Classification

Physics Abstracts

61.00 — 61.16D — 42.30

Structure analysis by crystallographic image processing

Hommage à Jean Baptiste Joseph Fourier (1768-1830)

Sven Hovmöller

Structural Chemistry, Stockholm University, S-106 91 Stockholm, Sweden

(Received October 03, 1990; accepted February 20, 1991)

Résumé. — Il est possible de déterminer la structure cristalline d'un solide inorganique à l'échelle atomique en associant la microscopie électronique haute résolution avec un traitement cristallographique de l'image. La micrographie électronique constitue une image agrandie mais distordue de l'objet. Par l'analyse quantitative de cette image, il est possible de déterminer les distortions et de les corriger pour la plupart. Il est aussi possible de déterminer la symétrie du cristal et, avec une précision d'au moins 0,1 Å, la position des atomes. Le traitement d'image est basé sur l'analyse de Fourier des données expérimentales contenues dans la micrographie. La symétrie cristallographique est déterminée et est utilisée largement pour retrouver une image pratiquement non distordue.

Abstract. — It is possible to determine the crystal structure of inorganic compounds to atomic resolution, by combining high resolution electron microscopy (HREM) with crystallographic image processing (CIP). An electron micrograph is a magnified, but distorted, image of the object. By quantitative analysis of the image it is possible to determine the distortions and correct for most of them. It is also possible to determine the symmetry of a crystal and even to find the atomic positions with an accuracy of at least 0.1 Ångström. The image processing is based on Fourier analysis of the experimental data in the form of a micrograph. The crystallographic symmetry is determined and made maximal use of in restoring a nearly distortion-free image.

1. Historical background.

Jean Baptiste Joseph Fourier was 21 years old when the French Revolution broke out, just as was Napoleon Bonaparte. Their lives went to a large extent in parallel and in dependence of each other. Both were interested in mathematics, politics and military affairs. Fourier accompanied Napoleon to Egypt in 1798 and was governor of part of Egypt for three years. Fourier did not have a too good success in military affairs and returned to France in 1801. From then on he concentrated all his efforts in science, mainly the problem of heat conductivity. He discovered what is now called Fourier's theorem in 1807.

Fourier's theorem states that any periodic (repeating) function can be described as the sum of a series of simple sine and cosine waves. The Paris Academy wanted to encourage Fourier to extend and improve his theory, so they selected the question of the propagation of heat as the subject of the *Grand Prix de Mathématiques* for 1812. Among the committee members who gave the first prize to Fourier were Laplace, Lagrange and Legendre. However, they were not completely happy with the rigour of the proofs so the paper was not published in the *Mémoires de l'Académie des Sciences* until 1824 [1], when Fourier had become the Secretary of the Academy.

2. Fourier analysis.

Today Fourier analysis is an indispensable mathematical tool which is used in a wide range of applications in physics, from telecommunications to crystallography. In this paper the usefulness of Fourier analysis for structure determination of inorganic crystals by electron microscopy will be demonstrated. In molecular biology similar procedures are used routinely in many laboratories for the study of thin protein crystals, especially of membrane proteins. This work was pioneered by Unwin and Henderson in 1975 in the study of bacteriorhodopsin to 7 Å resolution [2], and has recently led to the first atomic resolution structure determination of a membrane protein by electron microscopy [3].

Many physical properties of crystals are periodic functions, with the unit cell dimensions as the period. According to Fourier it is therefore possible to describe for example the electron density of a crystal equally well in real space and in the form of its Fourier series, in what crystallographers call reciprocal or Fourier space.

A Fourier series is in the form

$$a_0 + (a_1 \cos x + b_1 \sin x) + (a_2 \cos 2x + b_2 \sin 2x) + \dots$$

where a_0 is a constant term, namely the average of the function over the whole period. The terms a and b are called the Fourier coefficients and they represent the amplitude of each component wave. In a parallel to music the first term ($a_1 \cos x + b_1 \sin x$) can be called the fundamental tone and the following terms the overtones. Note that the frequencies of the overtones are simply integer multiples of the fundamental tone. Translated to the world of X-ray crystallography it means that the electron density can be described as:

a_0 = the average electron density (about $1 \text{ e}^-/\text{\AA}^3$) + $\Sigma (a_n \cos nx + b_n \sin nx)$ = many waves of positive and negative contributions that modulate the electron density.

