Acta Crystallographica Section A
Foundations of Crystallography
ISSN 0108-7673

Received 28 May 2001
Accepted 6 August 2001

A new phasing method based on the principle of minimum charge

Pavel Kalugin

Laboratoire de Physique des Solides, Université Paris-Sud, 91405 Orsay, France. Correspondence e-mail: kalugin@lps.u-psud.fr

A new method for phase determination in X-ray crystallography is proposed. The method is based on the so-called 'minimum-charge' principle, recently suggested by Elser [*Acta Cryst.* (1999), A**55**, 489–499]. The electron-density function ρ is sought in the form $\rho(\mathbf{x}) = |\psi(\mathbf{x})|^2$, where ψ is an n-component real function. The norm $\int |\psi(\mathbf{x})|^2 \, d\mathbf{x}$ is minimized under the constraint imposed by the measured data on the amplitudes of Fourier harmonics of ρ. Compared with the straightforward implementation of the 'minimum-charge' scheme, the method attenuates the Gibbs phenomenon and is also capable of extrapolation of the diffraction data beyond the set of measured amplitudes. The method is applicable to quasicrystals under the condition that the number of components n of the function ψ is bigger than the dimensionality of the 'atomic surface'. It has been successfully tested on synthetic data for a Fibonacci chain and octagonal tiling. In the latter case, the reconstructed density map shows the shape of the atomic surface, despite relatively low resolution data.

1. Introduction

The vast majority of the existing direct methods of X-ray structure determination approach the phase problem as a problem of constrained minimization. The quantity to minimize plays the role of the likelihood functional, optimization of which is subject to constraints imposed by the known structure-factor amplitudes. The choice of this quantity is a matter of trade-off between three requirements. First, this functional should approximately represent the common notion about the likelihood of a given density function. Second, it should be effectively computable and, finally, it should allow for efficient minimization. Usually none of these requirements are fully satisfied. Consider, for instance, the traditional direct methods based on the conditional probability distribution of the structure invariants (Harker & Kasper, 1948; Karle & Hauptman, 1950). The derivation of the expressions for the conditional probability takes as a starting point the assumption that unconditional probability of the density distribution (its *Bayesian prior*) is a translationally invariant measure on the ensemble of N δ-like atoms in a unit cell (Cochran, 1955). Clearly this approach misses the physical constraint on a minimal distance between the atoms. Although this prior measure allows one to express the conditional probability of the phase invariants in a closed form (Cochran, 1952, 1955; Hauptman & Karle, 1953), the resulting formulae are effectively computable only for invariants of small order. As a result, instead of the true conditional probability function, the practical algorithms use some sort of *ad hoc* approximation, usually based on a combination of the triplets and quartets [see, for example, the '*Shake-n-Bake* algorithm' (Weeks *et al.*, 1993)]. However, even these simplified functionals are not easy to optimize. The existing phase refinement methods are based on iterative procedures and can easily become stuck in a local minimum of a functional.

Recently, Elser (1999) suggested the so-called *principle of minimum charge*. According to this principle, the correct set of phases should minimize the average charge density (the unknown Fourier component with $\mathbf{k} = 0$, which has to be added to the density to make it non-negative). The rationale behind this principle is that the functions satisfying it tend to have shallow highly degenerated minima and spiky maxima, which is what one would expect of an atomistic density function. Compared with the solid statistical foundations of the conventional direct methods, the principle of minimum charge may look arbitrary. However, this comparison is not fair because, on the way from the first principles to the practical implementation of the conventional methods, many *ad hoc* assumptions are made. In contrast, the principle of minimum charge is almost readily applicable in the phase refinement algorithms. Additional advantage of this principle is that it could be used without modification for the determination of the structure of quasicrystals and non-commensurate crystals. This is especially important because the conventional methods fail when applied to these structures. Indeed, a naive attempt to approximate quasicrystals by crystals with a very large number N of atoms in the unit cell gives rise to meaningless results. As N tends to infinity, the normalized structure factors diverge as $E_{\mathbf{H}} = O(N^{1/2})$ and the triplet amplitudes as $A_{\mathbf{HK}} = O(N)$, leading to the conclusion that the phases of all triplets

$\Phi_{\mathbf{HKL}} = 0$. This divergence reflects the fact that the correlations between atomic positions in quasicrystals are anomalously strong with respect to what one would expect from the Bayesian prior described above.

A straightforward implementation of the principle of minimum charge [*e.g.* that proposed by Elser (1999)] implies solving a minimax problem. Indeed, the set of phases $\Phi_{\mathbf{K}}$ satisfies the principle if the deepest minimum of the 'density with zero average',

$$\rho'(\mathbf{r}) = 2\sum_{\mathbf{K}} |F_{\mathbf{K}}| \cos(\mathbf{K}\cdot\mathbf{r} - \Phi_{\mathbf{K}}), \tag{1}$$

over $\mathbf{r}$ has the maximal possible value over all possible sets $\Phi_{\mathbf{K}}$. Robust algorithms for finding a global saddle point are much more difficult to design than those for finding a global minimum or maximum. This is because the methods which are usually applied to prevent the algorithm from becoming stuck with a local minimum (*e.g.* simulated annealing or multiple runs) do not work in the case of local saddle points. In this article we propose a version of the principle of minimum charge which can be formulated as a problem of global minimization, and not as a minimax problem, allowing one to circumvent the shortcomings of a straightforward approach.

The naive implementation of the principle of minimum charge suffers from another drawback, which is due to the so-called Gibbs phenomenon. In order to understand this problem, consider the case when the true values of all phases $\Phi_{\mathbf{K}}$ are known. Suppose also that the correction for the atomic form factors is included in the values of the structure factors $F_{\mathbf{K}}$ in (1), or in other words that the atoms are point-like. Owing to the Gibbs phenomenon, any finite sum of the form (1) presents negative 'bumps' around the positions of atoms. The depth of these artifacts may be as big as 20% of the peak height. This picture is clearly different from multiply degenerate shallow minima that one would expect of a function satisfying the principle of minimum charge. In other words, the correct set of phases gives a density function (1) which is suboptimal from the point of view of this principle. Any further optimization of it may only introduce errors in the values of phases. The proposed method significantly attenuates the importance of the Gibbs phenomenon, although it does not remove it completely.

2. Method

2.1. Representation of the density function

The cornerstone of the new method is the representation of the atomic density function. The traditional way of reconstructing this function consists of using finite Fourier sums (1) with possible application of weights aimed to attenuate the Gibbs phenomenon. A different approach is used here. The density function is approximated by a square of a finite Fourier sum. More precisely, the density function $\rho(\mathbf{x})$ is modelled as

$$\rho(\mathbf{x}) = |\psi(\mathbf{x})|^2, \tag{2}$$

where $\psi(\mathbf{x})$ is a multi-component real function. Each component ψ_α of ψ has a finite Fourier spectrum,

$$\psi_\alpha(\mathbf{x}) = \sum_{\mathbf{K}\in\Lambda} \tilde{\psi}_{\alpha,\mathbf{K}} \exp(i\mathbf{K}\cdot\mathbf{x}), \tag{3}$$

where Λ is a finite subset of reciprocal-lattice vectors. The complex coefficients $\tilde{\psi}_{\alpha,\mathbf{K}}$ satisfy the condition

$$\tilde{\psi}_{\alpha,-\mathbf{K}} = \tilde{\psi}^*_{\alpha,\mathbf{K}}. \tag{4}$$

Before going any further, let us discuss the rationale behind this representation. Recall that the density function satisfying the principle of minimum charge should possess the following properties:

(i) The amplitudes of its Fourier components for the set of measured Bragg reflections should correspond to the measured data $F_{\mathbf{K}}$.

