High Resolution Electron Microscopy: 
From Imaging towards Measuring

S. Van Aert(1), A.J. den Dekker(1), A. van den Bos(1), and D. Van Dyck(2)
(1)Department of Applied Physics, Delft University of Technology
Lorentzweg 1, 2628 CJ Delft, The Netherlands
Phone: +31 15 278-1823, Fax: +31 15 278-4263, E-Mail: s.vanaert@tn.tudelft.nl, WWW: http://www.tn.tudelft.nl/mmr
(2)Department of Physics, University of Antwerp (RUCA)
Groenenborgerlaan 171, 2020 Antwerp, Belgium
Phone: +32 3 218-0258, Fax: +32 3 218-0318, E-Mail: dvd@ruca.ua.ac.be

Abstract – High resolution electron microscopy is discussed as a measuring, rather than an imaging technique. It is shown that the interpretation of the images could greatly benefit from a quantitative instead of a qualitative approach accompanied by quantitative statistical experimental design.

Keywords – quantitative model-based high resolution electron microscopy, parameter estimation, precision, experimental design.

I. INTRODUCTION

Scientists manage to control the structure of materials on an ever finer scale. The last decades are characterized by an evolution from macro- to micro-, and more recently, to nanotechnology. Examples are numerous: nanoparticles, nanotubes, quantum transistors, layered magnetic and superconducting materials [1]. The interesting properties of these materials are mainly related to their nanostructure. In parallel, one sees an evolution in solid state theory where materials properties are increasingly better understood from first principles theoretical calculations. The merging of these fields will enable materials science to evolve into materials design, that is, from describing and understanding towards predicting of materials properties. Materials properties are strongly connected to the electronic structure, which in turn is critically dependent on the atom positions. In order to cope with the requirements needed for theoretical calculations, materials characterization techniques should be able to measure atom positions down to a very high precision (order of 0.01 Å). Classical X-ray and neutron techniques will fail to accomplish this task, because of the inherent aperiodic character of nanostructures. Scanning probe techniques, on the other hand, cannot provide information below the surface. Only fast electrons interact sufficiently strongly with matter to provide local information to atomic scale. Therefore, in the near future, high resolution electron microscopy (HREM) is probably the most appropriate technique for very precise measurement of atom positions [2].

The purpose of this paper is to describe how HREM will ever be able to measure atom positions with a precision of the order of 0.01 Å, even with a microscope with only 1 Å resolution. In our view, to achieve this goal, the microscope should be turned from a qualitative imaging instrument into a quantitative measuring instrument.

In section II, it is explained how the resolution of the electron microscope has been increased over the years in order to improve visual interpretation of the images. One is inclined to think that a precision of 0.01 Å requires a resolution of 0.01 Å, which is far beyond the present possibilities. However, precision and resolution are different things. Precision is attained by quantitative interpretation of the images. In section III, it is described how this could be done. Section IV shows how the precision can further be increased if quantitative structure determination is accompanied by quantitative statistical experimental design.

II. BEYOND ONE ÅNGSTRÖM RESOLUTION

Since the very beginning of HREM, it is standard practice to characterize the structure of materials by visual interpretation of the images or by visual comparison with computer simulations. The resolution of the conventional microscope was not high enough to visualize the individual atoms that constitute matter. This would require a resolution beyond 1 Å. In an electron microscope, one has to distinguish two resolution measures, the so-called point resolution and information limit [3]. Point resolution is defined as the smallest detail that can be interpreted directly from the experimental images in terms of the structure, provided that the object is very thin. The point resolution is nowadays about equal to 2 Å. Classical X-ray and neutron techniques will fail to accomplish this task, because of the inherent aperiodic character of nanostructures. Scanning probe techniques, on the other hand, cannot provide information below the surface. Only fast electrons interact sufficiently strongly with matter to provide local information to atomic scale. Therefore, in the near future, high resolution electron microscopy (HREM) is probably the most appropriate technique for very precise measurement of atom positions [2].

The purpose of this paper is to describe how HREM will ever be able to measure atom positions with a precision of the order of 0.01 Å, even with a microscope with only 1 Å resolution. In our view, to achieve this goal, the microscope should be turned from a qualitative imaging instrument into a quantitative measuring instrument.

