Acta Crystallographica Section A
Foundations of Crystallography
ISSN 0108-7673

Received 28 November 2007
Accepted 15 April 2008

Convergence study of a Schrödinger-equation algorithm and structure-factor determination from the wavefunction

Kostas Bethanis,[a]* Pavlos Tzamalis,[a] Athanassios Hountas[a] and Georges Tsoucaris[b]

[a]Physics Laboratory, Department of Science, Agricultural University of Athens, 75 Iera Odos, Votanikos, Athens 118-55, Greece, and [b]Centre de Recherche de Restauration des Musees de France, C2RMF-UMR 171 du CNRS, Palais du Louvre, 75001 Paris, France. Correspondence e-mail: kbeth@aua.gr

The algorithm [Bethanis, Tzamalis, Hountas & Tsoucaris (2002). *Acta Cryst.* A**58**, 265–269] which reformulates the quantum-mechanical problem of solving a Schrödinger (S) equation in a crystallographic context has been upgraded and tested for many aspects of convergence. The upgraded algorithm in reciprocal space aims at determining a wavefunction $\Phi_{\mathbf{H}}$ such that (*a*) $\Phi_{\mathbf{H}}$ fulfils the S equation within certain precision and (*b*) $\Phi_{\mathbf{H}}$ minimizes by least squares the differences between the calculated structure factors from the wavefunction and the observed ones. Calculations have been made with three molecules (11, 41 and 110 non-H atoms in the asymmetric unit) for different numbers of initially given phases. Three main questions have been addressed: (I) Does the iterative calculation of the wavefunction converge? (II) Do the calculated wavefunctions converge to a unique set of $\Phi_{\mathbf{H}}$ values independent of the initial random set of $\Phi_{\mathbf{H}}$? (III) Is the calculated $\Phi_{\mathbf{H}}$ set a good approximation of a wavefunction able to produce within certain errors the correct values of the phases of the structure factors? Concerning questions (I) and (II), our results give a strong hint about fast convergence to a unique wavefunction independent of the arbitrary starting wavefunction. This is an essential prerequisite for practical applications. For question (III) in the case closer to the *ab initio* situation, the final mean phase error, respectively, for the three structures is 3, 26 and 28°. The combination of (*a*) and (*b*) in the upgraded algorithm has been proved crucial especially for the results concerning the larger structures.

1. Introduction: one-electron molecule and direct methods

X-ray diffraction diagrams allow the moduli of the structure factors $E_{\mathbf{H}}$ of the crystal structures to be obtained. The determination of the phases of the structure factors by using the observed moduli (solution of the phase problem) can be achieved by direct methods (DM) (Hauptman & Karle, 1953). The structure factors are then introduced into the calculation of Fourier series yielding the electron-density map

$$\rho^0(\mathbf{r}) = \sum_{\mathbf{H}} E_{\mathbf{H}} \exp(-2\pi i \mathbf{H}\cdot\mathbf{r}). \quad (1)$$

Localization of maxima on this electron-density map leads to the determination of the atomic positions in the unit cell according to the following sequence:

observed moduli of $E_{\mathbf{H}} \rightarrow$ direct methods (DM)
$\rightarrow$ calculated phases of $E_{\mathbf{H}} \rightarrow$ electron-density map
$\rightarrow$ positions of maxima $\rightarrow$ atomic coordinates.

At this point, one could wonder how this electron density [equation (1)] is linked to the quantum-mechanics (QM) wavefunction. We have shown (Bethanis *et al.*, 2002) that this link can be exemplified through an extremely simplified 'no inter-electronic interaction' model where *all electrons of a molecule are stripped off except one*. The resulting multi-atom ion recalls the well known example of H_2^+. Within this one-electron model, a fundamental objective of QM is to determine, by means of the Schrödinger equation (abbreviated as 'S equation'), the wavefunction $\psi(\mathbf{r})$, $\mathbf{r} \in \Re^3$, and finally an electron density for the one-electron model *given the coordinates of the nuclei*.

$$\rho(\mathbf{r}) = |\psi(\mathbf{r})|^2. \quad (2)$$

Clearly, this electron density is different from the theoretical electron density obtained from the multi-electron wavefunction. We consider thus in this work a kind of 'inverse' use of the S equation in the sense that we wish now to determine *the atomic coordinates of a crystal structure by means of the S equation given the X-ray diffraction data* $|E|^{obs}$. A similar

'inverse' formulation is used by Spence (1998) for the case of electron scattering from a periodic potential described by the S equation.

Summarizing, in the above sequence, the mathematical theory of DM is replaced by an algorithm entirely based upon the S equation. However, the 'philosophy' of the application of DM accumulated in the past decades is very useful in the practical implementation of our S-equation-based algorithm.

We are thus led to reformulate the one-electron S equation in a way adapted to the crystallographic information, *i.e.* the knowledge of the structure factors. The key for the reformulation resides in the use of reciprocal space and the subsequent replacement of the whole set of the atomic coordinates by the *infinite* set of structure factors. These two sets of data are clearly mathematically equivalent by Fourier transformation (hereafter FT). However, the behaviour of the S equation in the case of a *limited* number of structure factors is not known to our knowledge. The problem is even more difficult in the case where only the moduli of a limited set are known (*ab initio* problem in the crystallographic sense).

The algorithm devised in the paper of Bethanis *et al.* (2002) has shown the pertinence of this approach to the phase problem and the feasibility of the process in a particular case. However, further developments are necessary for a routine application in crystal structure determination and phase extension. A prerequisite for further advancements and standard use of the algorithm is the examination of the convergence of the S-equation algorithm in different cases and for different parts of the proposed algorithm.

It is to be noted that the idea of combining QM methods with experimental X-ray diffraction data has been used for extracting experimental wavefunctions (Jayatilaka & Grimwood, 2001). Karle *et al.* (1998) have also used the crystallographic information to enhance quantum-mechanical calculations. Conversely, in the present work we wish to borrow the QM mathematical machinery as an alternative to the DM theories; clearly this is a major change in approaching the phase problem.