(ii) The global minimum of the function should be highly degenerate (see Fig. 1).

The traditional way to represent the density function by a finite Fourier sum automatically guarantees that the condition (i) is satisfied. However, the condition (ii) cannot be expressed in a closed form as an explicit constraint on the phases $\Phi_{\mathbf{K}}$. On the other hand, degenerate global minima occur naturally in the functions of type (2) when $\psi(\mathbf{x})$ has multiple zeros. At the same time, the condition (i) can be explicitly imposed on the function (2). This constraint can be formulated as a system of equations of the fourth order on the variables $\tilde{\psi}_{\alpha,-\mathbf{K}}$ (see Appendix *A*).

Mention should be made of the criterion for choice of the support Λ for the Fourier spectrum of the function ψ in (3). Clearly, if the set Λ is too small, the condition (i) may be impossible to satisfy. This happens, for example, if there are wave vectors $\mathbf{L}$ belonging to a set Ξ of measured data which cannot be represented as $\mathbf{L} = \mathbf{H} - \mathbf{K}$, where $\mathbf{H}, \mathbf{K} \in \Lambda$. On the other hand, choosing the set Λ equal to the set Ξ guarantees that any constraint on the amplitudes of Fourier components of density $\tilde{\rho}_{\mathbf{K}}$,

$$\tilde{\rho}_{\mathbf{K}} = \int \rho(\mathbf{x}) \exp(-i\mathbf{K}\cdot\mathbf{x})\, d\mathbf{x}, \tag{5}$$

Figure 1
Phase refinement based on the principle of minimum charge leads to the reconstructed density maps with highly degenerate shallow minima.

where $\mathbf{K} \in \Xi$, can be satisfied in the representation (2). Indeed, by assuming the following constraint on the amplitudes of $\tilde{\rho}_{\mathbf{K}}$,

$$|\tilde{\rho}_{\mathbf{K}}| = a_{\mathbf{K}},$$

then setting

$$\tilde{\psi}_{\alpha,\mathbf{K}} = \begin{cases} \varepsilon^{-1} & \text{if } \mathbf{K} = 0, \quad \alpha = 0 \\ (\varepsilon/2)a_{\mathbf{K}} & \text{if } \mathbf{K} \neq 0, \quad \alpha = 0 \\ 0 & \text{if } \alpha \neq 0 \end{cases}$$

gives the correct Fourier components of ρ in the limit $\varepsilon \rightarrow +0$,

$$|\tilde{\rho}_{\mathbf{K}}| = a_{\mathbf{K}} + O(\varepsilon).$$

Note that in this case the support of the spectrum of ρ is roughly twice as large in the reciprocal space as the set Λ. In other words, the representation (2) allows for extrapolation of the structure factors. One can further extrapolate the experimental data by choosing an even larger set Λ. The limit of possible improvement of the resolution in this way is set by the occurrence of spurious peaks in the density map in the case of over-extrapolation owing to the increased sensitivity of the result to the errors in the measured data.

2.2. Optimization criterion and constraints

The representation (2) of the density is a non-negative function. Hence the average value of ρ can be used as a figure of merit for the optimization; in other words, the principle of minimum charge regains its original meaning. It should be emphasized that, instead of the problem of finding a global saddle point (global maximum of global minima), we deal here with an easier problem of constrained minimization. The role of parameters is played by the Fourier components $\tilde{\psi}_{\alpha,\mathbf{K}}$ of the function ψ. The average density is expressed through $\tilde{\psi}_{\alpha,\mathbf{K}}$ as

$$\langle \rho \rangle = \sum_{\mathbf{K} \in \Lambda} \sum_{\alpha} |\tilde{\psi}_{\alpha,\mathbf{K}}|^2. \qquad (6)$$

The parameters $\tilde{\psi}_{\alpha,\mathbf{K}}$ can be considered as a real vector in the Mn-dimensional space, where M is the number of points in the set Λ and n is the number of components of ψ. The principle of minimal charge thus boils down to the problem of constrained minimization of the norm of this vector.

Consider now the constraints imposed by the measured data on the values $\tilde{\psi}_{\alpha,\mathbf{K}}$. Let us assume that the amplitudes of the structure factors $|F_{\mathbf{K}}|$ are measured for the wave vectors $\mathbf{K}$ belonging to a subset Ξ of the reciprocal lattice. Then the seemingly obvious constraint on the density function $\rho(\mathbf{x})$ from (2) consists of the requirement that

$$|\tilde{\rho}_{\mathbf{K}}| = |F_{\mathbf{K}}| \qquad \text{for any } \mathbf{K} \in \Xi.$$

This condition, expressed in terms of the Fourier components of ψ from (3), takes the form

$$\left| \sum_{\alpha} \sum_{\mathbf{H},\mathbf{K}-\mathbf{H} \in \Lambda} \tilde{\psi}_{\alpha,\mathbf{H}} \, \tilde{\psi}_{\alpha,\mathbf{K}-\mathbf{H}} \right| = |F_{\mathbf{K}}| \qquad \text{for any } \mathbf{K} \in \Xi. \quad (7)$$

This constraint, however, does not take into account the effect of the Gibbs phenomenon. The importance of this effect is clear from Fig. 2. This plot shows the function $\rho(\mathbf{x})$ obtained by the minimization of the average density (6) under the constraints (7). The structure factors used here correspond to a one-dimensional crystal with one atom per unit cell, *i.e.* $|F_{\mathbf{K}}| = 1$ for all $\mathbf{K} \in \Xi$. Clearly the phasing on Fig. 2 is incorrect. The reason for this is that the correct set of phases would correspond to a function with deep negative bumps around the 'atom' because of truncation of data in the reciprocal space. The optimization routine attempts to flatten these bumps out and align all minima on the same level, giving rise to an incorrect solution.

