Inverse elastic scattering with adaptive regularization and meshes

D Aubry, A-L Hamon and G Puel

Laboratoire MSS-Mat. UMR CNRS 8579, Ecole Centrale Paris
Grande Voie des Vignes, 92290 Châtenay-Malabry, France

E-mail: denis.aubry@ecp.fr

Abstract. A simplified modelling of the inverse problem is proposed in the case of a scattering experiment such as in transmission electron microscopy (TEM). The two major tools are the use of the finite element method (FEM) for the wave equation, derived from a Helmholtz type equation, and a Tikhonov regularization for the misfit function. The equations for the wave function, the adjoint state and the optimal potential are established. The direct problem is illustrated with a 2D case of iron in the [011] zone axis using a Yukawa potential. The obtaining of the regularization by a coarse mesh rather than by the usual penalisation parameter is discussed.

1. Introduction
The solutions of an inverse problem are well known for being unstable and not unique. In [1], we studied the influence of Tikhonov’s regularization in the specific case of the subterranean cavities detection using experimental data obtained by an interferometric satellitar device. The corresponding inverse problem consisted in identifying the Young’s modulus field using the measured displacements on the surface of the studied ground. We showed that in this specific example, the use of a classical Tikhonov’s regularization term was not sufficient to identify a whole spatial field of properties.

Therefore we introduced a strategy using finite element methods (FEM) with a classical adaptive meshing in order to regularize the inverse problem. The first step consisted in using two distinct meshes: one associated to the discretization of the sought spatial field (e.g. Young’s modulus), which could be adaptively refined, and one associated to the resolution of the usual mechanical problems (e.g. direct and adjoint state problems). In [1], we showed that the choice of a coarse mesh to discretize the sought spatial field made possible a good regularization of the problem. In a second step, the introduction of local error estimators could drive the refinement of the coarse mesh associated to the sought parameters [2].

Wave scattering due to a heterogeneous medium is a generic problem analysed by such books as [3], [4] or [5] for review. In the present case, we study the case of electron-matter elastic interaction such as what occurs in transmission electron microscopy (TEM) experiments [6]. In the direct problem, the potential created by the sample is known at the electron scale, and we derive, at the atomic scale, the crystalline structure, the defects nature and organisation… which are confronted to experimental results. In the inverse problem, the sample potential is derived from experimental scattering information. Thus the inverse problem consists in determining in which case the potential can be estimated uniquely or, if not, what the ambiguous data are due to.
In the case of a central potential which corresponds to the present study, the scattering information that we get is the scattering amplitude at a given energy. In that case, the existence and the uniqueness of the inverse problem have been already analysed [3].

1. Physical problem set-up

The specific case of inverse electron scattering is developed using the same general iterative strategy as described in [1] and [2]. Before interacting with the studied specimen $\Omega_s$, an incident electron is considered as a plane wave function $\psi_i$ associated to a wave vector $k_i$ which propagates in an empty space $\Omega_e$ (Figure 1). When removing the time harmonic factor:

$$\psi_i (x) = \hat{\psi}_i \exp \left( i \mathbf{k}_i \cdot \mathbf{x} \right)$$  \hspace{1cm} (1)

The sample domain $\Omega_s$ is characterized by a potential $V_s$ due to the electrons and nuclei of the sample. $V_s$ vanishes rapidly outside $\Omega_s$. The goal of the study is to determine that potential through electron scattering measurements.

In the whole free space, the electron wave function $\psi_i$ satisfies the Schrödinger equation:

$$-\frac{1}{2} \Delta \psi_i = E_i \psi_i \hspace{1cm} (2)$$

in atomic units (au: distances in Bohr, energies in Hartree). The plane wave solution gives:

$$E_i = ||\mathbf{k}_i||^2/2$$

The whole system to be considered in a scattering experiment – the incident electron and the particles of the sample – has to satisfy the full Schrödinger equation [4]. $\psi_{es}$ is the total wave function of the system of total energy $E$, and the potential is decomposed as a sum of potentials taking into account the interaction of the sample with itself, $V_s$, and with the incoming electron, $V_{es}$:

$$-\frac{1}{2} \left( \Delta_x + \Delta_{x_s} \right) \psi_{es} (x, x_s) + \left[ V_{es} (x, x_s) + V_s (x_s) \right] \psi_{es} (x, x_s) = \text{E}_{es} \psi_{es} (x, x_s)$$  \hspace{1cm} (3)