2. The one-electron molecule and the Schrödinger equation in direct and reciprocal space

First, the S equation in the position space is rearranged as shown below in equation (3). The variable $\mathbf{r} \in \Re^3$ denotes the position vector of the electron. The Hamiltonian operator $\hat{H} = -(\Delta/2) + V(\mathbf{r})$ is composed of the Laplacian operator $-\Delta/2$ (kinetic term) and the potential operator $V(\mathbf{r})$ acting by ordinary multiplication. We then take the FT of both members of equation (3) and we obtain the S equation in reciprocal space (called momentum space in physics), equation (4). We adopt the usual crystallographic notation $\mathbf{H}$ for a vector of reciprocal space, as well as the abbreviation 'reciprocal wavefunction' $\Phi(\mathbf{H})$ and 'reciprocal potential' $W(\mathbf{H})$:

$$\psi(\mathbf{r}) \overset{\text{FT}}{\longleftrightarrow} \Phi(\mathbf{H}) \quad V(\mathbf{r}) \overset{\text{FT}}{\longleftrightarrow} W(\mathbf{H}).$$

In this FT, the Laplacian operator becomes ordinary multiplication by $2\pi^2H^2$; the ordinary multiplication $V(\mathbf{r})\psi(\mathbf{r})$ is transformed into the convolution of $W(\mathbf{H})\otimes\Phi(\mathbf{H})$ [equation (4)].

$$[-(\Delta/2) - \varepsilon]\psi(\mathbf{r}) = -V(\mathbf{r})\psi(\mathbf{r}) \tag{3}$$

$$\updownarrow \text{FT}$$

$$(2\pi^2\mathbf{H}^2 - \varepsilon)\Phi(\mathbf{H}) = -W(\mathbf{H}) \otimes \Phi(\mathbf{H}) \tag{4}$$

It is noted that the reciprocal wavefunction $\Phi(\mathbf{H})$ has a fundamental physical meaning in QM: the amplitude squared of the wavefunction, $|\Phi(\mathbf{H})|^2$, is the probability density of the electron's momentum. We will refer to the S equation without distinguishing between equation (3) in direct space and equation (4) in reciprocal space.

Within the context of the one-electron-molecule model of the present paper, the potential operator contains only the nucleus–electron attractive Coulomb term:

$$V(\mathbf{r}) = -\sum_j \frac{Z_j}{|\mathbf{r} - \mathbf{r}_j|} \overset{\text{FT}}{\longleftrightarrow} W(\mathbf{H}) = -\frac{s_2}{\pi}\frac{E(\mathbf{H})}{H^2}, \tag{5}$$

where $\mathbf{r}_j$ denotes the nuclei positions and s_2 will be defined below [see equation (6)]. This equation shows that in the FT the knowledge of the set of the atomic coordinates $\mathbf{r}_j$ is equivalent to an infinite set of structure factors $E(\mathbf{H})$. The crucial relation (5) makes possible the connection between the one-electron S equation and the crystallographic problem: the reciprocal potential $W(\mathbf{H})$ is equal, apart from a positive constant, to $-E(\mathbf{H})/H^2$, where $E(\mathbf{H})$ is the normalized structure factor whose theoretical value is defined in crystallography by equation (6) for a point-charge scattering model (Tsoucaris *et al.*, 2000, 2003):

$$E(\mathbf{H}) = \frac{1}{s_2}\sum_j Z_j \exp(2\pi i\mathbf{H}\cdot\mathbf{r}_j), \quad s_2 = \left(\sum_j Z_j^2\right)^{1/2}, \tag{6}$$

where Z_j is the atomic number of the jth atom in the unit cell.

For a periodic crystalline structure, the convolution integral in equation (4) is replaced by a discrete sum. Thus, the S equation (4) in momentum space is written as equation (7) which represents a precise relation between the crystallographic structure factors E and the reciprocal wavefunction Φ:

$$(2\pi^2H^2 - \varepsilon)\Phi_{\mathbf{H}} = \frac{s_2}{\pi}\sum_{\mathbf{K}}\frac{E_{\mathbf{K}}\Phi_{\mathbf{H}-\mathbf{K}}}{K^2}. \tag{7}$$

In the crystallographic problem, the theoretical values of the moduli of $E_{\mathbf{K}}$ are approximated by the experimental structure-factor magnitudes $|E_{\mathbf{K}}|^{\text{obs}}$ obtained by a standard calculation from the diffraction pattern (Giacovazzo, 2006). We write then:

$$(2\pi^2H^2 - \varepsilon)\Phi_{\mathbf{H}} \approx \frac{s_2}{\pi}\sum_{\mathbf{K}}\frac{|E_{\mathbf{K}}|^{\text{obs}}\exp(i\varphi_{\mathbf{K}})\Phi_{\mathbf{H}-\mathbf{K}}}{K^2}. \tag{8}$$

We can now express the problem of crystal structure determination within a quantum-mechanical context as: determine the reciprocal wavefunction $\Phi_{\mathbf{H}}$ which satisfies the S equation (8), where the observed moduli of the structure factors are introduced in the expression $E_{\mathbf{K}} \simeq |E_{\mathbf{K}}|^{\rm obs} \exp(i\varphi_{\mathbf{K}}^{\rm calc})$. The determination of the phases $\varphi_{\mathbf{K}}^{\rm calc}$ is the object of the iterative procedure shown in §3.

The energy ε is a negative quantity for bound states so that the coefficient $(2\pi^2H^2 - \varepsilon)$ in the left parts of equations (7) and (8) is always positive. This fact allows us to calculate the values of $\Phi_{\mathbf{H}}$ for all $\mathbf{H}$ by dividing the right-hand member of equation (8) with this positive coefficient. We thus obtain equation (9) which is the base of the iterative process described in §3:

$$\Phi_{\mathbf{H}} = \frac{1}{(2\pi^2H^2 - \varepsilon)} \frac{s_2}{\pi} \sum_{\mathbf{K}} \frac{1}{K^2} |E_{\mathbf{K}}^{\rm obs}| \exp(i\varphi_{\mathbf{K}}) \Phi_{\mathbf{H}-\mathbf{K}}. \quad (9)$$

It is interesting to note that, in a recent publication (Karabiyik, 2007), equation (8) has been reformed to include relativistic corrections based on a simplified spin-free one-component Dirac equation. The relativistic Dirac formulation has led to a formula differing from S equation (9) by higher terms in the denominator $(2\pi^2H^2 - 2\alpha^2\pi^4H^4 + 2\alpha^4\pi^6H^6 - \varepsilon)$, where α is the fine-structure constant defined as $\alpha = 1/c$.

Another fundamental relation between the reciprocal wavefunction $\Phi_{\mathbf{K}}$ and the usual crystallographic structure factors $F_{\mathbf{K}}$ is the convolution equation corresponding to the FT of the electron density, equation (2):

$$\sum_{\mathbf{H}} \Phi_{\mathbf{H}} \Phi_{\mathbf{H}-\mathbf{K}}^* = F_{\mathbf{K}} = |F_{\mathbf{K}}| \exp(i\varphi_{\mathbf{K}}) \overset{\rm FT}{\longleftrightarrow} \psi(\mathbf{r})\psi^*(\mathbf{r}) = \rho(\mathbf{r}). \quad (10)$$

Thus the calculation of the reciprocal wavefunction $\Phi_{\mathbf{H}}$ leads to structure-factor determination. The 'quality' of the calculated $\Phi_{\mathbf{H}}$ is evaluated by comparing the phases $\varphi_{\mathbf{K}}$ determined by equation (10) with the correct ones.