The standard approach to reduction of the Gibbs phenomenon consists of using the windowing in the Fourier domain. As applied to (7), the windowing consists of multiplying the amplitudes of the structure factors by weights,

$$|\tilde{\rho}_{\mathbf{K}}| = \left| \sum_{\alpha} \sum_{\mathbf{H},\mathbf{K}-\mathbf{H} \in \Lambda} \tilde{\psi}_{\alpha,\mathbf{H}} \, \tilde{\psi}_{\alpha,\mathbf{K}-\mathbf{H}} \right| = w_{\mathbf{K}} |F_{\mathbf{K}}|. \qquad (8)$$

The choice of the coefficients $w_{\mathbf{K}}$ is a matter of trade-off between softening the truncation of data on the reciprocal lattice and preserving the spatial resolution of the density map. In the field of digital signal processing, there exist many windowing formulae designed to reduce the magnitude of the Gibbs phenomenon, *e.g.* Welch or Hann windows (Rabiner & Gold, 1975; Bloomfield, 1976). Nonetheless, instead of using these formulae, we shall derive the expressions for $w_{\mathbf{K}}$ which are optimized from the point of view of the discussed method. Ideally, the weights $w_{\mathbf{K}}$ should guarantee that the global minimum of the average density (6) with the constraints (8) corresponds to the density map ρ with the correct phases. The weights should be a function of the sets Ξ and Λ only, and should not depend on the values of the structure factors $F_{\mathbf{K}}$. Unfortunately, such ideal weights do not exist. This is clear at

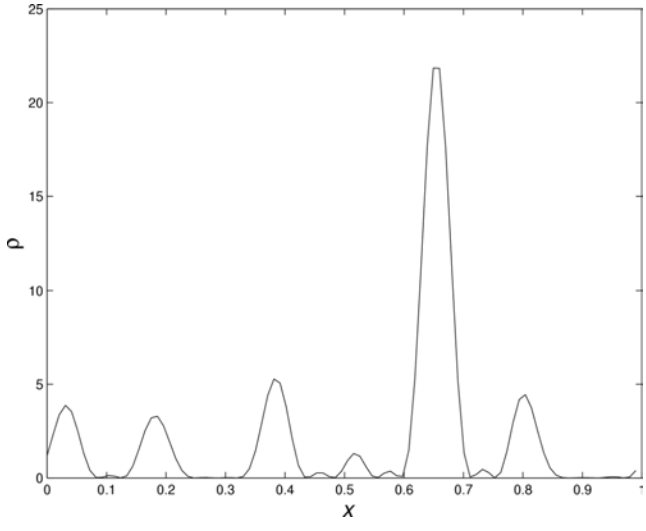

Figure 2
Artifacts occurring owing to the Gibbs phenomenon. The original structure is a one-dimensional crystal with one atom per unit cell. The function ψ has two components ($n = 2$). The set Ξ includes the first ten vectors of the reciprocal lattice, and the set Λ includes the same vectors, their opposites and the zero vector.

least from the fact that perfect phasing is impossible when the set of measured data is too small compared with the number of atoms in the unit cell. Nevertheless, as long as the resolution of the measured data is sufficiently high, there exists an almost optimal set of weights. More specifically, the weights $w_{\mathbf{K}}$ chosen in such a way as to guarantee the correct phasing for the structure with one atom per unit cell should produce reasonable results for other structures as well, provided that the peaks in the density map do not overlap.

Let us start the construction of the optimal weights with the case of a one-dimensional crystal with one atom in a unit cell. First of all, note that the constraints in (7) with $|F_{\mathbf{K}}| = 1$ are equivalent to those of (8) with $|F_{\mathbf{K}}| = 1/w_{\mathbf{K}}$. In other words, one can interpret the density map of Fig. 2 as a result of correct phasing of *a priori* unknown structure-factor amplitudes $|F_{\mathbf{K}}| = 1/w_{\mathbf{K}}$. Clearly, the number of peaks in this density function depends on the number of constraints in (7). One naturally expects that the minimization of the global charge with only one constraint would lead to some very simple structure. Indeed, as shown below, if the set Ξ in (7) contains only one wavevector $\mathbf{K}_0$, which is equal to the elementary period of the reciprocal lattice, the minimization of the average density gives rise to a structure with one atom in a unit cell. Let $\tilde{\psi}_{\alpha,\mathbf{K}} = \tilde{\eta}_{\alpha,\mathbf{K}}$ be the solution of (7) with the single constraint $|F_{\mathbf{K}_0}| = 1$. One can use these values to construct the following set of weights:

$$w_{\mathbf{K}} = \frac{\left|\sum_{\alpha}\sum_{\mathbf{H},\mathbf{K}-\mathbf{H}\in\Lambda}\tilde{\eta}_{\alpha,\mathbf{H}}\,\tilde{\eta}_{\alpha,\mathbf{K}-\mathbf{H}}\right|}{\sum_{\mathbf{K}\in\Lambda}\sum_{\alpha}|\tilde{\eta}_{\alpha,\mathbf{K}}|^2}. \tag{9}$$

This set of weights guarantees that the density map obtained with a single constraint will satisfy the principle of minimum charge for any number of constraints. Indeed, $\tilde{\psi}_{\alpha,\mathbf{K}} = w_{\mathbf{K}_0}^{1/2}\,\tilde{\eta}_{\alpha,\mathbf{K}}$ obeys (8) with the weights (9) and $|F_{\mathbf{K}}| = 1$ for all $\mathbf{K} \in \Xi$. In other words, the weights $w_{\mathbf{K}}$ are optimal for the set Λ.

Let us actually compute the optimal weights for the case when the set Λ includes the nodes $-N, -N+1, \ldots, N$ of the one-dimensional reciprocal lattice. We suppose that $\psi(\mathbf{x})$ has only one component and use simplified notations by writing $\tilde{\psi}_k$ or F_k instead of $\tilde{\psi}_{\alpha,\mathbf{K}}$ or $F_{\mathbf{K}}$ (here k is the number of the node). In the case when only the amplitude of the structure factor F_1 is known, there is only one constraint on the values $\{\tilde{\psi}_i\}$,

$$\left|\sum_{k=-N+1}^{N}\tilde{\psi}_k\tilde{\psi}^*_{k-1}\right| = w_1|F_1|. \tag{10}$$

Introduction of a Lagrange multiplier means that the constrained minimization of the average density can be replaced by finding an unconstrained extremum of the following quantity,

$$\sum_{k=-N}^{N}\tilde{\psi}_k\tilde{\psi}^*_k + \lambda\left|\sum_{k=-N+1}^{N}\tilde{\psi}_k\tilde{\psi}^*_{k-1}\right|. \tag{11}$$

Note that the second sum in (11) can always be made real and positive by an appropriate shift δx of the coordinate system and multiplication of $\tilde{\psi}_k$ by $\exp(ik\delta x)$. Then, if expression (11) has an extremum at a given $\{\tilde{\psi}_i\}$, the same is true for the expression

$$\sum_{k=-N}^{N}\tilde{\psi}_k\tilde{\psi}^*_k + (\lambda/2)\sum_{k=-N+1}^{N}(\tilde{\psi}_k\tilde{\psi}^*_{k-1} + \tilde{\psi}^*_k\tilde{\psi}_{k-1}). \tag{12}$$

Differentiation with respect to $\tilde{\psi}$ gives the equation for the point of extremum,

$$\tilde{\psi}_k + (\lambda/2)(\tilde{\psi}_{k-1} + \tilde{\psi}_{k+1}) = 0. \tag{13}$$

(Here we assume $\tilde{\psi}_{-N-1} = \tilde{\psi}_{N+1} = 0$.) In other words, the solution of the principle of minimum charge with the constraint (10) should be an eigenstate of a lattice Laplace operator with zero boundary conditions beyond the set Λ. Straightforward calculations show that the function

$$\tilde{\psi}_k = C\cos[\pi k/2(N+1)] \tag{14}$$

gives the smallest value of the average density. The density (2) is a sum of peaks centred at the lattice nodes

$$\rho(x) = \sum_{n\in Z} f\,[(2N+2)(x-n)], \tag{15}$$

where the shape of an individual peak is given by the following formula,

$$f(u) = C[\cos(\pi u)/(1-4u^2)]^2. \tag{16}$$

As one might expect, in the limit $N \to \infty$ the density map tends to a sum of δ functions corresponding to a crystal with one atom per unit cell. The formula (9) gives the values for the optimal weights,

$$w_k = \frac{\sin[(2N+3-|k|)\,\theta] + (2N+3-|k|)\sin(\theta)\cos(k\theta)}{\sin[(2N+3)\,\theta] + (2N+3)\sin(\theta)\cos(k\theta)}, \tag{17}$$

where $\theta = \pi/2(N+1)$. The results of testing of these weights with the structure factors corresponding to one atom in the unit cell are shown in Fig. 3.

