.
- ²⁴ M. J. Buerger, *Vector Space*, J. Wiley & Sons, Inc., New York 1959.
- ²⁵ If there are atoms with a shape defined by an atomic scattering factor f (instead of points), the shapes of the convolution peak are defined by f^2 (see the text-books on crystallography). f^2 shows a stronger decrease with scattering angle as f . Therefore the convolution peaks are broader.
- ²⁶ This is in fact not true; they contain important information about the internal structure of q_1 .
- ²⁷ W. Hoppe, *Z. Elektrochemie u. angew. physikal. Chemie* **61**, 1076 [1957].
- ²⁸ R. Huber, *Acta Cryst.* **19**, 353 [1965].
- ²⁹ It would certainly be interesting, if one could also collect structural information via inelastic scattering. E. g. in the case of carbon it is wellknown that on the average only every fourth scattering event is elastic. The collection of this information is possible, if a transmission electron microscope is used. But unfortunately inelastic scattering is in general delocalized and therefore a structure derived from inelastic scattering will show up in a very bad resolution. This has been predicted in³⁰ and experimentally verified in³¹. Whether a parallel trace analysis with inelastic scattering could nevertheless give important additional information cannot yet be decided.
- ³⁰ W. Hoppe, *Z. Naturforsch.* **26a**, 1155 [1971].
- ³¹ M. Isaacson, J. P. Langmore, and H. Rose, *Optik* **41**, 92 [1974].
- ³² L. E. Thomas, C. E. Humphrey, and D. T. Grubb, *Radiat. Effects* **3**, 89 [1970].
- ³³ K. Kobayashi and K. Sakaoku, *Quantitative Electr. Micr.*, Ed. G. F. Bahr and E. H. Zeitler, Williams & Wilkins, Baltimore 1965, p. 359 ff.
- ³⁴ J. Wall, *Proc. 30th Ann. Conf. EMSA* 1972, p. 186.
- ³⁵ We propose to do the measurements at Helium-temperature. In this case the small molecular products would remain at least partially in the specimen and they could be found in the analysis and coordinated.
- ³⁶ or, if cone-tilting will be used, the apex of the cone.
- ³⁷ It is obvious that the superposition of all images leads to the image which would be achieved by the "conventional" three-dimensional imaging process.
- ³⁸ zero except within one resolution element (δ -function).
- ³⁹ One can show that the unusual shape of the "star image point" can be converted into the shape of a point function using more sophisticated reconstruction schemes (e.g. q -filtered backprojection). This conversion can be done independently from the statistical significance.
- ⁴⁰ W. Hoppe, *Acta Cryst. A* **26**, 414 [1970].
- ⁴¹ R. Langer and W. Hoppe, *Optik* **24**, 470 [1966/67].
- ⁴² in a nonideal configuration the contrast becomes lower and therefore $2/\varepsilon$ greater.
- ⁴³ see Fig. 8 in³⁰. $|r^*|$ in the case treated above is $r^* = a/\lambda = 0.354 \text{ \AA}^{-1}$.
- ⁴⁴ see again³⁰, Figure 8.
- ⁴⁵ The value of ε from⁴¹ seems to be very high. It should be noted that, depending on the atomic model, the scattering amplitudes differ very much in the literature, especially for low- Z atoms and low scattering angles. Therefore the values of ε calculated from the scattering amplitudes of different authors vary up to a factor of about 2.5 for carbon and $\vartheta_{\max} = 0.013$. The following table gives some examples for $Z=6$:

Source for $f(\vartheta)$	atomic model	$\vartheta_{\max} = 0.013$	$\varepsilon \vartheta_{\max} = 0.026$
F. Lenz, <i>Z. Naturforsch.</i> 9a , 185 [1954].	Wentzel	0.039	0.100
J. A. Ibers and J. A. Hoerni, <i>Acta Cryst. A</i> 7 , 405 [1954]	TFD	0.054	0.097
H. P. Hanson et al., HFS Atomic Structure Factors, Welch Found., 1964	HFS	0.034	0.087
H. Raith, <i>Acta Cryst. A</i> 24 , 85 [1968]	TFD	0.022	0.068
O. Scherzer, <i>Ber. Bunsenges. Physik. Chem.</i> 74 , 1154 [1970]	korr.		
J. Haase, <i>Z. Naturforsch.</i> 25a , 936 [1970]	Wentzel	0.034	0.087
J. Haase, <i>Z. Naturforsch.</i> 25a , 1219 [1970]	HFS	0.034	0.085
	TFD	0.029	0.074

(private communication by D. Typke).

- ⁴⁶ V. E. Cosslett, *Ber. Bunsenges. Physikal. Chem.* **74**, 1171 [1970].