In section II, it is explained how the resolution of the electron microscope has been increased over the years in order to improve visual interpretation of the images. One is inclined to think that a precision of 0.01 Å requires a resolution of 0.01 Å, which is far beyond the present possibilities. However, precision and resolution are different things. Precision is attained by quantitative interpretation of the images. In section III, it is described how this could be done. Section IV shows how the precision can further be increased if quantitative structure determination is accompanied by quantitative statistical experimental design.
The former three possibilities aim at a reduction of the point resolution. The latter two restore the details that are present between the point resolution and the information limit. In high-voltage electron microscopy, the accelerating voltage of the electron microscope is increased up to 1 MeV and beyond [5]. The disadvantage here is the increase of the radiation damage of the specimen and the high cost. The correction of the spherical aberration [6] in the microscope allows to correct for a lens defect that causes a point object to be imaged as a disk of finite size. By using multipole lenses, spherical aberration can be cancelled out. One of the main difficulties is the complicated procedure for the alignment of the large number of electrostatic and magnetic optical elements. HAADF STEM [7] reduces the point resolution, although at the expense of an enormous loss of imaging electrons. Both last methods, focal-series reconstruction as well as off-axis holography, retrieve the exit-plane wave function, that is the complex electron wave function at the exit-plane of the specimen. Ideally, this exit-plane wave function (or, shortly, exit wave) is free from any imaging artefacts, which means that the visual interpretability of the reconstruction is enhanced considerably for thin objects when compared to the original experimental images. In this case, the resolution will be determined by the information limit. The focal-series reconstruction method [8, 9] reconstructs the exit-plane wavefunction from a series of images collected at different focus values. Off-axis holography [10] is based on the original idea of Gabor [11], where the exit-plane wave function is retrieved from the interference between the object wave and a reference wave.

With the increasing need for precise determination of the atomic arrangements of non-periodic structures in materials design, and control of nanostructures, a resolution of 1 Å resolution is still insufficient. Quantitative structure determination at the subångström level is needed. Materials science of the future will require precision instead of resolution. The atom positions will have to be measured within picometer precision.

III. THE ROAD TO QUANTITATIVE STRUCTURE DETERMINATION

The goal of quantitative high resolution electron microscopy is to extract the structural information (parameters) of an object. Images should therefore be considered as data planes from which this structural information has to be retrieved quantitatively. Ultimately, one wishes to determine structural parameters, such as the atom positions, as precisely as possible.

Quantitative extraction of structural information requires a model for the electron-object interaction, the transfer in the microscope, and the image detection, i.e., the ingredients needed to perform a computer simulation of the experiment. Nowadays, the electron-object interactions and the image formation process are sufficiently well understood to allow the successful derivation of such models and several commercial software packages for HREM image simulations are available [12], [13]. The object model, describing the interaction with the electrons, consists of the assembly of electrostatic potentials of the constituting atoms. Since for each atom type the electrostatic potential is known, the object parameters (required as an input to an image simulation program) reduce to atom numbers, coordinates, Debye Waller factors, object thickness and orientation. These parameters are unknown beforehand and have to be estimated from the experiment. The formation of an image in the microscope is also characterized by a set of parameters, such as defocus, spherical aberration constant of the objective lens, voltage, pixel size of the camera, etc. These parameters, which are needed as input to an image simulation program as well, are either known beforehand with sufficient accuracy, or not, in which case they have to be estimated from the experiment. The quantification of the structural parameters is now done by fitting the model to the experimentally obtained images. The goodness of fit between model and experiment, i.e., between the image obtained from the image simulation program using a set of model parameters as input and the image obtained experimentally, is then evaluated using a criterion such as likelihood function. For each set of model parameters, one can thus calculate the value of the criterion of goodness of fit. The parameters for which this value is optimum then correspond to the estimate of the model structure.

In a sense, one is looking for the optimum value of a criterion on a parameter space whose dimension is equal to the number of parameters to be estimated. This search for the global optimum of the criterion of goodness of fit is an iterative numerical procedure. Generally, the dimension of the parameter space is very high. This means that a very large variety in images can be obtained by a variation in model parameter values. Consequently, it is quite possible that the optimization procedure ends up at a local optimum instead of at the global optimum of the criterion of goodness of fit, so that the wrong structure is suggested. To solve this dimensionality problem, that is, to find a pathway to the global optimum in the parameter space, a good starting structure is required. Finding such a starting structure is not trivial, since due to two encoding processes, details in the images do not necessarily correspond to features in the atomic
structure. The first encoding process is the dynamic scattering of the electrons on their way through the object. The second encoding process is the formation of the image in the electron microscope. As a consequence, the structural information of the object may be strongly delocalised. Additionally, the images are always disturbed by noise, which further complicates direct interpretation. However, it has been shown that good starting structures can be found by using direct methods, which are described in section II. For example, focal-series reconstruction methods in a sense invert, or equivalently, decode the image process. Consequently, the thus obtained exit wave is much more related to the object structure, providing a directly interpretable resolution close to the information limit (ca. 1 Å), which just surpasses the limit beyond which individual atom columns can be discriminated [14], [15]. Focal-series reconstruction methods thus yield an approximate structure that can be used as a starting point in a final refinement by fitting with the original images.