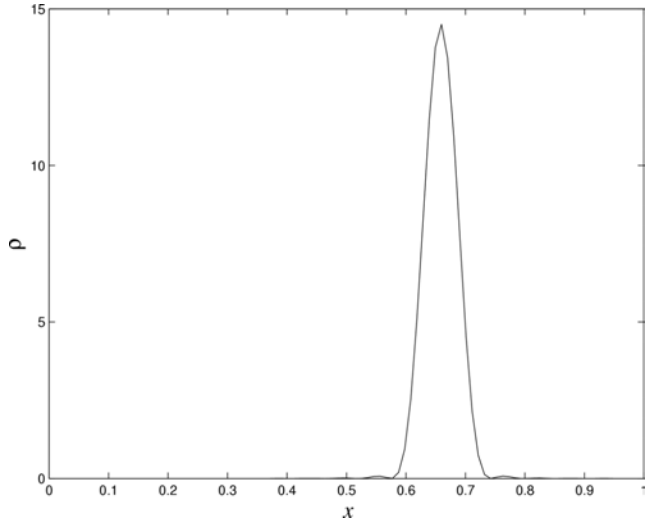

Figure 3
Reconstructed density map for the same test case as that of Fig. 2, but with constraints (8). The values of the weights are given by (17).

Consider now the ways to generalize the above construction of the optimal weights to more than one dimension. The constraints (8) may uniquely determine an atomistic structure only if the wavevectors from the set Ξ span the entire reciprocal lattice (Gabardo, 1999). The minimal number of constraints is thus equal to the dimensionality d of the crystal. Solving the problem of minimum charge with the known structure factors for the set of wavevectors $\Xi = \{\mathbf{K}_i\}$, $i = 1, \ldots, d$, yields to a formula similar to (13),

$$\tilde{\psi}_{\alpha,\mathbf{K}} + \sum_{i=1}^{d} (\lambda_i/2)(\tilde{\psi}_{\alpha,\mathbf{K}-\mathbf{K}_i} + \tilde{\psi}_{\alpha,\mathbf{K}+\mathbf{K}_i}) = 0, \qquad (18)$$

where $\tilde{\psi}_{\alpha,\mathbf{K}} = 0$ for $\mathbf{K} \notin \Lambda$. The weights can be obtained from the solution $\tilde{\psi}_{\alpha,\mathbf{K}}$ of this equation using (9).

Contrary to the one-dimensional case, (18) does not yield to a unique set of optimal weights. First of all, the choice of the wave vectors $\mathbf{K}_i$ is ambiguous. Clearly, d vectors $\mathbf{K}_i$ spanning the entire reciprocal lattice form its basis. However, there are many ways to choose a basis of a lattice in more than one dimension. The other problem is related to the choice of Lagrange multipliers λ_i. In (13), the solution ψ_k is always one of the eigenstates of the Laplace operator and the role of the parameter λ is restricted to mere selection of the eigenstate. In contrast, $\tilde{\psi}_{\alpha,\mathbf{K}}$ in (18) may vary continuously with λ_i. In order to resolve the ambiguity, one needs to recall that the density corresponding to the solution of the finite difference equation (18) describes the structure with one atom at the origin of the unit cell. In other words, the function $\psi(\mathbf{x})$ should have a sharp peak at the origin, *i.e.* its Fourier components $\tilde{\psi}_{\alpha,\mathbf{K}}$ should vary slowly across the domain Λ. The requirement that (18) should admit such a slowly varying solution determines the choice of parameters $\mathbf{K}_i$ and λ_i.

2.3. Quasicrystals and incommensurate structures

Consider now the application of the discussed method to the deterministic quasicrystals and incommensurate structures in general. These structures can be conveniently described using the so-called superspace or 'cut-and-project' method (de Wolff *et al.*, 1981; Kalugin *et al.*, 1985; Duneau & Katz, 1985; Elser, 1986). According to this approach, the density function of quasicrystals can be obtained as a cut through a periodic function in a space of higher dimensionality. The Fourier spectrum of the structure is obtained as a projection of the spectrum of the periodic function on the cut direction, and thus consists of a discrete sum of δ functions. If the direction of the cut is incommensurate with the periodicity, each δ peak in the Fourier spectrum of the quasicrystal corresponds to a node of the reciprocal lattice of the periodic function. By this means, the phase problem for quasicrystals can be reformulated in a conventional way as a problem of phasing for a periodic function in a space of higher dimensionality.

For the case of point-like atoms, the density function of a quasicrystal in real space is a discrete sum of δ functions. The corresponding periodic density function consists of δ-like distributions on submanifolds, commonly referred to as 'atomic surfaces' (Janssen, 1986; Bak, 1986). This brings up again the question of optimality of the weight factors $w_{\mathbf{K}}$ from (8). Indeed, the weights obtained following the method described in the previous section are appropriate for distributions of narrow non-overlapping peaks. Atomic surfaces clearly do not fall into this category. As a result, one could expect that using the weight factors optimized for point-like atoms might give rise to incorrect results in the case of atomic surfaces. Nevertheless, lacking any better alternative, we applied these weights for quasicrystals as well. Despite concerns, the results of the numerical tests described below show no significant distortion of the density map.

Another problem emerges when the atomic surfaces are discontinuous, which is the case for all known deterministic structure models of real materials. The point is that, owing to the finite resolution in the reciprocal space, the boundary of the atomic surface is inevitably smoothed out. As a result, when the physical space cuts through an atomic surface near its boundary, the amplitude of the corresponding peaks in the density map is reduced. Similarly, when the cut misses the atomic surface by a short distance, a 'phantom' peak appears in the density map. In reality, both of the above phenomena occur simultaneously, because of the so-called 'closedness' property of the atomic surfaces (Katz, 1989). This results in the occurrence of double peaks of reduced intensity, which are often separated by a distance much smaller than the typical interatomic spacing. A careful analysis shows that such double peaks in places group together to form more complex clusters (Kalugin & Katz, 1993). It should be emphasized that all these artifacts are not specific to the phasing method and will exist in any reconstructed quasicrystalline density map. They should not be confused with the phason disorder in real quasicrystals (Lyonnard *et al.*, 1996; de Araújo *et al.*, 1996; de Boissieu *et al.*, 1995), which manifests itself in the occurrence of true partially occupied sites at the same positions.