where $x$ stands for the description of the incident electron position and $x_s$ for the description of all other particles. To uncouple the terms due to the incoming electron from the terms due to the sample, as a quick justification for the following equation, we can approximate the total wave function $\psi_{es}$ by a product of the sample eigenfunction $\psi_s$ with that of the electron $\psi_e$. A complete justification can be found in [4], so that the following equation has to be satisfied by $\psi_e$:

$$-\frac{1}{2} \Delta_x \psi_e (x) + V(x) \psi_e (x) = \left( E - E_s \right) \psi_e (x) \hspace{1cm} (4)$$

where $E_s$ stands for the sample energy with:

$$V(x) = \int \left[ V_{es} (x, x_s) \right] \psi_s (x_s) |^2 \, dx_s.$$

In the particular situation of inelastic scattering when $||\mathbf{k}||$ is not constant, the response of the whole system needs to be computed. In the elastic scattering case, a common approximation used in the case of TEM is that the velocity of the incident electron and consequently $E_i$ are very large and given (kinetic energy around 200 keV, i.e., 7350 Hartree), so that $(E - E_s)$ is finally approximated by the energy of the incident electron $E_i$:

$$-\frac{1}{2} \Delta_x \psi_e + V \psi_e = E_i \psi_e \hspace{1cm} (6)$$

The wave function $\psi$ is decomposed into the incident one $\psi_i$ and the diffused one $\psi_d$:

$$\psi = \psi_i + \psi_d \hspace{1cm} (7)$$
and the unknown diffused wave function $\psi_d$ satisfies the following equation obtained from equations (6) and (7):

$$\frac{1}{2} \Delta_x \psi_d + (E_i - V) \psi_d = V \psi_i$$  (8)

Yet, considering that $V$ is small compared to $E_i$ (e.g. $V_{\text{max}} \approx 0.26$ Hartree for carbon), that equation is very often simplified into the following Helmholtz equation:

$$\frac{1}{2} \Delta_x \psi_d + E_i \psi_d = V \psi_i$$  (9)

In the conventional TEM environment, the outgoing wave function ($\psi_i + \psi_d$) is magnified by a series of magnetic lenses allowing an analysis both in imaging or in diffraction mode (i.e. Fourier transform mode). The intensity of the final wave function $\psi_f$ is then collected on devices such as a screen, an imaging plate or a CCD camera. $\psi_f$ results from the convolution of the exit wave function ($\psi_i + \psi_d$) considered on the exit plane of the sample with the transfer function of the TEM, which mostly takes into account the effect of the defocus and of the objective lens spherical aberration. In the present case, we do not take into account the transfer function of the microscope, but we compute ($\psi_i + \psi_d$) on any virtual plane outside the sample.

The diffused wave function $\psi_d$ is a function of the crystal potential $V$. In the identification problem considered here the inverse relationship is looked for.

$$\sum_{\Sigma_m}$$

2. Direct problem computation

The direct problem is a Helmholtz type equation with $E_i$ very large. The order of magnitude of the wavelength is $\lambda_i = 0.05$ au whereas the interatomic distance is around 5 au. The total thickness of the sample is about $5.10^3$ au.

2.1. Classical image computation

Classically, the Green's function $g$ of the Helmholtz problem is considered:

$$g(r) = \exp(i k, r)/r$$

After multiplying the Helmholtz equation (9) by $g$ and integrating it by parts, the following expression for $\psi_d$ is obtained:

$$\psi_d(x) = \int_{\Omega_e} V(y) \psi_i(y) g(x - y) dy = \left[ (V \psi_i) * g \right](x)$$  (10)

The back and forth reflections between the atoms are not taken into account.
2.2. FEM direct problem using a paraxial approximation

The multislice method ([7], [8], [9] and [10]), which is classically used, divides the sample into slices with respect to the thickness. The size of these slices is chosen more or less empirically. But the FEM computation of the Helmholtz equation is not submitted to the different assumptions implied by the multislice method. All the waves are taken into account and the potential \( V \) of the sample may not be periodic e.g. when defects are present. However the main shortcomings come from the high frequency feature of the problem which implies very large meshes (around \( 10^9 \) degrees of freedom - DOFs - for one single crystalline cell, assuming 10 DOFs per wavelength). To circumvent this difficulty, a paraxial approximation of the solution is first adopted with the following new unknown \( \psi_d \):

\[
\psi_d (x) = \tilde{\psi}_d (x) \exp \left( i k_i x \right)
\]  