The first Fourier coefficient after the constant gives information about if the density is concentrated in broad bands through the unit cell, in which case the amplitude is large, or if the density is more uniformly spread out over the unit cell, in which case the a_1 and b_1 terms are small. As more and more coefficients are added, i.e. as the resolution is increased, finer and finer bands of density are added and the structure becomes more and more clear, until finally atomic resolution is reached. The scattering process by electrons is different from that of X-rays in that also the nuclei contribute to the scattering of electrons, and that electrons are scattered more than a thousandfold stronger than are X-rays. However, the electron diffraction pattern of a thin crystal closely resembles that of the X-ray diffraction pattern of a thicker crystal, and most principles known from X-ray diffraction can be applied in a resembling way to electron microscopy, but with some important differences noted below.

Since atoms are separated by some 1.5 to 2.0 Å, it is necessary to include Fourier coefficients to more or less this value in order to see all atoms. However, the distance between metal atoms in oxides is in the order of 3.5 to 4.0 Å, so these can be separated already at a resolution of about

3 or 4 Å. For a crystal with a unit cell dimension of 12 Å, the different Fourier coefficients along that axis represent periodicities of 12 Å, 6 Å ($= 12/2$), 4 Å ($= 12/3$), 3 Å ($= 12/4$) etc.

3. Diffraction patterns and Fourier transforms.

Each Fourier coefficient is described by three numbers; the frequency, the amplitude and the phase. The frequency tells which Fourier coefficient we are talking about. The amplitude is the size of the coefficient, and it is calculated as $\sqrt{(a^2 + b^2)}$. The phase is $\arctan(b/a) + 180^\circ$ if a is negative.

The amplitude is very simple to explain in physical terms as the strength of the planes of density through the crystal. The phase is really no more difficult to explain; it is the position of the maximum of each plane, relative to a fixed point in the unit cell, the origin.

For a 3-dimensional function, for example a crystal, the Fourier transform also becomes 3-dimensional. A 3D crystal is a superposition of density waves in different directions as well as with different frequencies. The diffraction pattern of a crystal, obtained by X-rays or electrons, is just the Fourier transform of the crystal structure, except for two things. The intensities of a diffraction pattern are the squares of the amplitudes, but more seriously, the phase information of the Fourier transform is lost. This causes the well-known phase problem in crystallography. Without the phase information it is not possible to solve a structure even if all amplitudes are known, since we do not know where to add and where to subtract the electron density in the unit cell. The phase problem is overcome in X-ray crystallography by different methods of estimating the phases.

4. Phases in EM images.

Electron microscopy has a major advantage over X-ray and electron diffraction: the phases are recovered in the EM image. The reason for this is that the diffracted electrons are focussed into an image, something which is not possible to do with X-rays. Just as the crystal, the electron micrograph is also a periodic function in real space. It is an enlarged but more or less distorted image of the crystal. The phases can easily be retrieved in numerical form from the Fourier transform of the image. The procedure for this is to digitize the negative, or rather a small part of it, and then calculate the Fourier transform by computer.