The starting structure obtained via a direct method may still be insufficiently close to the global optimum of the criterion of goodness of fit to guarantee convergence. In order to find a better starting structure, one also has to decode the first encoding process mentioned above, that is, the dynamic scattering of the electrons on their way through the object. Decoding the dynamic scattering is of a higher complexity than decoding the imaging process. It requires advanced knowledge of quantum mechanics. An all embracing solution has not yet been found. Different routes to achieve this goal are currently investigated. For instance, a simple and invertible albeit approximate channeling theory has been proposed, applicable if the crystal object is perfectly oriented along a zone axis [16]. Then it can be shown that the electrons are trapped in the positive potential of the columns. Each atom column then, in a sense, acts as a channel for the electrons. If the distance between the columns is not too small, a one-to-one correspondence between the exit wave and the object structure is established [16], [17]. Channeling theory allows us to derive an analytical expression for the exit wave, parametric in the projected positions of the constituent atom columns of the structure, their energy and the object thickness. The energy of the column is a parameter related to the projected “weight” of the column, which is a function of the atom number, the distance between successive atoms along a column and the Debye-Waller factor [18]. Obviously, one may expect to obtain only projected information and ambiguity about the types and distance of atoms along a column can only be removed by combining information from different zone axis orientations [19]. Fitting the analytical expression for the exit wave to the reconstructed exit wave, with respect to the model structure parameters, provides the experimenter with an approximate structure that can then be used as an improved starting point for a final refinement by fitting with the original images.

IV. PRECISION AND EXPERIMENTAL DESIGN

As described above, with the resolution becoming sufficient to reveal individual atom columns, a structure can be characterized completely by the (projected) positions of its constituent atom columns, their energy, the object thickness and orientation. Quantitative structure determination by means of HREM imaging now becomes a statistical parameter estimation problem, the image pixel values being the observations from which the parameters of interest have to be estimated. If the model is correct, the precision with which these parameters can be estimated is only limited by the presence of noise. Measurement of the unknown parameters may result in higher precisions if it is accompanied by quantitative statistical experimental design. The limiting factor in the design is the amount of electrons that interact with the material during the experiment, because too long exposure times may cause material damage and/or specimen drift.

Due to the inevitable presence of noise, the observations will always fluctuate randomly and may therefore be modeled as stochastic variables. By definition, a stochastic variable is characterized by its probability density function (distribution), while a set of stochastic variables has a joint probability density function. The expectation (mean value) of each observation may be described by a model, parametric in the quantities to be estimated. Given such a model, use of the concept of Fisher information allows the experimenter to determine the highest precision, that is, the lowest possible variance, with which a parameter can be estimated unbiasedly from a set of observations assumed to obey a certain distribution [20]-[22]. Thus, it is possible to derive an expression for the lower bound on the variance with which the position of an atom column can be estimated in a quantitative HREM experiment. This lower bound, which is called the Cramér Rao Lower Bound (CRLB), is a function of both object parameters, such as the weight of the columns and the specimen thickness, and microscope parameters, such as defocus, spherical aberration constant, energy spread of the electrons, as well as the electron dose used to form the image. This means that the CRLB varies with the experimental conditions, of which at least some are adjustable (within certain practical limits). The optimal experimental design of an HREM experiment is then given by the experimental conditions that correspond to the lowest CRLB. It can be found by minimizing the CRLB with respect to the adjustable experimental parameters. Notice, that the optimal experimental design may be different for different objects under investigation. In our view, the way in which HREM, aiming at the highest precision, should be approached is represented in Fig. 1.
The principle of precision based optimization of the experimental design can be illustrated as follows. Suppose that one is interested in a measurement of the distance \( d \) between two atoms (or two atom columns in projection) from an HREM experiment. The distances, or equivalently the atom positions, are strongly connected to the materials properties, and are therefore, the most interesting structural parameters. Let us call \( \rho \) the width of the image of the atom and \( N \) the total number of detected electrons available to visualize this atom. The width \( \rho \) in fact represents the “resolution” in the sense of Lord Rayleigh. It depends on both object and microscope parameters. It can be shown [23]-[25] that the highest attainable precision, in terms of the standard deviation \( \sigma \), with which the distance can be measured is given by the following rule of thumb:

\[
\sigma = \frac{2\rho^2}{\sqrt{Nd}} \quad \text{if} \quad d \leq \sqrt{2}\rho, \\
\sigma = \frac{\sqrt{2}\rho}{\sqrt{N}} \quad \text{if} \quad d \geq \sqrt{2}\rho.
\]

(1)

From (1), it is clear that if one wants to optimize the design of an HREM experiment or to decide between different methods, or to develop new techniques, one has to keep in mind that it is not only the resolution that counts but also the detected electron dose. For distances smaller than \( \sqrt{2}\rho \), the precision increases proportionally with the distance. For distances larger than \( \sqrt{2}\rho \), the precision becomes independent of the distance. In that case, the variance \( \sigma^2 \) is equal to two times the variance of the estimated position of an isolated atom.