(11)

where \( k_i \) is the incident wave vector, along the microscope axis in the case of parallel illumination. Even if, formally, this paraxial approximation does not imply that the diffracted wave vectors are close to the incident one \( k_i \), it fits well experimental TEM conditions, where all the beams diffracted with an angle superior to approximately 30 mrad are not to be taken into account, since they are truncated by the microscope's transfer function.

Plugging this expression into the Helmholtz equation (9) gives the following equation:

\[
\frac{1}{2} \Delta \tilde{\psi}_d + i k_i \nabla \tilde{\psi}_d = V \tilde{\psi}_i
\]

(12)

since

\[
E_i = \| k_i \|^2 / 2.
\]

Equation (12) is thus linear with respect to \( k_i \), a key-point with respect to the numerical computation. One should note that contrarily to paraxial approximations as commonly used, the second order derivative in the wave vector direction is not neglected.

The boundary conditions are of the evanescent type [11] which means that,

on \( \Sigma_i \):

\[
\frac{\partial \tilde{\psi}_d}{\partial n} = i \| k_i \| \tilde{\psi}_d
\]

(13)

where \( n \) is the unit outer vector, normal to the boundary of the domain.

This boundary condition has to be transferred to \( \tilde{\psi}_d \) which implies that,

on \( \Sigma_i \):

\[
\frac{\partial \tilde{\psi}_d}{\partial n} = i \left( \| k_i \| - k_i \cdot n \right) \tilde{\psi}_d
\]

(14)

3. Inverse problem formulation

The inverse problem consists in finding the potential field \( V \) such that the electron wave modulus \( |\psi_m|^2 = |\tilde{\psi}_i + \tilde{\psi}_d|^2 \) matches the measured intensity \( |\psi_m|^2 \) on \( \Sigma_m \) at best. The following misfit function \( J(V) \) is chosen ([12] and [13]):

\[
J(V) = \frac{1}{4} \int_{\Sigma_m} \left( |\tilde{\psi}_i + \tilde{\psi}_d|^2 - |\psi_m|^2 \right)^2 + \frac{\alpha}{2} \int_{\Omega} \left( V - V_0 \right)^2
\]

(15)

where \( \alpha \) is the regularization parameter, and \( V_0 \) is an a priori chosen potential close to the one sought, for instance the potential of the crystal with no defects. In that case, the inverse problem consists in finding a defect in the crystal through the modification induced in the potential.

Other norms may be more adapted to this problem like the superficial gradient of the measured intensity on \( \Sigma_m \).
3.1. Adjoint state and optimality equations

The resolution is made through the computation of an adjoint state problem, similar to the direct
electron scattering formulation: using both latter problems, we can easily express the derivative of the
cost function $J(V)$. The adjoint state $z$ is defined by [14]:

inside $\Omega$:

$$1/2 \Delta z - i \mathbf{k}_i \nabla z = 0 \quad (16.1)$$

on $\Sigma_m$:

$$\frac{1}{2} \left[ \frac{\partial z}{\partial n} \right] = \left( \left| \psi_i + \psi_d \right|^2 - \left| \psi_m \right|^2 \right) \left( \psi_i + \psi_d \right) \quad (16.2)$$

thus, on $\Sigma_\infty$:

$$\frac{\partial z}{\partial n} = i \left( \left| \mathbf{k}_i \right|^2 + \mathbf{k}_i \mathbf{n} \right) z \quad (16.3)$$

The derivative of the cost function is then given by:

$$\delta J = \int_{\Sigma_m} \left[ \left( \left| \psi_i + \psi_d \right|^2 - \left| \psi_m \right|^2 \right) \left( \psi_i + \psi_d \right) \delta \psi_d^* \right] + \alpha \int_{\Omega} \left( \left( V - V_0 \right) \delta V \right) \quad (17)$$