5. Distortions.

There is only one snag. The electron micrograph is not a perfect enlarged image of the crystal. The image has suffered a certain deterioration due to dynamical scattering inside the crystal, electron optical distortions in the microscope and finally random noise due to the limited number of electrons that make up the image, and the silver crystals in the photographic plate. This can all be summarized as:

$$\text{Structure} + \text{dynamical events} + \text{optical distortions} + \text{noise} \rightarrow \text{image}$$

The random noise level is surprisingly low in HREM pictures of inorganic crystals. In fact, the random noise is only a few percent of the signal, and it is so low that it causes problems! With a higher random noise level it becomes easier to determine the level of defocus and the astigmatism, by looking at the white noise background in the F.T. We have estimated the random noise level by comparing two adjacent but non-overlapping areas of a very uniform crystal of $\text{K}_7\text{Nb}_{15}\text{W}_{13}\text{O}_{60}$ [4]. The amplitudes were identical to within 7 percent, and the phases differed by as little

as 5 degrees between the two parts of the image. These values should be compared to the systematic (non-random) noise level of the same image which were much higher. This crystal, which is orthorhombic [5] must show mm-symmetry of the diffraction pattern, i.e. reflection (hk) and $(h, -k)$ should have the same amplitudes. For the same image as analysed above, the symmetry R -value was as bad as 66%, and the phase residual from the centrosymmetric case (all phases 0 or 180 degrees) was 17 degrees. This shows clearly that systematical errors, due to electron optical distortions and crystal tilt are much more severe problems than the random noise. The random noise is essentially eliminated by simply averaging all the scanned unit cells, i.e. applying P1 symmetry to the crystal. The strength of CIP lies in its ability to overcome also a large part of the systematical errors.

Dynamical effects are more serious than the random noise, and have caused great worries among electron microscopists, even to the extent that the consensus seems to be that the image hardly bears any resemblance to the original crystal structure. For this reason image simulations are commonly used as a way to verify the structure determinations by HREM. However, the dynamical effects can be eliminated to a very large extent by choosing a thin area (<5 nm for metal oxides) of the crystal for image processing, and by applying Crystallographic Image Processing, CIP.

6. Electron microscopy.

High resolution electron micrographs of thin areas of wedge-shaped crystals are recorded, for example at 200 kV. The better the alignment of the microscope and of the crystal is, the easier are the following procedures. A single image taken near Scherzer focus is sufficient for the analysis, but it is also possible to combine information from several images, in order to check the reproducibility and accuracy of the experimental data.

7. Symmetry determination.

It is important to know the symmetry of a crystal in order to make full use of the experimental data and obtain maximal possible resolution. It is possible both to over- and underestimate the symmetry. When overestimating the symmetry (for example assuming 4-fold symmetry for a crystal which in fact only has 2-fold symmetry) chemically different regions of the unit cell will be averaged. The result is a smearing out of the structural details. If, on the other hand, the symmetry is underestimated, we do not take the full advantage of the symmetry, and the resulting map will be more noisy than was necessary.

Probably the most common way to determine the space group in the electron microscope is by convergent beam electron diffraction (CBED). However, it is also possible to determine the symmetry of a crystal by combining electron diffraction and CIP. For a single projection only the 2D projection symmetry will be determined, and usually it is necessary to have several projections in order to determine the 3D space group.

The symmetry determination using electron diffraction data is in principle similar to space group determination in X-ray crystallography, but there are practical differences. Because of the almost perfectly kinematical data X-ray diffraction and the high resolution (often 1 Å or better) a lot of good quality data is collected and it is possible to make an accurate space group determination. Unlike the HREM image, electron diffraction data are unaffected by most of the distortions of the objective lens, and they go to high resolution. It must be emphasized that a slight deviation from perfect symmetry in the diffraction pattern quite often is only a reflection of slight misalignment of the crystal, and should not be taken as an indication of a lower symmetry. Similarly the

presence of systematically absent reflections due to screw axes may be an effect of dynamical scattering in the thicker parts of the diffracting crystal. In X-ray diffraction the presence of a single weak so-called forbidden reflection should be interpreted as a proof of the absence of a screw axis. In electron diffraction there may well be some intensity in all the forbidden reflections for a crystal with a screw axis. Whenever the odd index reflections are much weaker than the even ones, a screw axis must be considered.