In turn, $F_{\mathbf{K}}$ is a complex number related to the normalized structure factor $E_{\mathbf{K}}$ by a positive factor $k_{\mathbf{K}}$ well known in crystallography (Giacovazzo, 2006). We then write

$$E_{\mathbf{K}} = k_{\mathbf{K}} F_{\mathbf{K}} \quad \text{with } k_{\mathbf{K}} > 0. \quad (11)$$

Therefore, the phases of $F_{\mathbf{K}}$ and $E_{\mathbf{K}}$ are identical. Thus $\varphi_{\mathbf{K}}$ determined by equation (10) can also be recycled in equation (8), a fact that plays an important role in the development of a self-consistent potential calculation developed in the second part of the algorithm described in §3.

3. Algorithm in reciprocal space

The general algorithm operating entirely in reciprocal space is based on a *self-consistent iterative method* which comprises two parts.

3.1. First part

The reciprocal wavefunction of the preceding cycle $n-1$, namely $\Phi_{\mathbf{H}-\mathbf{K}}^{(n-1)}$, is introduced into the right-hand member of equation (9*a*). The initially given phases $\varphi_{\mathbf{K}}^{(0)}$ and the corresponding moduli $|E|_{\mathbf{K}}^{\rm obs}$ are introduced in equation (9*a*) and they are kept constant throughout all cycles of the algorithm:

$$\Phi_{\mathbf{H}}^{(n)} = \frac{1}{(2\pi^2H^2 - \varepsilon)} \frac{s_2}{\pi} \sum_{\mathbf{K}} \frac{1}{K^2} |E_{\mathbf{K}}|^{\rm obs} \exp(i\varphi_{\mathbf{K}}^{(0)}) \Phi_{\mathbf{H}-\mathbf{K}}^{(n-1)}. \quad (9a)$$

No other E's are used in this part. However, the calculated Φ set provides values of $\varphi_{\mathbf{K}}$ through equation (10) not only for the reflections initially endowed with known phases but also for the unphased part of the observed reflection set. Thus, the initial structure-factor information is transferred and capitalized into the Φ set *via* the S equation. The calculated phases will be used for the evaluation of the mean phase error (MPE) [equation (17)] as stated in §4. Note that the first part uses the S equation alone. This part, though it may sometimes concern only a very small subset with initially known phases, is very important because it produces a Φ set endowed with structural information to be exploited in the second part. It is this wavefunction $\Phi_{\mathbf{H}}$ that presides over the whole algorithm.

In the present algorithm, we consider the value of the energy ε as an adjustable semi-empirical parameter and we set $\varepsilon = -N_{\rm asym}$ (= number of atoms in the asymmetric unit). Another possibility is to evaluate the value of ε for each iteration by

$$\varepsilon^{(n)} = \sum_{\mathbf{H}} 2\pi^2H^2 \Phi_{\mathbf{H}}^{(n-1)} \Phi_{\mathbf{H}}^{*(n-1)} - \sum_{\mathbf{H}} \sum_{\mathbf{K}} \frac{s_2}{\pi} \frac{E_{\mathbf{H}-\mathbf{K}}^{(n-1)} \Phi_{\mathbf{K}}^{(n-1)}}{|\mathbf{H}-\mathbf{K}|^2} \Phi_{\mathbf{H}}^{*(n-1)}.$$

The calculations carried out with these two possibilities (recomputed energy or fixed value) do not affect significantly the convergence of the iterative process.

3.2. Second part

If at the end of the iterative calculations of the first part the phase information conveyed by the Φ set is not satisfactory (for instance when the initially known set $\varphi_{\mathbf{K}}^{(0)}$ is very limited), then this Φ set can be the initial input into a more involved calculation. This calculation constitutes the second part of the algorithm which differs from the first part by two elements.

(I) One of these elements concerns the use of the S equation itself. The phases calculated by equation (10) of the initially unphased part of the diffraction data [denoted $\varphi_{\mathbf{K}}^{(n-1)}$ in equation (9*b*)] are not only evaluated but also reintroduced stepwise into the S equation (9*b*) at each cycle:

$$\Phi_{\mathbf{H}}^{(n)} = \frac{1}{(2\pi^2H^2 - \varepsilon)} \frac{s_2}{\pi} \sum_{\mathbf{K}} \frac{1}{K^2} |E_{\mathbf{K}}|^{\rm obs} \exp(i\varphi_{\mathbf{K}}^{(n-1)}) \Phi_{\mathbf{H}-\mathbf{K}}^{(n-1)}. \quad (9b)$$

The stepwise procedure has been adapted from the standard DM practice. The calculated Φ set provides values for all the phases at each step but only a small subset (in our tests about ten new phases at each step) is accepted and introduced in equation (9*b*) along with the corresponding observed moduli as extended terms of the potential expression $W_{\mathbf{K}}^{(n-1)} = E_{\mathbf{K}}^{(n-1)}/K^2$. The identification of the accepted subset of phases is made according to the usual criteria for phase selection.

Table 1
Results of tests of the Schrödinger-equation algorithm.

N_{asym}: No. of non-H atoms in the asymmetric unit. M_{refl}: No. of independent observed reflections with $|E| > 1.2$. MPE_{seq}: $MPE_{sequential}$. MPE_{dif}: $MPE_{differential}$.

(*a*) Crystal structure

	N_{asym}	Space group	Data resolution (Å)	M_{refl}
Sigi	11	$P\bar{1}$	0.81	292
Tcnq	41	$P1$	0.80	632
Pn1a	110	$P2_1$	1.10	802

(*b*) Results

	No. of initial phases	No. of iteration cycles	Φ-MPE_{seq} (°)	Φ-MPE_{dif} (°)	*E*-MPE for M_{refl} (°)
Case (*a*)					
Sigi	292	11	0.15	1.0	0.4
Tcnq	632	12	0.06	1.0	19
Pn1a	802	17	0.08	1.0	35/6†
Case (*b*)					
Sigi	11	126	0.006	1.0	16
Tcnq	41	21	0.23	1.1	41/22†
Pn1a	110	56	0.01	1.0	36/28†
Case (*c*)					
Sigi	6	166	0.008	1.7	34/3†
Tcnq	8	114	0.36	1.08	61/26†
Pn1a	10	35	0.7	0.9	77/28†

† The second part of the algorithm was also activated.

Similar to the first part of the algorithm, the initially given phases are not refined. Thus, in the second part, new extended potential terms are introduced along with the initial ones, refining stepwise the potential in reciprocal space. This recalls a self-consistent potential method in quantum mechanics. We note that the calculated Φ set is recycled in the S equation in two ways: (i) directly as $\Phi^{(n-1)}$ values and (ii) indirectly by the calculated $\varphi^{(n-1)}$ values through equation (10). This method (first part followed by self-consistent potential expansion) along with a starting multisolution procedure has been used for the *ab initio* determination of the crystal structure of TCNQ with 41 non-H atoms (Bethanis *et al.*, 2002). Nevertheless, tests with larger structures showed that further developments are necessary for a routine application in crystal structure determination.