2.4. Choice of the number of components of ψ

Up to this point, the number of components of the function $\psi(\mathbf{x})$ has been of no importance to us. However, because of the iterative nature of the optimization algorithm, this parameter plays an important role in the case of quasicrystalline structures. The problems that arise in the case when the number of components of ψ is too small are clear from Fig. 4. This figure depicts a one-dimensional atomic surface in the two-dimensional space crossed by a line of zeros of the one-component function ψ. The resulting hole can only be removed by pushing it towards the boundary of the atomic surface. It may occur, however, that pushing the hole inwards decreases the average density. In this case, the algorithm will converge to a local minimum.

The above problem could be avoided if the function ψ had two components. In this case, the zeros of ψ are points. Should such a point fall onto the atomic surface, it could be easily pushed away by a small perturbation of ψ. The same applies to the case of more than two dimensions. In this case, the number of components of ψ should be bigger than the dimensionality of the atomic surface.

Table 1
One-dimensional test structure.

'Charge'	x
2.0	0.0
1.5	0.6
1.2	0.25
1.0	0.43
1.0	0.8

3. Results

The results presented in this section were obtained numerically using the algorithm of constrained minimization described in Appendix *A*. The algorithm was tested on the sets of synthetic data including a one-dimensional crystal and a quasicrystal, as well as a two-dimensional quasicrystalline structure. No rotational symmetry was assumed in any of the cases. When the structure actually possessed additional symmetry, it was recovered as a result of optimization.

Owing to its iterative nature, the algorithm of Appendix *A* does not guarantee convergence to the global minimum of the average density. It is worth noting, however, that for all tested structures the correct phasing corresponds to the deepest of the found minima.

3.1. One-dimensional crystals

The test structure includes five point-like atoms. Their 'charges' and fractional coordinates x are given in Table 1. As the structure does not possess central symmetry, the reconstructed density can correspond to any of the enantiomorphs with equal probability.

The density map shown in Fig. 5 has been obtained with the first 12 structure-factor amplitudes. The function ψ had two components. The support of the Fourier spectrum of each component Λ from formula (3) included the wavevectors from the interval $[-12, 12]$. The positions of five peaks in Fig. 5 correspond to the coordinates of atoms in Table 1 up to a global translation. The amplitudes of the peaks are also qualitatively recovered. One can remark, however, that the two smallest peaks in Fig. 5 are noticeably different, although they correspond to identical atoms. The value of the average density is equal to 6.636, which is about 1% smaller than the total charge of 6.7 in Table 1. The iterative optimization, which converged to the global minimum solution in 14 of 20 trials when starting with random normally distributed $\tilde{\psi}_{\alpha,\mathbf{K}}$, was 70%, giving a 70% success rate.

The test structure from Table 1 was also used to benchmark the extrapolating capabilities of the algorithm. The input data was restricted to the first nine structure-factor amplitudes. Note that the structure recovery with any smaller data set would be impossible because a one-dimensional structure of N atoms is described up to a shift by $2N - 1$ real parameters (N weights of atoms plus $N - 1$ relative coordinates). By contrast, the set Λ from (3) included all wavevectors from the interval $[-50, 50]$. By this means, the algorithm extrapolated the 'experimental' scattering data to a region of the reciprocal space roughly 11 times larger than that where they were initially defined. The reconstructed density map is shown in Fig. 6. The positions of atoms are perfectly accurate up to a global translation and inversion. Note also that despite a smaller resolution of the input data (which is only about 0.65 of the smallest interatomic distance), the amplitudes of the peaks are restored with a higher accuracy than in Fig. 5. The value of the average density is equal to 6.690, which is also much closer to the total charge of 6.7 in Table 1. This can be explained by the overlapping of wider peaks in Fig. 5, which makes the correct solution suboptimal. As a result, the algorithm produces slightly distorted solutions, which have smaller values of the average density. Narrower peaks in Fig. 6 make this phenomenon much less visible.

One could anticipate a lower success rate in the latter example because of the bigger number of parameters in $\tilde{\psi}_{\alpha,\mathbf{K}}$ to fit. Contrary to the expectations, convergence to the global minimum was obtained in 19 trials out of 20.

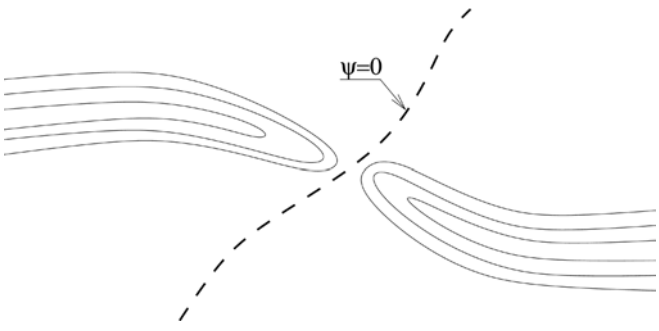

Figure 4
An unremovable hole in the atomic surface on the reconstructed density map. Such holes occur when a line of zeros of the one-component function ψ crosses the atomic surface.

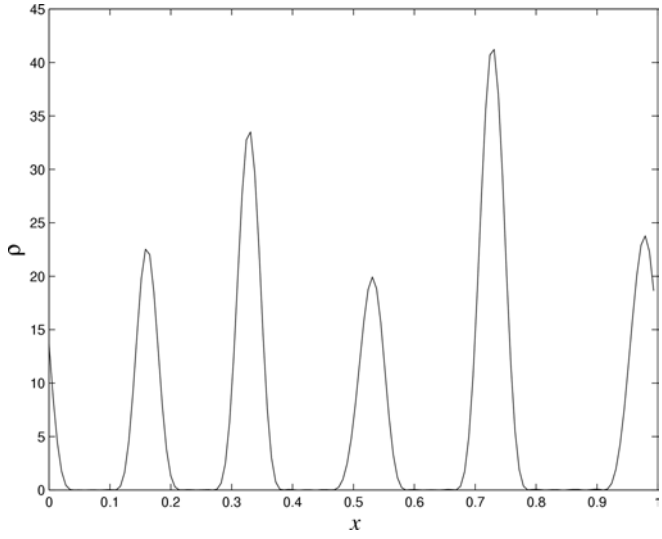

Figure 5
The restored density map corresponding to the structure from Table 1. The input data includes first 12 structure-factor amplitudes. The function ψ has two components and is defined with the same resolution.

3.2. One-dimensional Fibonacci chain

A Fibonacci chain is probably the best known example of a one-dimensional quasicrystal. The chain consists of identical atoms arranged in a straight line. The distance between neighbouring atoms may take two values, usually referred to as 'short' and 'long' segments of the chain. The ratio of lengths of short and long segments is $\tau = (5^{1/2} - 1)/2$. The short and long segments alternate following a quasiperiodic Fibonacci sequence, which is usually defined recursively (Senechal, 1995). The same structure can be obtained following the conventional 'cut-and-project' method (see Fig. 7).