A further help in symmetry determination comes from the phases, as they are read out from the Fourier transform of the HREM image. Pairs of reflections (h, k) , $(-h, k)$ have equal phases if the h axis is a 2-fold rotation axis, but differ by $(h \cdot 180^\circ)$ if h is a screw axis. Since there are many more general reflections (h, k) than axial reflections $(h, 0)$ this method of looking for screw axes also gives better statistics than just looking at a few axial reflections.

8. Optical diffraction.

An optical diffractometer is a very useful tool for assessing the qualities of different images. In amorphous regions it is easy to see the defocus value, and check for possible astigmatism. In the crystalline area it is possible to observe crystal tilt; the diffraction pattern is symmetrical in the thin part of the crystal, but in the thicker part only the reflections on the tilt axis remain to high resolution.

Beam tilt causes more serious distortions than crystal tilt [6], but are not detected in the optical diffractometer, since beam tilt mainly affects the phases and not the amplitudes, which are the only things we see in the optical diffractometer.

The areas that are useful for image processing are selected from the thin edges of the crystals in the optical diffractometer. Only areas showing the full symmetry of the crystal and highest possible resolution (as judged by the presence of sharp diffraction points with high indices) are selected.

9. Digitizing an electron micrograph.

In order to perform image analysis it is necessary to translate the analog information of the micrograph into numerical form. Until now this has been a tedious and expensive step. Microdensitometers at the cost of some FF 1 million had to be used until the CCD cameras very recently became available for about one tenth of this cost. Meanwhile the time needed to scan an image decreased from several minutes to a fraction of a second. With an image grabber the image can be directly accessed by a standard personal computer.

The information on an electron micrograph is stored in the silver grains, each with a size of about $1 \mu\text{m}^2$. When the micrograph is digitized, the grey level (from for example 0 for white to 256 for black) is measured over square areas much larger than the silver grains. The magnification of the electron micrograph should be chosen such that areas of about 20 to $50 \mu\text{m}$ could be used. The scan size also depends on the resolution, as determined on the optical diffractometer. A typical example is the following: an image taken at 500 000 times magnification shows diffraction peaks to $2.5 \text{ \AA}$ resolution. $2.5 \text{ \AA}$ times 500 000 equals $125 \mu\text{m}$. The finest details of the crystal are then $125 \mu\text{m}$. The scanning raster must be more than 2 times finer than the finest details we expect to find. Thus a scan of 40 or $50 \mu\text{m}$ would be appropriate in this case. No more information is picked up by using a finer raster, but the scanned area will be smaller. An area of some 100 unit cells is often scanned, and this may represent about one cm^2 on the micrograph.

10. Fourier transformation.

The Fourier transformation of the scanned image is calculated by some standard mathematical package. If the amplitude part of the Fourier transform is calculated (by calculating $\sqrt{a^2 + b^2}$ for every point of the transform) it can be displayed on a computer graphics screen. What we see then is the same diffraction pattern as we saw on the optical diffractometer.

The lattice can now be indexed and the basic lattice vectors can be determined. This must be done very accurately using a least squares procedure.

The amplitude of each diffraction peak is calculated by integrating over a small area (for example 3 by 3 pixels) around the exact predicted lattice point. The phase is determined as the phase value (calculated as mentioned earlier) at the pixel closest to the predicted position of the peak.

11. Origin refinement.

The phases are, as mentioned above, the position of the maximum of a density wave in the micrograph, relative to a common origin. Initially this origin is a random point in the unit cell, namely the center of the scanned area. It is necessary to move the origin to a more reasonable place, such as a center of symmetry or a 4-fold axis etc. There are special computer programs that search for the positions of different symmetry elements in the unit cell. They also give a phase residual, from which the certainty of the symmetry can be estimated. When the position of a symmetry element is found, this is taken as the new origin or reference point for all the reflections. When the origin is shifted by a distance (x, y) the phase of each reflection is changed by $360^\circ \cdot (h \cdot x + k \cdot y)$.