The rule of thumb, given in (1) can further be illustrated with an example. Annular dark-field scanning transmission electron microscopy (ADF STEM) is often used to provide direct structural information. In ADF STEM, an electron probe is formed by demagnifying an electron source with a set of electromagnetic lenses. The resulting probe scans over the specimen in a two-dimensional raster. A detector is placed in the back focal plane beyond the specimen, where a diffraction pattern is formed. The detector is provided with a central hole, so that only the scattered electrons are detected. The part of the diffraction pattern that is collected, is integrated and displayed as a function of the position of the probe. More details can be found in [7]. It is important to realize that the resolution in ADF STEM is inversely proportional to the probe size. From this point of view, it is clear that new instrumental developments aim at smaller probes. However, a smaller probe also requires a larger hole in the detector, so as to guarantee that the non-scattered electrons are eliminated from detection. As a side effect, the number of detected electrons decreases drastically. Mathematically, this means that \( \rho \) as well as \( N \) in (1) decrease. In this respect, it is not clear whether smaller probes result in more precise measurement of the distance between two atom columns. This issue is investigated in more detail in [25]-[28]. From the optimization procedure, as explained above, it follows that the optimal probe size depends strongly on the material under study. Consider, for example, a Si-crystal in its [110]-direction, which has the classic “dumbbell” pairs of Si-columns in its projected structure. In Fig. 2, a simulated image of this structure is shown under the optimal experimental conditions, in terms of attainable precision. The white blobs represent two Si-columns, which are not visually separable. The distance \( d \) between these two atom columns is equal to 1.36 Å. The attainable precision, in terms of the standard deviation \( \sigma \), is equal to 0.0012 Å (for a given total recording time).

In Fig. 3, an ADF STEM image of the same structure is imaged, but this time with a much smaller probe, and thus a larger hole in the detector. The total recording time is the same as in Fig. 2. It is clear, that due to the higher resolution, the Si-columns are visually separable but, due to a huge loss of electrons in the detector, the signal-to-noise ratio is very low. As a result, the standard deviation \( \sigma \) is equal to 0.05 Å, which is about 40 times larger than what can be expected from an experiment with optimal design. This example shows that intuition may be misleading for optimization of the design of quantitative experiments. Although Fig. 3 would be preferred based on resolution criteria related to direct visual interpretation, Fig. 2 would result in much higher precisions. It should be mentioned, however, that an experiment as presented in Fig. 3 can provide a good starting structure for a final refinement. The vital importance of good starting structures has been pointed out in the previous section.

Another interesting question is if correction of the spherical aberration constant in a transmission electron microscope will yield a higher precision. This subject is discussed in [29]. It turned out that, indeed, this correction improves the point resolution, but it also shifts the whole passband (in the transfer from the electron wave field at the exit plane to the image plane) to higher spatial frequencies. Hence, for light atoms, which provide only limited scattering at high angles, the optimal spherical aberration constant need not be very low.
Fig. 2. Simulated ADF STEM image of a Si-crystal in its [110]-direction at the optimal conditions. The white blobs represent two Si-columns, which are not visually separable.

Fig. 3. Simulated ADF STEM image of a Si-crystal in its [110]-direction at experimental conditions providing visually separable atom columns.

Yet another recent progress is the development of a monochromator [30], which improves the information limit. However, it is shown in [31], that in most cases a monochromator doesn’t pay off in terms of precision, since the improvement of resolution is accompanied by a reduced loss of electrons.

V. CONCLUSIONS

Recent progress and future prospects of quantitative structure determination from HREM images have been discussed in a parameter estimation framework. In the authors’ view, the electron microscope should be turned from a qualitative imaging instrument into a quantitative measuring instrument in order to cope with future developments in materials science. Furthermore, it has been shown that quantitative structure determination could greatly benefit from an accompanied quantitative statistical experimental design.

ACKNOWLEDGMENT

The research of Dr. A.J. den Dekker has been made possible by a fellowship of the Royal Netherlands Academy of Arts and Sciences.

REFERENCES