(II) Such a new development is the other element of the second part of the algorithm which is achieved by introducing an additional criterion stemming from comparison between $|E|^{obs}$ and $|E|^{calc}$ obtained from equation (10): the minimization of the *R* factor [equation (12)] by varying the Φ variables:

$$R = \frac{\sum_{\mathbf{H}} |E_{\mathbf{H}}^{obs} - E_{\mathbf{H}}^{calc}|}{\sum_{\mathbf{H}} |E_{\mathbf{H}}^{obs}|} = \frac{\sum_{\mathbf{H}} |E_{\mathbf{H}}^{obs} - k_{\mathbf{H}} \sum_{\mathbf{K}} \Phi_{\mathbf{K}} \Phi_{\mathbf{K}-\mathbf{H}}^{*}|}{\sum_{\mathbf{H}} |E_{\mathbf{H}}^{obs}|}. \quad (12)$$

A similar least-squares (LS) procedure has been introduced by Sayre (1972) and Sayre & Toupin (1975) within the DM context. A main advantage of this element is that the algorithm uses from the start of part 2 the whole set of observed moduli $|E|^{obs}$ whereas only the $|E|^{obs}$ of the initial phased set of reflections have been used in the first part. The values of $\Phi_{\mathbf{K}}$ minimizing the *R* expression (12) are substituted in equation (9*b*).

3.3. Summary

The general algorithm, after capitalizing on the initially given information by the iterative procedure described in the first part, activates the second part which calculates a Φ set fulfilling both conditions: satisfying the S equation (9*b*) with self-consistent potential and minimizing the *R* factor [equation (12)]. At the end of the iterative calculations, the inverse FT of $\Phi_{\mathbf{H}}$ will provide the usual wavefunction $\psi(\mathbf{r})$ in direct space and therefore, the 'electron density' [equation (2)].[1]

4. Convergence study

First we note that in the usual quantum-mechanical problems the set of atomic coordinates is the *a priori* information introduced in the S equation [right-hand part of equation (5) in direct space]. In reciprocal space, this set is equivalent to an infinite set of structure factors [right-hand part of equation (5) in reciprocal space]. In this paper, we wish to examine the 'quality' of the wavefunction $\Phi_{\mathbf{H}}$ considering gradually decreasing structure-factor initial information expressed as initial phase information per atom. The information initially given within the crystallographic context is in fact a function of the structure size which can be expressed in terms of N_{asym}, the number of unique non-H atoms in the asymmetric unit. Thus we distinguish three cases:

(*a*) a '*quasi-infinite*' set of $E_{\mathbf{H}}$ phases is considered a set of the order of $10N_{asym}$ phases. Clearly, this set of phases is able to reveal all the atomic positions of the crystal structure;

(*b*) a limited set of phases (N_{asym} phases);

(*c*) a very small subset of phases (for instance 6 to 10; this is considered as a 'close to *ab initio*' case).

The convergence for all cases is evaluated with basic criteria pertaining to the following three questions.

(I) *Does the iterative calculation of the wavefunction* $\Phi_{\boldsymbol{H}}$ *converge?*

For this, we evaluate the 'sequential convergence' of the calculated reciprocal wavefunction $\Phi_{\mathbf{H}}$ with the following.

(i) The mean phase error between the $\Phi_{\mathbf{H}}$ values calculated in the $(n-1)$th and *n*th iteration cycle:

$$\Phi\text{-MPE}_{sequential} = \sum_{\mathbf{H}} |\omega_{\mathbf{H}}^{(n)} - \omega_{\mathbf{H}}^{(n-1)}|/M, \quad (13)$$

where $\omega_{\mathbf{H}}^{(n)}$ = phase of $\Phi_{\mathbf{H}}^{(n)}$ and *M* is the number of Φ's.

[1] Illustrative schemes for the first and second parts of the algorithm, additional figures for the variation of all indices in all cases and for all structures, diagrams showing the convergence for the Tcnq structure with decreasing given information, and data for the unpublished Tcnq structure have been deposited with the IUCr. These are available from the IUCr electronic archives (Reference: AU5063). Services for accessing these data are described at the back of the journal.

(ii) The R factor for the moduli ($R_{\rm moduli}$) of the $\Phi_{\mathbf{H}}$ values calculated in the $(n-1)$th and nth iteration cycle:

$$\Phi\text{-}R_{\rm moduli\_sequential} = \sum_{\mathbf{H}} ||\Phi_{\mathbf{H}}^{(n)}| - |\Phi_{\mathbf{H}}^{(n-1)}||/\sum_{\mathbf{H}} |\Phi_{\mathbf{H}}^{(n-1)}|. \quad (14)$$

(II) *Does the calculated wavefunction $\Phi_{\boldsymbol{H}}$ converge to a unique set of $\Phi_{\boldsymbol{H}}$ values independently of the initial random set of $\Phi_{\boldsymbol{H}}$?*

Here, we evaluate the 'differential convergence' between $\Phi_{\mathbf{H}}$ sets originating from two different seeds in the MS random-number-generator subroutine with the following indices:

$$\Phi\text{-MPE}_{\rm differential} = \sum_{\mathbf{H}} |\omega_{\mathbf{H}}^{(n){\rm iseed1}} - \omega_{\mathbf{H}}^{(n){\rm iseed2}}|/M \quad (15)$$

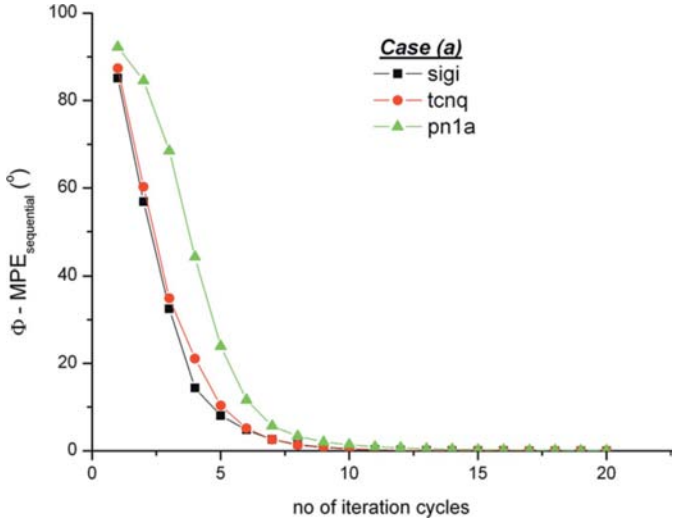

Figure 1
Convergence of Φ-MPE$_{\rm sequential}$ in case (*a*) for the three structures.