The algorithm has been tested on a set of 40 independent reflections with Miller indices $h^2 + k^2 \leq 25$. Although the structure in Fig. 7 possesses central symmetry, this symmetry was not imposed as an additional constraint on ψ. The reconstructed density map is shown in Fig. 8. The optimization has yielded the correct solution with 100% success rate for 50 trials with random initial conditions.

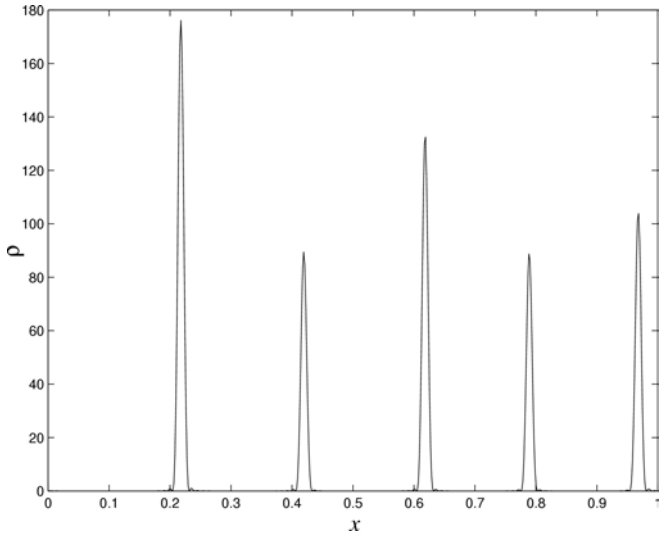

Figure 6
Density map for the test structure of Table 1 extrapolated from the low-resolution data.

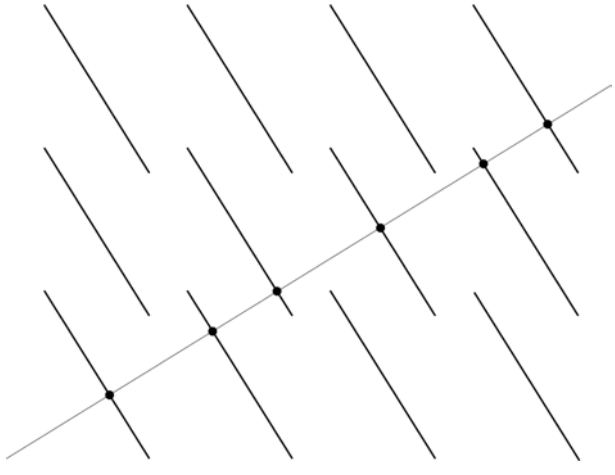

Figure 7
The 'cut-and-project' method of construction of the Fibonacci chain. The one-dimensional physical space cuts the two-dimensional lattice of periodically arranged segments ('atomic surfaces'). The atoms, shown as filled circles, are located at intersections of 'atomic surfaces' with the physical space.

3.3. Two-dimensional octagonal Ammann–Beenker tiling

Ammann–Beenker octagonal tiling (Ammann *et al.*, 1992; Beenker, 1982) refers to quasiperiodic covering of a plane by squares and 45° rhombi. Structure models associated with this tiling can be constructed by decorating each tile with atoms. The simplest decoration consists of placing atoms at vertices of the squares and rhombi. The resulting structure can be obtained by the 'cut-and-project' technique from a periodic density function in four-dimensional space. The corresponding atomic surfaces are two-dimensional perfect octagons, one per unit cell.

The input data included the reflections with $h^2 + k^2 + l^2 + m^2 \leq 5$ (where h, k, l and m stand for four-dimensional Miller indices). There are 68 pairs of opposite nodes of the reciprocal lattice satisfying the above inequality. The 8*mm* symmetry of the diffraction pattern makes only 14 of them independent. However, as mentioned above, the point symmetry was not taken into account. As a result, all 68 pairs of reflections were considered as independent. The function ψ had three components. Note that, as no central symmetry is imposed, the coefficients $\tilde{\psi}_{\alpha,\mathbf{K}}$ from (3) are complex numbers.

The optimal set of coefficients $\tilde{\psi}_{\alpha,\mathbf{K}}$ returned by the algorithm should be converted to the density map in the physical space. As the direction of the physical space is incommensu-

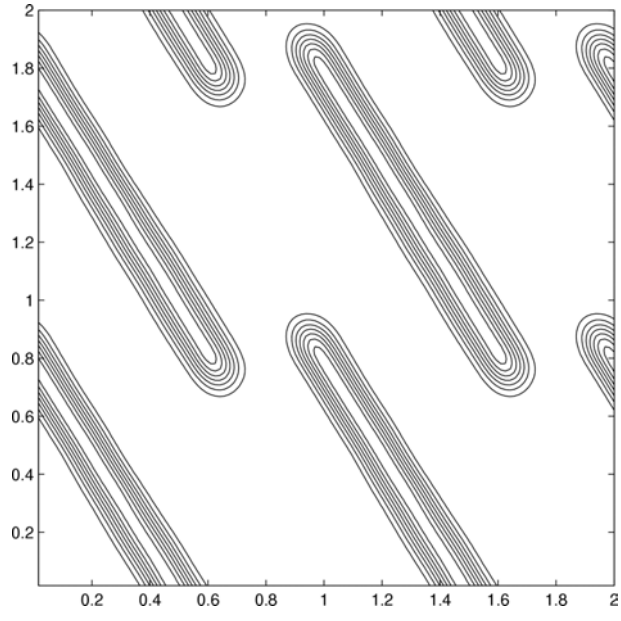

Figure 8
Reconstruction of the atomic surfaces of the one-dimensional Fibonacci chain. The input data include 40 independent reflections with $h^2 + k^2 \leq 25$. The function ψ has two components. The area of the density map shown includes four unit cells. The axes represent fractional coordinates.

rate with the four-dimensional periodic lattice, this conversion implies the Fourier transform of unevenly spaced data. This precludes using standard FFT algorithms, which significantly slows down displaying the results. A way around this problem consists of tilting the physical space slightly to make it commensurate with the lattice. This is equivalent to replacing the quasicrystal by a close approximant (Duneau & Audier, 1994).

Fig. 9 depicts the reconstructed density map for an approximant to Ammann–Beenker tiling. The approximant is obtained by replacing the vectors $\mathbf{u}_1 = (2^{1/2}/2, 1/2, 0, -1/2)$ and $\mathbf{u}_2 = (0, 1/2, 2^{1/2}/2, 1/2)$ spanning the physical space by the rational vectors

$$\begin{aligned} \mathbf{u}'_1 &= (5/7, 1/2, 0, -1/2), \\ \mathbf{u}'_2 &= (0, 1/2, 5/7, 1/2). \end{aligned} \tag{19}$$

Note that the peaks on the density map substantially overlap because of low resolution of the input data. This is also confirmed by the fact that the algorithm yields an average density as low as 0.718 of its correct value, suggesting significant over-optimization. Nevertheless, most of the atomic positions are correctly resolved, as can be seen from a comparison with the vertices of the ideal approximant tiling shown in Fig. 10. The double peaks which are expected owing to the smoothened edges of the atomic surface (see the explanation above) are not fully resolved. They manifest themselves as elongated features on the density map. The optimization algorithm has converged to the correct solution with a 100% success rate of ten trials with random initial conditions.