12. Phase adjustments.

Most inorganic crystals have a centrosymmetric projection. In this case all phases must be exactly 0° or 180° . If the phase residual is less than about 20 degrees then it is a simple matter to assign the correct phase to all reflections, except perhaps a few very weak ones. At a certain resolution it starts to become very difficult to assign the phases with any confidence, as seen for example by symmetry-related reflections giving mis-matching phase indications. It is advised to cut out the data above this resolution.

The main reason for phases deviating from the ideal values of 0° or 180° is beam tilt.

An extensive description of different procedures that can be used for eliminating the different optical distortions in the case of a protein crystal is given in [7].

13. Amplitude adjustments.

The amplitude adjustment can be done in a simple way, by just averaging the symmetry-related reflections and giving each of the reflections this average value. Although more sophisticated schemes are possible, this crude way is often quite sufficient, because the amplitudes need not be all that accurate to solve the structure. It is far more important that the phases are correct.

One of the main reasons for amplitudes losing their symmetry is crystal tilt. Diffraction points along the tilt axis are unaffected by the tilt, but the points perpendicular to the tilt axis decrease in amplitude, proportionally to the tilt angle multiplied by the crystal thickness.

Also the C.T.F., astigmatism and beam tilt influence the amplitude values. While the crude way of improving amplitudes just by averaging over symmetry-related reflection does give good

results to 2.5 Å resolution, it would be desirable to make a more precise correction for the C.T.F., astigmatism, beam tilt etc. This should in principle be possible, since only a very limited number of distortion parameters need to be found, and the large number of reflections, at least for crystals with unit cells larger than about 10 Å, give a lot of highly accurate data. The mathematical problem lies of course in the fact that all the different distortions are non-linear and interfere with each other.

14. Calculation of a density map.

The list of amplitudes and phases, with the correct crystallographic symmetry imposed as above, are used as input to an inverse Fourier transform. The result is a virtually distortion-free density map.

If the resolution is better than 3 Å all metal atoms in oxides will be resolved in a 3D map. A resolution around 2 Å is needed to resolve also the oxygens, or to resolve metal atoms in an alloy. Many metal oxides, especially among the much studied niobium oxides, have one short crystal axis (3.9 Å). A projection down that axis shows the whole structure with all metal atoms well resolved. For structures with metal atoms closer in a projection it is necessary to increase the resolution or take images from more than one direction in order to resolve such metal atoms.

It is worth mentioning that it is easier to correct for a large part of the distortions than it is to estimate them accurately. However, because of the abundance of accurate data (especially for large and highly symmetrical unit cells) the data can also be used for a quantitative estimate of the distortions, beam tilt, defocus, astigmatism and crystal tilt.

15. The difference between solving a structure and refining it.

A structure determination consists of two very different parts; solving the structure and refining it. Both logically and experimentally these two things are completely different, yet they are often confused. In order to solve a structure all tricks are allowed. One may guess the structure, but this becomes impossible when the structure is not extremely simple. The structure can also be deduced from a series of similar compounds, the structures of which are already known. Finally there are all the methods of phase estimations in X-ray crystallography (The Patterson method, direct methods, isomorphous replacement etc.) which may give at least part of the structure.

The refinement of a structure is a different matter. Here there are no such things as guesses. The preliminary atomic positions are given, and from those a calculated Fourier transform is obtained (with phases now, since we have a model function in real space!) and compared with the experimental diffraction data. The atomic positions are shifted very slightly (0.1 Å or less) by a least squares program that tries to match the calculated and observed diffraction patterns.

Crystallographic Image Processing, CIP, is a method to solve crystal structures. It is not a method for refining them. Unfortunately it seems a very difficult task to refine a structure from EM or electron diffraction data, because of the relatively low amount of data points, and more seriously, because of the poor quality, especially of the amplitude data and the electron diffraction data.

What then is image simulations? It is not a method for solving structures, since the structure model is needed as input to the simulation program. And it is also not a structure refinement, since there is no least squares refinement of the model against experimental data involved. In fact the comparison of the simulated image with the experimental one is not done quantitatively by comparing the amplitudes and phases of the two images, but only qualitatively by looking at them

Crystallographic Image Processing (CIP)

1. Select the best area of an HREM image near the edge. Digitize the selected area for example as 256×256 pixels of $40 \mu\text{m} \times 40 \mu\text{m}$.