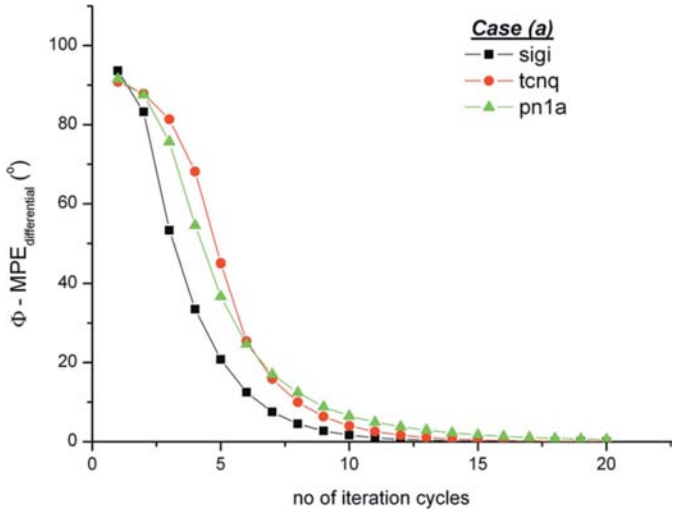

Figure 2
Convergence of Φ-MPE$_{\rm differential}$ in case (*a*) for the three structures.

$$\Phi\text{-}R_{\rm moduli\_differential} = \sum_{\mathbf{H}} ||\Phi_{\mathbf{H}}^{(n){\rm iseed1}}| - |\Phi_{\mathbf{H}}^{(n){\rm iseed2}}||/\sum_{\mathbf{H}} |\Phi_{\mathbf{H}}^{(n){\rm iseed1}}|, \quad (16)$$

where $\Phi^{(n){\rm seed1}}$ and $\Phi^{(n){\rm seed2}}$ are, respectively, the values of the Φ set calculated in the nth iteration cycle and initially generated from number seed1 = 123456 and seed2 = 427925 in the MS random-number-generator subroutine. Similar calculations have also been made with other seed numbers not given here yielding similar results.

(III) *Is the calculated $\Phi_{\boldsymbol{H}}$ set a good approximation of a wavefunction able to produce (within certain error) the correct values of the phases of the structure factors through equation*

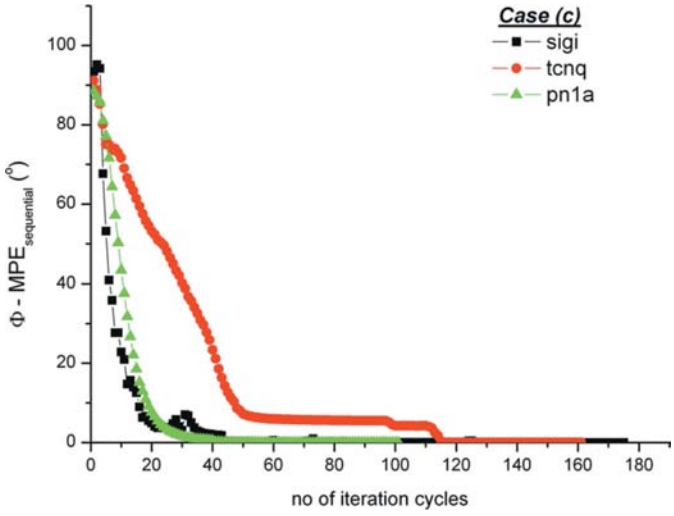

Figure 3
Convergence of Φ-MPE$_{\rm sequential}$ in case (*c*) for the three structures. The index evaluates the stability of the Φ set when only the first part of the algorithm is activated.

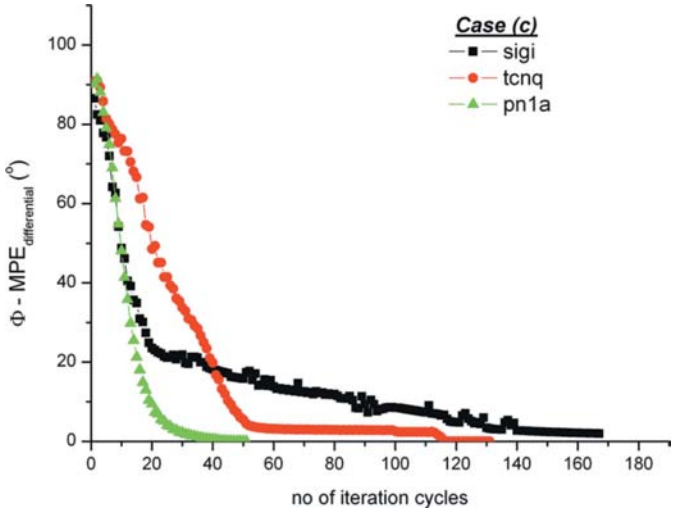

Figure 4
Convergence of Φ-MPE$_{\rm differential}$ in case (*c*) for the three structures. The index evaluates the convergence between two initially random Φ sets throughout the iterative procedure described in the first part of the algorithm.

(10) and thus determine the atomic positions of the crystal structure?

We evaluate thus the convergence of the calculated towards the correct set of phases using the corresponding *E*-MPE between calculated and correct *E*'s:

$$E\text{-MPE} = \sum_{\mathbf{H}} |\varphi_{\mathbf{H}}^{\text{calc}} - \varphi_{\mathbf{H}}^{\text{correct}}|/(\text{No. of calculated phases}), \quad (17)$$

where φ^{calc} is the phase calculated from equation (10) and φ^{correct} is the phase obtained from the known coordinates.

5. Results

We examine questions of convergence for three cases of different initially given information (*a*), (*b*) and (*c*) and three known structures with different numbers of atoms in the asymmetric unit and with different crystallographic symmetries (Table 1). The data of the two structures with code name Sigi and Pn1a have been obtained from the list of the *SHELX97* manual (Sheldrick, 1993, 2008). The data for the structure called Tcnq [4,5-bis(methylthio)-1,3-dithiol-2-ylium bis(TCNQ), $C_{29}H_{15}N_8S_4$] were kindly provided by D. Mentzafos (Bethanis *et al.*, 2002).[2]

In all cases, we use observed structure factors with $|E| > 1.2$. The Φ set corresponds to a *P*1 structure respecting the rule of Friedel reflections. Initial values of phases are arbitrarily assigned to the whole Φ set using the MS random-number-generation subroutine. The initial values of moduli $|\Phi_{\mathbf{H}}^{(0)}|$ are arbitrarily assigned as 1. In all cases, the initially given phase information is first transferred in the first part of the algorithm from the set of the known $E_{\mathbf{H}}$'s and capitalized into the whole set of the $\Phi_{\mathbf{H}}$ *via* iterations of the S equation (9). The indices given in equations (13)–(16) help in evaluating the convergence of the algorithm. The overall quality of the information transfer is evaluated through the values of *E*-MPE, equation (17). In order to simplify the presentation, we display in Table 1 and in Figs. 1–6 a selection of the above results. Additional figures for the variation of all indices in all cases and for all structures are deposited as supplementary material.[3] In Table 1, we give only the values of Φ-MPE and *E*-MPE. Φ-R_{moduli} factors are consistent with the corresponding Φ-MPE's.

The results allow one to evaluate the following.