But we easily get from the adjoint state definition (16.1) and (16.2), the sensitivity equation (17)
and the vanishing conditions of the wave function (14) and of the adjoint state (16.3):

$$\int_{\Omega} \left( - \frac{1}{2} \nabla \delta \psi_d \cdot \nabla \left[ \psi_i \right] + i \mathbf{k}_i \nabla \delta \psi_d \cdot \nabla \left[ z \right] \right) + \int_{\Sigma_m} \left( \frac{1}{2} \nabla \delta \psi_d \cdot \mathbf{n} \right) \left( \nabla \left[ z \right] \right) = \int_{\Omega} \left( \left[ \psi_i \right] \delta \psi_d \right) \quad (18)$$

$$\int_{\Omega} \left( \frac{1}{2} \nabla \delta z \cdot \nabla \left[ \psi_i \right] - i \mathbf{k}_i \nabla \delta \psi_d \cdot \nabla \left[ \psi_i \right] \right) = \int_{\Sigma_m} \left[ \left( \left| \psi_i \right|^2 - \left| \psi_m \right|^2 \right) \left( \psi_i + \psi_d \right) \delta \psi_d \right] \quad (19)$$

so that:

$$\int_{\Sigma_m} \left[ \left( \left| \psi_i \right|^2 - \left| \psi_m \right|^2 \right) \left( \psi_i + \psi_d \right) \delta \psi_d \right] = \int_{\Omega} \left( \left[ \psi_i \right] \delta \psi_d \right) \quad (20)$$

and the optimality of the sought potential $V$ gives:

$$\psi_i \left| z \right| = \alpha \left( V - V_0 \right) \quad (21)$$

As a summary we have now the coupled system of equations (12)-(14)-(16)-(21) with unknowns
$\{ \psi_i, z, V \}$. This system may be solved by an iterative approach, typically either a fixed point or
Newton schemes. Furthermore, the unknown potential $V$ may be eliminated thanks to the optimality
equation (21) and plugged into the direct and adjoint equations.

3.2. Multimesh numerical scheme

The regularization of the inverse problem results from the use of a classical Tikhonov’s regularization
through the potential $V_0$. It is however important here to remark that independant discretization may be
employed for the direct problem with unknown $\psi_d$ and for the parameter $V$ estimation. The
variational formulation of the direct problem reads for the discretized unknowns $\{ \psi_{dh}, z_h, V_{dh} \}$
associated to the continuous unknowns $\{ \psi_i, z, V \}$, for any test functions $\{ w_{dh}, y_h, W_h \}$:

$$\forall w_h, \quad \int_{\Omega} \left( - \frac{1}{2} \nabla \psi_{dh} \cdot \nabla w_h + i \mathbf{k}_i \nabla \psi_{dh} \cdot \nabla \psi_i \right) + \int_{\Sigma_m} \left( \frac{1}{2} \left( \mathbf{n} \cdot \nabla \psi_{dh} \right) \right) w_h = \int_{\Omega} \left( V_{dh} \psi_i w_h \right) \quad (22)$$

$$\forall W_h, \quad \int_{\Omega} \left[ \psi_i W_h + \alpha \left( V_{dh} - V_0 \right) \right] W_h = 0 \quad (23)$$

with a similar expression for the adjoint state $z_h$. Although not considered here, different
discretizations could also be considered for the direct and adjoint state because the former is more sensitive to the potential description whereas the latter to the experimental error smoothness.

The proposed scheme here accounts for the fact that no potential derivative appears in the formulation which implies that a coarser mesh can be employed for this quantity. Bangerth [15] has remarked that this procedure acts as a regularization itself because the discretized potential during the adaptive meshing belongs to a restricted space which constrains its possible range.

For a given number of DOFs, it is more efficient to use larger meshes coupled with polynoms of higher degree. That avoids well known pollution artefacts.

4. Numerical examples

![Figure 2](image1.png)

**Figure 2.** Computation of $|\Psi_i + \Psi_d|^2$ using the paraxial approximation in the 2D direct problem for $\alpha$-iron in a [011] zone axis. Unit length are in atomic unit (au: $a_0 \approx 5.10^{-11}$ m). The sample potential $V$ is superimposed on the figure in order to visualize the position of the sample nuclei. The sample thickness is equal to $2\text{ nm}$, still below experimental values. The number of adaption levels is equal to 2, the number of DOFs to approximately $6.10^5$.