2. Calculate the Fourier transform of the scanned area and determine the lattice.

3. Extract the amplitudes and phases from the Fourier transform.

Index (h, k)	Observed amplitudes		Observed phases	
	(h, k)	(h, -k)	(h, k)	(h, -k)
1 1	1020	1459	147	161
2 1	2378	2270	158	327
2 2	2891	2763	338	337
3 1	1558	1697	328	330
3 2	1533	1760	134	308
etc.				

4. Impose crystallographic constraints on amplitudes and phases.

Index (h, k)	Deduced amplitudes		Deduced phases	
	(h, k)	(h, -k)	(h, k)	(h, -k)
1 1	1239	1239	180	180
2 1	2324	2324	180	0
2 2	2827	2827	0	0
3 1	1628	1628	0	0
3 2	1647	1647	180	0
etc.				

5. Calculate an undistorted density map of the structure by inverse Fourier transformation of the corrected amplitudes and phases.

Fig. 1. — Schematic overview of the steps involved in CIP.

by eye. Image simulations is a widely used method for confirming the correctness of a structure which has been derived from HREM images. Image simulation is very useful for structures closely related to other, known structures, but it cannot be used to solve large, completely unknown structures, because the number of possible models to test becomes almost infinite.

16. Applications.

The structures of many niobium oxides are especially easy to determine by HREM, since they often have projections along a short (3.9 Å) unit cell axis, in which all metal atoms are seen resolved and non-overlapping. We have studied three niobium oxides; $K_7Nb_{15}W_{13}O_{80}$ [5], $Na_3Nb_{12}O_{31}F$ [8] and $Cs_8Nb_{54}(O, F)_{146}$ [9], and in each case we were able to determine the atomic coordinates to within at least 0.1 Å accuracy.

A slightly more difficult group of compounds are the perovskites, where the metal atoms are closer together in projections, and where the question often is to localize oxygen vacancies. Good results have been obtained for two perovskites $Ca_4Fe_2Ti_2O_{11}$ [10] and $Ba_2Fe_2O_5$ [11].

References

- [1] FOURIER J.B.J., *Mém. Acad. Sci.* **4** (1822) 185; **5** (1822) 153.
- [2] UNWIN P.N.T. and HENDERSON R., *J. Mol. Biol.* **94** (1975) 425.
- [3] HENDERSON R., BALDWIN J.M., CESKA T.A., ZEMLIN F., BECKMAN E. and DOWNING K., *J. Mol. Biol.* **213** (1990) 899.
- [4] SJÖGREN A., Doctoral dissertation, Chem. Comm. Univ. Stockholm No 8, (1988).
- [5] HOVMÖLLER S., SJÖGREN A., FARRANTS G., SUNDBERG M. and MARINDER B.-O. *Nature* **311** (1984) 238.
- [6] SMITH D.J., SAXTON W.O., O'KEEFE M.A., WOOD G.J. and STOBBS W.M. *Ultramicroscopy* **11** (1983) 263.
- [7] HENDERSON R., BALDWIN J.W., DOWNING J.W., LEPAULT J. and ZEMLIN F. *Ultramicroscopy* **19** (1986) 147.
- [8] LI D.X. and HOVMÖLLER S., *J. Solid State Chem.* **73** (1988) 5.
- [9] WANG D.N., HOVMÖLLER S., KIHILBORG L. and SUNDBERG M. *Ultramicroscopy* **25** (1988) 303.
- [10] HOVMÖLLER S., ZOU X.D., WANG D.N., GONZALEZ-CALBET J.M. and VALLET-REGI M., *J. Solid State Chem.* **77** (1988) 316.
- [11] ZOU X.D., HOVMÖLLER S., PARRAS M., GONZALEZ-CALBET J.M. and VALLET-REGI M., to be published.