(i) The '*stability*' of the final Φ set: the values of $\Phi^{(n-1)}$ introduced in the right-hand member of equation (9) yield sensibly the same values of $\Phi^{(n)}$. This is best illustrated in case (*a*).

(ii) The '*ability*' of the final Φ set to produce through equations (10) and (11) calculated phases close to the correct ones for the unphased part of observed reflections. This is illustrated in cases (*b*) and (*c*).

Concerning the '*stability*', the values of the convergence indices are very small after a reasonable number of calculations: for instance, the final Φ-$\text{MPE}_{\text{sequential}}$ is a small fraction of a degree and the final Φ-$\text{MPE}_{\text{differential}}$ converges to values of the order of a degree. The number of the corresponding cycles varies from 11 to 166. The convergence is slower in cases (*b*) and (*c*) where the initial phase information is much poorer than that of case (*a*). However, in all cases and all structures, the Φ-$\text{MPE}_{\text{differential}}$ is small and this corroborates the outstanding fact that the final set of phases is nearly independent of the randomly assigned set of initial reciprocal wavefunction $\Phi_{\mathbf{H}}$. In Figs. 1–4, the smooth decreasing of the Φ-MPE sequential and differential indices is shown for cases (*a*) and (*c*). Similar figures have been obtained for case (*b*).

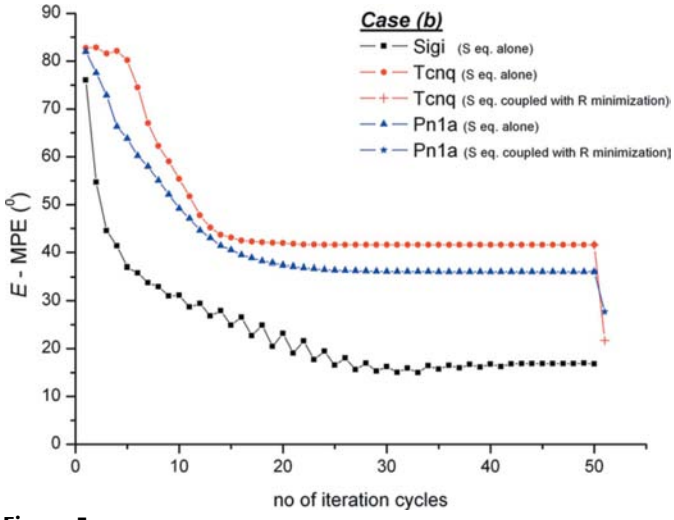

Figure 5
E-MPE in case (*b*) for the three structures. The second part of the algorithm (S equation coupled with *R* minimization) has been activated after the 50th iteration cycle for the structures with high (>30°) *E*-MPE values. The *E*-MPE values are then improved dramatically after just one iteration cycle.

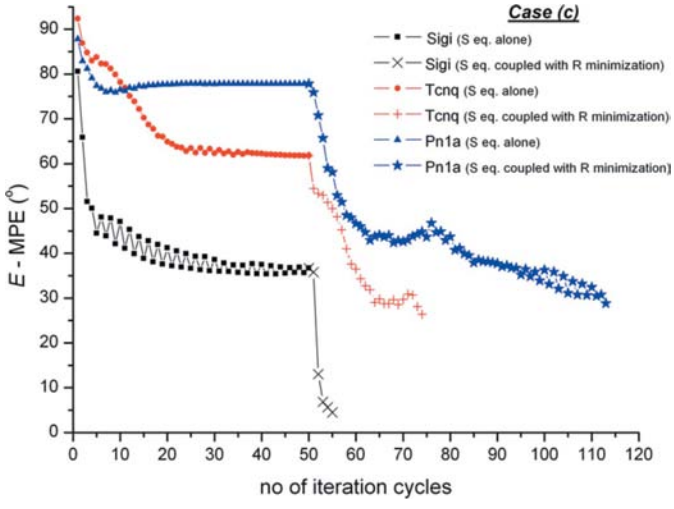

Figure 6
E-MPE in case (*c*) for the three structures. Until the 50th iteration cycle [only the 50 iteration cycles of the first part are shown instead 166 or 114 (see Table 1) because *E*-MPE has already been satisfactorily stabilized at this number of cycles], only the first part of the algorithm is activated. The high *E*-MPE values are dramatically improved after the 50th cycle by activating also the second part (S equation coupled with *R* minimization).

[2] Unit-cell parameters, atomic coordinates and experimental $|E|^{\text{obs}}$ have been deposited as supplementary material. See deposition footnote.
[3] See deposition footnote.

Concerning the '*ability*' of the final Φ set to produce correct phases $\varphi_{\mathbf{H}}$, crystallographic experience has shown that sets of phases with *E*-MPE say less than 30° are sufficient to reveal most if not all of the atomic positions. The *E*-MPE given in Table 1 corresponds to the full set of observed reflections ($M_{\rm refl}$). In the cases/structures where the *E*-MPE after the first part of the algorithm (S equation alone) is larger than 30°, the activation of the second part where the *R*-factor minimization is coupled with the S-equation fulfilment with self-consistent potential (§3) significantly improves the results. The latter fact is illustrated by using the abbreviation, for instance 77/28, to distinguish the results respectively without and with the second part of the general algorithm. The footnote symbol in Table 1 identifies the occurrence of such improvement.

In case (a), the algorithm is only meant to reproduce *via* the wavefunction $\Phi_{\mathbf{H}}$ the whole set of known phases whereas, in cases (*b*) and (*c*), the algorithm also produces phases for the unphased part of the whole set of reflections.

In case (b), clearly this unphased part covers the largest fraction of the whole set. For instance, for structure Tcnq this unphased part comprises 632 − 41 = 591 reflections. In other words, the algorithm achieves a phase extension from 41 to 632 phases. As shown in Fig. 5 for the structures Tcnq and Pn1a after 50 iteration cycles of the first part of the algorithm, the *E*-MPE has been stabilized to a value larger than 30° and the second part should also be activated. This activation is sufficient to lead to satisfactory *E*-MPE values (<30°) after just one iteration cycle.

In case (c), the initial data are extremely poor; in particular, for the $P2_1$ structure only 10 out of the selected 802 reflections have known phases. We note the extreme 'dilution' of the initial structure-factor information (10 reflections) into the large initially random Φ set (1600 Φ's). As one could expect, the finally stabilized Φ set after the calculations of the first part alone produces through equation (10) large values of *E*-MPE (34, 61, 77°) for the whole set of reflections $M_{\rm refl}$. However, having used this Φ set as the initial input to the *second part* of the algorithm described in §3, the results are then dramatically improved as shown in Table 1. In Fig. 6 is shown the variation of the *E*-MPE by using initially (until the 50th iteration cycle) only the first part of the algorithm and then by activating also the second part (after the 50th cycle). The final *E*-MPE values for the whole set of reflections $M_{\rm refl}$ are respectively 3, 26 and 28°. In Figs. 7, 8 and 9 are the electron-density maps drawn by using the final calculated phases which show that all the atomic positions are revealed for all crystal structures and thus the general algorithm is able to solve the close to *ab initio* problem. The *ab initio* solution can be achieved with the additional use of multisolution algorithms as has been shown by Bethanis *et al.* (2002).