One could also consider cutting the four-dimensional density function along different directions. For instance, a cut in the direction of the atomic surface would reveal its shape. Under practical conditions, this shape will be distorted because the cut may not pass exactly through the centre of the four-dimensional peak corresponding to the atomic surface. Such a cut is shown in Fig. 11. Despite the low resolution of the input data, one can clearly see a faceted octagonal shape.

4. Summary and discussion

We have described a new method of structure determination based on the minimum-charge principle. The method possesses extrapolating capabilities and can restore atomic positions from low-resolution data (about 65% of the interatomic distance). It is also applicable to quasicrystals and incommensurate structures. As the current implementation of

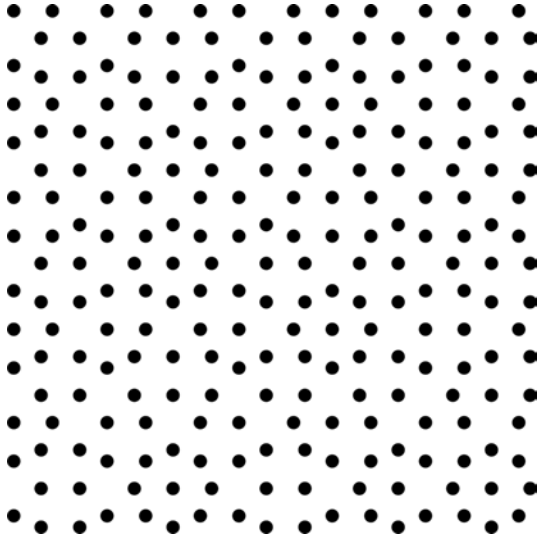

Figure 10
The vertices of the approximant to Amman–Beenker tiling with the physical space spanned by the vectors (19).

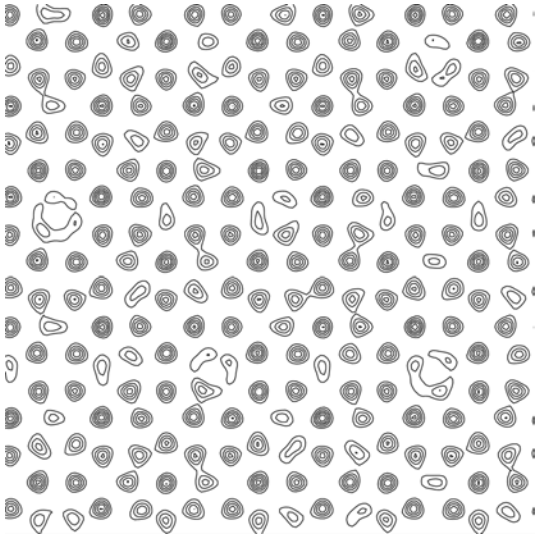

Figure 9
Density map corresponding to the vertices of an approximant to Ammann–Beenker tiling, reconstructed from the reflections with $h^2 + k^2 + l^2 + m^2 \leq 5$.

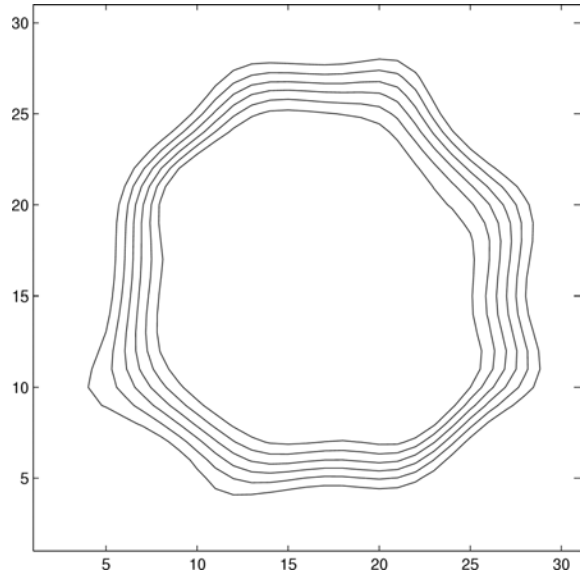

Figure 11
The cut through the reconstructed density map of Ammann–Beenker tiling in the direction parallel to the atomic surface. The scale is arbitrary.

the algorithm was intended as a demonstration of principle, no special efforts were taken to make it more efficient. Making the new method practical for crystals with non-trivial point symmetry and for real quasicrystals requires further research.

The failure to take advantage of the symmetry of the structure is a major drawback of the current implementation. The problems related to the symmetry are most visible in the case of central symmetric structures. It is well known that in the case of central symmetry the structure factors $F_{\mathbf{K}}$ are real numbers (the same applies to some structure factors when the point symmetry group contains other elements of order 2). As a result, each of the constraints (8) on the absolute values of the Fourier components of $\rho(\mathbf{x})$ defines two disconnected manifolds in the nM-dimensional space of coefficients $\tilde{\psi}_{\alpha,\mathbf{K}}$. The total number of disconnected manifolds defined by all constraints is equal to 2^m, where m is the number of real structure factors. In its current implementation, the optimization algorithm sticks to one of these pieces in the early stages and then continues looking for the point of minimal average density on this piece only. In this way, the global minimum will most likely be missed.

The other problem is related to the fact that the symmetry of the function $\psi(\mathbf{x})$ does not necessarily coincide with that of the density function $\rho(\mathbf{x})$. In the general case, rotations and translations in real space are accompanied by orthogonal transformations in the n-dimensional space of components of ψ. In other words, the symmetry group of $\psi(\mathbf{x})$ is an extension of the symmetry group of $\rho(\mathbf{x})$ by a subgroup of $O(n)$. Generally speaking, the search for the solution corresponding to the global minimum of the average density should be performed over all such extensions.

APPENDIX *A*
Numerical algorithm

This section describes in detail the algorithm used to minimize $|\psi|^2$ with the constraints (8). The algorithm converges quadratically in the vicinity of a local minimum with the computational complexity of one iteration $O[(Mn)^3]$ and the storage requirement of the order $O[(Mn)^2]$. Here, M stands for the number of points in the set Λ in (3) and n is the number of components of $\psi(\mathbf{x})$. These properties are suitable for refining an already found approximate minimum. In other words, this algorithm should be used as a second stage in a two-stage optimization scheme. Nonetheless, the results of the testing described above show that this algorithm can work also in a single stage, starting with random values of parameters.

The minimal charge problem can be stated in the following abbreviated form:

$$\begin{aligned} &\text{minimize } |\psi|^2 \\ &\text{subject to } h_i(\psi) = c_i. \end{aligned} \tag{20}$$

Here, ψ is a vector in the real Mn-dimensional space. The minimization of its norm is subject to N non-linear equality constraints (20) representing the conditions (8). Note that, by squaring both sides of (8), h_i can be chosen in the form of uniform polynomials of the fourth degree in ψ. By this means, all derivatives of h_i are readily available in an analytic form. This makes appropriate application of the sequential quadratic programming (SQP) techniques of optimization (Biggs, 1975; Gill *et al.*, 1984). We have used the following scheme for an elementary SQP iteration:

(i) Compute the SVD (singular value decomposition) of the Jacobian $J = \nabla_\psi h$ of constraints: $J = USV$, where $UU^T = \hat{1}$, $VV^T = \hat{1}$ and S is an $N \times nM$ diagonal matrix. Denote the first N rows of V by $V_{\rm range}$ and the rest of the rows by $V_{\rm null}$. These matrices give the projectors correspondingly onto the range and null spaces of the constraints.