![Figure 3](image2.png)

**Figure 3.** FFT of the exit wave function $(\Psi_i + \Psi_d)$ taken on a virtual plane $\Sigma_m$ located close to the inferior plane of the sample. Diffracted beams clearly appear according to Bragg's law.

For the computation of the direct scattering problem, a 2D sample of pure $\alpha$-iron is considered (Figure 2). Parameters are chosen close to a 200 keV TEM, using $||k|| = 130$ rad.au$^{-1}$. The sample is oriented in the [011] zone axis, and the potential $V$ induced by the sample is simulated using a Yukawa potential [16]. As equation (12) is linear in the incident plane wave amplitude $\Psi_i$ and in the potential coefficient, both values are taken equal to 1. For better convergence, an absorption coefficient equal to
5.10⁻³ is chosen.

If we wish to take into account the microscope's lens aberrations, it is the Fourier transform of the wave function on a given plane which is of physical interest. Thus the fast Fourier transform (FFT) of the exit wave function is also shown (Figure 3). The diffraction of the electron wave function by the crystal appears clearly, obeying Bragg's law.

5. Conclusion

FEM are used here to solve a scattering inverse problem which consists in estimating a potential from the observed diffused wave intensity collected in the far field. For the direct problem involving large wave vectors such as in TEM, the paraxial approximation allows to treat larger domains than the Helmholtz equation. The equation is linear in \( k \), the incident wave vector, but contrarily to the Helmholtz equation, it is not selfadjoint anymore. The adjoint state satisfies a paraxial equation of the same type but with retropropagation in order to propagate the misfit information from the observation locus to the potential support.

As far as the inverse problem is concerned, the misfit function uses the Tikhonov regularization. In the present case, only a simple least square norm has been chosen but more refined ones, involving square derivatives for instance, might be used as well. Further work will obviously include such investigations together with the three-dimensional applications.

References

[1] Puel G, Tie B and Aubry D 2007 Adaptive approach and choice of norms related to the regularization of inverse problems *International Conference on Adaptive Modeling and Simulation ADMOS* 194-197

[2] Bangerth W, Joshi A and Sevick-Muraca E M 2005 Adaptive finite element methods for increased resolution in fluorescence optical tomography *Progr. Biomed. Optics Imag.* 6 318–329

[3] Newton R G 2002 *Scattering theory of waves and particles* (Mineola:Dover) Part III pp. 141-670

[4] Wang Z L 1995 *Elastic and inelastic scattering in electron diffraction and imaging* (New York:Plenum) Part I pp. 1-126

[5] Colton D and Kress R 1992 *Inverse acoustic and electromagnetic scattering theory* Applied Mathematical Sciences 93 (Berlin:Springer Verlag)

[6] Smith A R and Eyring L 1982 Calculation, display and comparison of electron microscope images modeled and observed *Ultramicroscopy* 8 65-78

[7] Williams D B and Carter C B 1996 *Transmission electron microscopy* (New York:Plenum)

[8] Kirkland E J 1998 *Advanced Computing in Electron Microscopy* (New York:Plenum)

[9] Ishizuka K and Uyeda N 1977 A new theoretical and practical approach to the multislice method *Acta Cryst.* A 33 740-749

[10] Stadelmann P A 1987 *EMS - a software package for electron diffraction analysis and HREM image simulation in materials science, Ultramicroscopy* 21 131-146

[11] Popov A.V. 2006 Computation of paraxial wave fields using transparent boundary conditions *Computational Mathematics and Mathematical Physics* 46:9 1595-1600

[12] Beilina L and Johnson C 2005 A posteriori error estimation in computational inverse scattering *J. Mathematical models and methods in applied sciences* 15:1 23–37

[13] Beilina L and Clason C 2006 An adaptive hybrid fem/fdm method for an inverse scattering problem in scanning acoustic microscopy *J. SIAM Sci.Comp.* 28:1 382-402

[14] Nintcheu Fata S and Guzina B B 2007 Elastic scatterer reconstruction via the adjoint sampling method *SIAM J. Appl. Math.* 67:5 1330–1352

[15] Bangerth W and Rannacher R 2003 *Adaptive finite element methods for differential equations* (Basel:Birkhäuser Verlag) Chapter 8 pp. 101-111

[16] Ashcroft N W and Mermin N D 2002 *Physique des solides* (Les Ulis:EDP Sciences) Chapter 2 p.43 and Chapter 17 pp. 406-407