5.1. Enhanced criterion of correctness from known space-group symmetry

A particular aspect of the correctness of the calculated phases is the fulfilment of the phase rules linking symmetry-

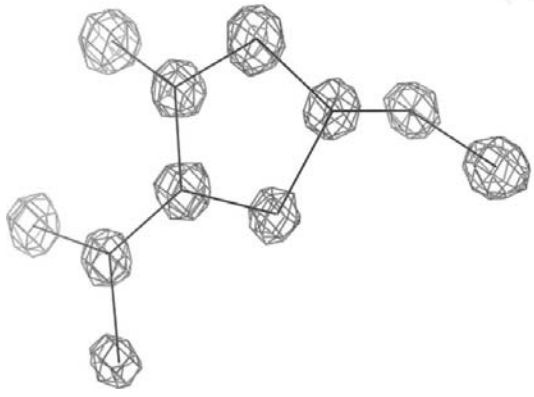

Figure 7
Final electron-density (e.d.) map for the Sigi structure ($N_{\rm asym}$ = 11). Number of initially known phases = 6, number of calculated phases = 292 with a final *E*-MPE = 3°.

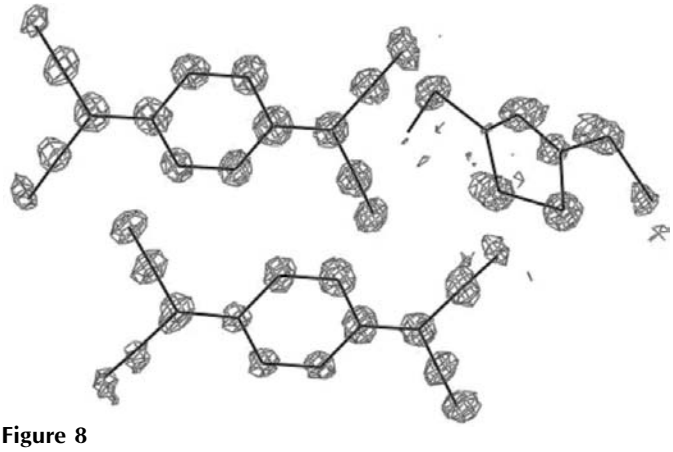

Figure 8
Final e.d. map for the Tcnq structure ($N_{\rm asym}$ = 41). Number of initially known phases = 8, number of calculated phases = 632 with a final *E*-MPE = 26°.

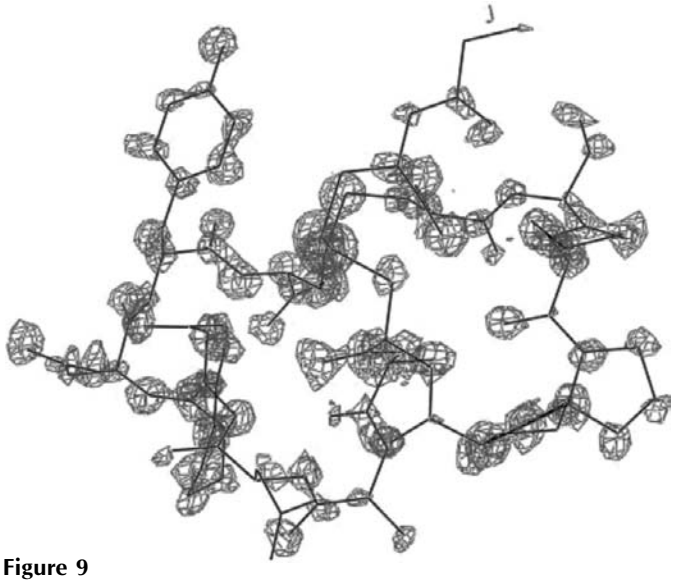

Figure 9
Final e.d. map for the Pn1a structure ($N_{\rm asym}$ = 110). Number of initially known phases = 10, number of calculated phases = 802 with a final *E*-MPE = 28°.

related reflections, whereas the initial Φ set lacks symmetry information (Tzamalis *et al.*, 2003). For instance, for the case (*b*) of the Pn1a structure, space group $P2_1$, reflections *hkl* and $\bar{h}k\bar{l}$ present a departure from ideal symmetry evaluated by the expression

$$\text{Sym\_}E\text{\_MPE} = \sum_{hkl} |\varphi_{hkl} - \varphi_{\bar{h}k\bar{l}} \pm k\pi| / (\text{No. of calculated phases}). \quad (18)$$

By activating only the first part of the general algorithm, we find at the end of the 40th cycle Sym_*E*_MPE = 1.6°. It is remarkable that this value is considerably lower than the value of *E*-MPE = 36° [equation (17)]. The phases of the Φ set present also a very low departure from ideal symmetry with a value of symmetry mean phase error Sym_Φ_MPE = 1.3°, where

$$\text{Sym\_}\Phi\text{\_MPE} = \sum_{hkl} |\omega_{hkl} - \omega_{\bar{h}k\bar{l}} \pm k\pi| / M. \quad (19)$$

The convergence of these two indices is shown in Fig. 10. The origin of this symmetry fulfilment resides in the initially given *E* values respecting the symmetry rules known in modulus and phase. The above low figures witness the successful transfer of information from *E*'s to Φ's performed by the S equation alone.

A particular case of symmetry is the space group $P\bar{1}$ where all structure factors are real numbers, *i.e.* the phases are 0 or π. Here also in cases (*a*) and (*b*) it is remarkable that an initial set of Φ's with random phases introduced only in the first part of the algorithm finally leads to a set of phases $\varphi_{\mathbf{H}}$ close to 0 or π. However, in case (*c*) of the centrosymmetric structure Sigi, the discrete character of the correct phases (0 or π) may cause a computational problem which is overcome by a large number of iterations (166).

6. Discussion

The proposed algorithm is based upon the Schrödinger equation in reciprocal space [equation (7)]. The connection with crystallography appears already in this equation where the $E_{\mathbf{K}}$ are the (theoretical) normalized structure factors for point atoms. This equation would be exactly fulfilled for an infinite number of terms in the second member. A first question arises about the degree of fulfilment for a finite number of terms with experimental E^{obs} in equation (9), a topic of fundamental interest which has not been studied to our knowledge. We have designed an algorithm to search for a valid approximation to the wavefunction $\Phi_{\mathbf{H}}$ which in turn would allow the calculation of unknown (or refined) structure factors *via* equations (10) and (11). A prerequisite for practical use is the convergence of this algorithm in different cases of *a priori* information and different sizes of crystal structures.

We come now to the three questions issued in §4.