(ii) Compute the Newton step in the range space,

$$\delta\psi_{\rm range} = V^T_{\rm range} S^{-1} U^T (c - h),$$

where S^{-1} stands for a square $N \times N$ diagonal matrix of inverse singular values.

(iii) Compute the row vector of approximate Lagrange multipliers,

$$\lambda = -2\psi^T V^T_{\rm range} S^{-1} U^T.$$

(iv) Compute the Hessian of the Lagrangian function $L = \psi^T\psi + \lambda h$,

$$H = \partial^2 L / \partial\psi\partial\psi.$$

(v) Compute the projection of H onto the null space of the constraints $H_{\rm null} = V_{\rm null} H V^T_{\rm null}$.

(vi) Find the search step $\delta\psi_{\rm null} = V^T_{\rm null} x$ in the null space. Ideally, this step should minimize the quadratic approximation to the variation of the Lagrangian,

$$\delta L \simeq 2\psi^T V^T_{\rm null} x + (1/2) x^T H_{\rm null} x.$$

Note that the quadratic form $H_{\rm null}$ is not necessarily positive definite.

(vii) Update ψ,

$$\psi := \psi + \delta\psi_{\rm null} + \delta\psi_{\rm range}.$$

(viii) Repeat steps (i)–(vii) until convergence.

The storage requirements of the algorithm are dominated by the necessity to store a dense $nM \times nM$ Hessian matrix H and its projection for steps (iv)–(vi). The speed bottlenecks are steps (v) and (vi), both requiring $O[(nM)^3]$ multiplications. The SVD at step (i) and the computation of the Hessian at step (iv) take $O[N^2(nM)]$ and $O[N(nM)^2]$ multiplications correspondingly. Taking into account that M scales as N, the contribution of these steps to the computation time may be non-negligible for small n.

The only non-trivial point in the algorithm consists of handling possibly non-positive definite Hessian matrix at step (vi). The Hessian often has non-positive eigenvalues when the trial point is far from the local minimum. Note also that for the considered problem the local minimum is always degenerate because of translational invariance in the real space and

rotational invariance in the space of the components of the field $\psi(\mathbf{x})$. As a result, the Hessian at the local minimum always has at least $d + n - 1$ zero eigenvalues. The standard approach to this problem consists of modifying the Hessian matrix to make it positive definite. We use the fact that diagonalization of a symmetric matrix is numerically stable to represent $H_{\rm null}$ in the form

$$H_{\rm null} = \sum_i \mu_i \nu_i \nu_i^T,$$

where ν_i form an orthogonal basis in the null space and μ_i are the eigenvalues of $H_{\rm null}$. The modified Hessian matrix is then given by the formula

$$\tilde{H}_{\rm null} = \sum_i \min\{|\mu_i|, \varepsilon\} \nu_i \nu_i^T.$$

As the matrix $H_{\rm null}$ is dimensionless, the regularization parameter ε can be set to some constant numeric value, which should be much bigger than the machine ε but still small enough to keep the convergence quadratic in the vicinity of the minimum. The matrix $\tilde{H}_{\rm null}$ is then used to compute the full Newton update step in the null space,

$$x = -2\tilde{H}_{\rm null}^{-1} V_{\rm null} \psi.$$

The author is grateful to E. Tatarinova for fruitful discussions.

References

Ammann, R., Grunbaum, M. & Shephard, G. C. (1992). *Discrete Comput. Geom.* **8**, 1–25.

Araújo, J. de, Gomes, A. & da Cunha, J. (1996). *Solid State Commun.* **97**, 1025–1028.

Bak, P. (1986). *Phys. Rev. Lett.* **56**, 861–864.

Beenker, F. P. M. (1982). Technical Report 82-WSK-04. University of Technology, Eindhoven, The Netherlands.

Biggs, M. (1975). *Towards Global Optimization*, edited by L. C. W. Dixon & G. P. Szergo, pp. 341–349. Amsterdam: North-Holland.

Bloomfield, P. (1976). *Fourier Analysis of Time Series – an Introduction.* New York: Wiley.

Boissieu, M. de, Boudard, M., Hennion, B., Bellissent, R., Kycia, S., Goldman, A., Janot, C. & Audier, M. (1995). *Phys. Rev. Lett.* **75**, 89–92.

Cochran, W. (1952). *Acta Cryst.* **5**, 65–68.

Cochran, W. (1955). *Acta Cryst.* **8**, 473–478.

Duneau, M. & Audier, M. (1994). *Lectures on Quasicrystals*, edited by F. Hippert & D. Gratias, pp. 283–333. Les Ulis: Les Editions de Physique.

Duneau, M. & Katz, A. (1985). *Phys. Rev. Lett.* **54**, 2688–2691.

Elser, V. (1986). *Acta Cryst.* A**42**, 36–43.

Elser, V. (1999). *Acta Cryst.* A**55**, 489–499.

Gabardo, J.-P. (1999). *J. Math. Anal. Appl.* **239**, 349–370.

Gill, P. E., Murray, W., Saunders, M. A. & Wright, M. H. (1984). *Computer-Aided Analysis and Optimization of Mechanical System Dynamics*, edited by E. E. J. Haug, pp. 679–697. Iowa City: NATO.

Harker, D. & Kasper, J. S. (1948). *Acta Cryst.* **1**, 70–75.

Hauptman, H. & Karle, J. (1953). *Am. Crystallogr. Assoc. Monogr.* No. 3. Pittsburgh, PA: Polycrystal Book Service.

Janssen, T. (1986). *Acta Cryst.* A**42**, 261–271.

Kalugin, P. & Katz, A. (1993). *Europhys. Lett.* **21**, 921–926.

Kalugin, P. A., Kitaev, A. Y. & Levitov, L. S. (1985). *JETP Lett.* **41**, 145–149.

Karle, J. & Hauptman, H. (1950). *Acta Cryst.* **3**, 81–187.

Katz, A. (1989). *Number Theory and Physics*, edited by J. M. Luck, P. Moussa & M. Waldschmidt. Berlin: Springer-Verlag.

Lyonnard, S., Coddens, G., Calvayrac, Y. & Gratias, D. (1996). *Phys. Rev. B*, **53**, 3150–3160.

Rabiner, L. R. & Gold, B. (1975). *Theory and Application of Digital Signal Processing.* Englewood Cliffs, NJ: Prentice Hall.

Senechal, M. (1995). *Quasicrystals and Geometry.* Cambridge University Press.

Weeks, C. M., DeTitta, G. T., Miller, R. & Hauptman, H. A. (1993). *Acta Cryst.* D**49**, 179–181. (http://www.hwi.buffalo.edu/SnB/.)

Wolff, P. de, Janssen, T. & Janner, A. (1981). *Acta Cryst.* A**37**, 625–636.