(I) The wavefunction $\Phi_{\mathbf{H}}$ obtained by the iterative calculation using solely the S equation in the first part of the algorithm converges in all examined cases and for all examined structures to certain values.

(II) The final Φ values obtained as described in (I) are sensibly independent of the initial random set.

(III) For sufficient initial phase information, these values lead to a structure close to the correct one. When the initial information is not sufficient (in particular in the 110 atom structure), these values are used as input for a more involved method (second part) where the moduli of the calculated structure factors are constrained to be close to the experimental ones by a least-squares calculation. This innovating constraint over the Bethanis *et al.* (2002) publication is efficient, leading to an electron density where all atoms appear. However, we have found that the least-square procedure alone (without using the S equation) never leads to a meaningful set of unknown phases.

In conclusion, the use of the Schrödinger-equation method by itself is not sufficient to solve the structure in the close to *ab initio* case (*c*) for the 110 atom structure unless a least-squares step is used. Thus, the least-squares part is required in addition to the 'strength' of the S equation for difficult problems. The present results are encouraging for an application of the algorithm as a phase-extension procedure in protein structures, and ultimately to *ab initio* determination. The initial set of phases required to achieve an *ab initio* determination for small molecules can be obtained by multisolution methods like the magic integer methods (White & Woolfson, 1975; Main, 1977). For protein *ab initio* determination, a combination of real- and reciprocal-space techniques could provide the initial information like the shake-and-bake method (DeTitta *et al.*, 1994; Weeks *et al.*, 1994; Hauptman, 1995), the peaklist-optimization procedure (Sheldrick & Gould, 1995) and the *SIR99* procedure (Burla *et al.*, 1999). The new elements of the present paper are likely to allow determination of larger structures after optimization of the present algorithm. On the other hand, the least-squares concept can be extended to other pertinent functions and other constraints can be developed within the present algorithm such as that of solvent flattening

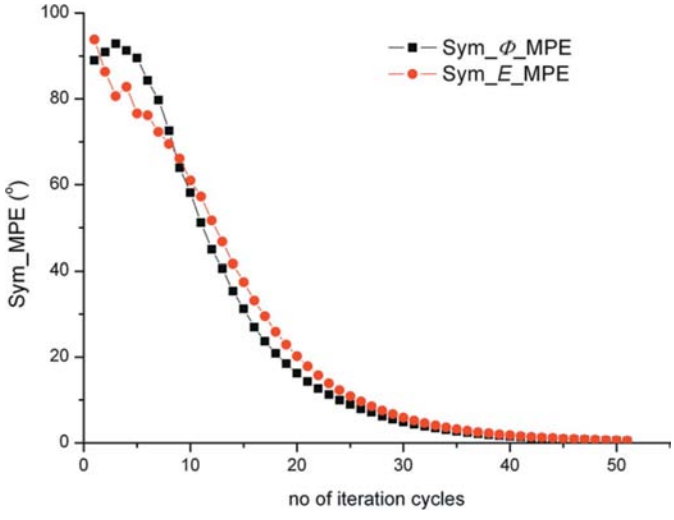

Figure 10
Convergence of indices showing the progressive re-establishment of the initially ignored $P2_1$ symmetry in case (*b*) for the Pn1a structure.

in protein structures or molecular envelopes (Zhang *et al.*, 2006). Note also that the sensitivity of lengthy iteration procedures to the rounding errors calls an increased attention to the convergence problems.

As a final remark, in a previous publication (Bethanis *et al.*, 2002), we have pointed out that 'the Schrödinger equation is a postulate of quantum mechanics'. The S-equation algorithm is conceptually different from any DM and thus it provides a new approach in crystallographic problems. The recent publication of a related algorithm based upon the Dirac equation (Karabiyik, 2007) leads to a formula close to our equation (7) and corroborates the physical background of this approach.

References

Bethanis, K., Tzamalis, P., Hountas, A. & Tsoucaris, G. (2002). *Acta Cryst.* A**58**, 265–269.

Burla, M. C., Camalli, M., Carrozzini, B., Cascarano, G. L., Giacovazzo, C., Polidori, G. & Spagna, R. (1999). *Acta Cryst.* A**55**, 991–999.

DeTitta, G. T., Weeks, C. M., Thuman, P., Miller, R. & Hauptman, H. A. (1994). *Acta Cryst.* A**50**, 203–210.

Giacovazzo, C. (2006). *International Tables for Crystallography*, Vol. B, *Reciprocal Space*, 1st online ed., edited by U. Shmueli, Section 2.2.4. Chester: International Union of Crystallography.

Hauptman, H. (1995). *Acta Cryst.* B**51**, 416–422.

Hauptman, H. & Karle, J. (1953). *The Solution of the Phase Problem. I. The Centrosymmetric Crystal*, Am. Crystallogr. Assoc. Monograph No. 3. New York: Polycrystal Book Service.

Jayatilaka, D. & Grimwood, D. J. (2001). *Acta Cryst.* A**57**, 76–86.

Karabiyik, H. (2007). *Theor. Chem. Acc.* **118**, 785–790.

Karle, J., Huang, L. & Massa, L. (1998). *Pure Appl. Chem.* **70**, 319–324.

Main, P. (1977). *Acta Cryst.* A**33**, 750–757.

Sayre, D. (1972). *Acta Cryst.* A**28**, 210–212.

Sayre, D. & Toupin, R. (1975). *Acta Cryst.* A**31**, S20.

Sheldrick, G. M. (1993). *SHELXL97*. University of Göttingen, Germany.

Sheldrick, G. M. (2008). *Acta Cryst.* A**64**, 112–122.

Sheldrick, G. M. & Gould, R. O. (1995). *Acta Cryst.* B**51**, 423–431.

Spence, J. C. H. (1998). *Acta Cryst.* A**54**, 7–18.

Tsoucaris, G., Bethanis, K., Tzamalis, P. & Hountas, A. (2000). *Direct Methods and Quantum Mechanics*, NATO Sciences Series, Vol. 33, *Twentieth Century Harmonic Analysis*, edited by J. S. Byrnes, pp. 384–386.

Tsoucaris, G., Bethanis, K., Tzamalis, P. & Hountas, A. (2003). Communication au Colloque de l'Association Française de Cristallographie, Caen, France.

Tzamalis, P., Bethanis, K., Hountas, A. & Tsoucaris, G. (2003). *Acta Cryst.* A**59**, 28–33.

Weeks, C. M., DeTitta, G. T., Hauptman, H. A., Thuman, P. & Miller, R. (1994). *Acta Cryst.* A**50**, 210–220.

White, P. S. & Woolfson, M. M. (1975). *Acta Cryst.* A**31**, 53–56.

Zhang, K. Y. J., Cowtan, K. D. & Main, P. (2006). *International Tables for Crystallography*, Vol. F, *Crystallography of Biological Macromolecules*, 1st online ed., edited by M. G. Rossman & E. Arnold, ch. 15.1. Chester: International Union of Crystallography.