Acta Crystallographica Section A
Foundations of Crystallography
ISSN 0108-7673

Received 7 May 2009
Accepted 21 September 2009

Improving the direct-methods sign-unconstrained S-FFT algorithm. XV

Jordi Rius* and Carles Frontera

Institut de Ciència de Materials de Barcelona (CSIC), Campus de la UAB, 08193-Bellaterra, Catalunya, Spain. Correspondence e-mail: jordi.rius@icmab.es

In order to extend the application field of the direct-methods S-FFT phase-refinement algorithm to density functions with positive and negative peaks, the equal-sign constraint was removed from its definition by combining ρ^2 with an appropriate density function mask [Rius & Frontera (2008). *Acta Cryst.* A**64**, 670–674]. This generalized algorithm (S_2-FFT) was shown to be highly effective for crystal structures with at least one moderate scatterer in the unit cell but less effective when applied to structures with only light scatterers. To increase the success rate in this second case, the mask has been improved and the convergence rate of S_2-FFT has been investigated. Finally, a closely related but simpler phase-refinement function (S_m) combining ρ (instead of ρ^2) with a new mask is introduced. For simple cases at least this can also treat density peaks in the absence of the equal-sign constraint.

1. Introduction

The moduli $|G|$ of the structure factors of the squared density distribution (ρ^2) are not measurable quantities. For a crystal structure with N equal scatterers in the unit cell following $P1$ space-group symmetry, $|G|$ can be derived from the observed $|E|$ values with $|G| = |E|/N^{1/2}$. However, this expression is not valid if positive and negative scatterers are present, because $|G|$ and $|E|$ are no longer proportional. To avoid this difficulty, Rius & Frontera (2008) introduced the squared-shape density function, *i.e.* $\rho_2(\mathbf{r}) = \rho^2(\mathbf{r})m(\mathbf{r})$, where $\mathbf{r}$ is an arbitrary point in the unit cell and $m(\mathbf{r})$ denotes a density function mask taking the value 1 if $\rho(\mathbf{r}) > t\sigma(\rho)$, -1 if $\rho(\mathbf{r}) < -t\sigma(\rho)$ or 0 if $|\rho(\mathbf{r})| \leq t\sigma(\rho)$ ($t \simeq 2.5$). This mask causes the peaks in ρ_2 to have the same shapes as in ρ^2 but preserves the signs they have in ρ. If G_2 represents the structure factors of ρ_2, this means that $|G_2|$ and $|E|$ are proportional and the expression $|G_2| = |E|/N^{1/2}$ is still applicable. The origin-free modulus sum function with the $|G_2|$'s as moduli takes the form (Rius & Frontera, 2008)

$$\begin{aligned} S_2(\Phi) &= \sum_{\mathbf{H}} (|G_{2,\mathbf{H}}| - \langle |G_2| \rangle) |G_{2,-\mathbf{H}}(\Phi)| \\ &= N^{-1/2} \sum_{\mathbf{H}} (|E_{\mathbf{H}}| - \langle |E| \rangle) |G_{2,-\mathbf{H}}(\Phi)|, \end{aligned} \tag{1}$$

in which Φ are the collectivity of phases of the E's and the calculated $|G_2(\Phi)|$'s are obtained by Fourier inverting $\rho^2 m$. The Φ phases can be optimized by maximizing S_2 with the direct methods S_2-FFT algorithm. Tests of S_2-FFT proved its viability for solving crystal structures with positive and negative peaks (Rius & Frontera, 2008). For structures (without negative scatterers) containing at least one moderate scatterer in the unit cell, the number of correct solutions for S_2- and S-FFT was similar (Rius, 1993; Rius *et al.*, 2007). However, for crystal structures with only light scatterers, the success rate of S_2-FFT was much smaller, although the true solutions were always characterized by the largest refined S_2 value. This result suggested that some improvements were still necessary.

2. An alternative mask for S_2

To improve the phase-refinement algorithm based on S_2, the mask definition is modified according to

$$\begin{aligned} m(\mathbf{r}) = &\quad 1 \text{ if } \rho(\mathbf{r}) > t\sigma(\rho) \\ &= -1 \text{ if } \rho(\mathbf{r}) < -t\sigma(\rho) \\ &= \quad a \text{ if } |\rho(\mathbf{r})| \leq t\sigma(\rho) \end{aligned} \tag{2}$$

with $t \simeq 2.5$ and where a is a random value between -1 and $+1$. This mask differs from the previously published one in the assignment of the random value a instead of zero for $|\rho(\mathbf{r})| \leq t\sigma(\rho)$. As described in §4, this modification has been tested by applying S_2-FFT to different data sets with better results.

3. The S_m function

For completeness, a new phase-refinement function (S_m) resulting from the simplest possible combination of ρ with an appropriate mask m is also investigated. In this function, the density function is modified according to the product

$$\rho_m = \rho m, \tag{3}$$

in which m takes the value

$$\begin{aligned} m(\mathbf{r}) &= 1 \text{ if } |\rho(\mathbf{r})| > t\sigma(\rho) \\ &= a \text{ if } |\rho(\mathbf{r})| \leq t\sigma(\rho) \end{aligned} \tag{4}$$

with $t \simeq 2.5$ and where a represents a random value between -1 and $+1$. To simplify the nomenclature and since no

Table 1
Application of the S_2-, S_m- and S-FFT phase-refinement algorithms to single-crystal data from compounds containing at least one moderate scatterer.

nsol = number of correct solutions; ntotal = total number of trials; ncycle = number of allowed cycles per trial; d_{min} = minimum d-spacing (Å). The number of correct solutions obtained with the old mask (Rius & Frontera, 2008) is given in parentheses in the S_2-FFT column.

					nsol/ntotal			ncycle		
Code	Formula	Z	Space group	d_{min}	S_2-FFT	S_m-FFT	S-FFT	S_2	S_m	S
Jul1	$C_9H_{10}N_3Cl_3$	8	$P4_12_12$	0.84	5(3)/25	4/25	3/25	40	40	40
Jul3	$C_{14}H_{22}S_2Si_2$	2	$P2_1/n$	0.76	10(8)/25	7/25	10/25	56	56	56
Jul4	$C_{11}Cl_{10}$	4	$P2_1/n$	0.70	6(6)/25	6/25	4/25	79	79	79
Jul5	$C_{11}HCl_9$	4	$Pca2_1$	0.70	25(22)/25	22/25	25/25	53	53	53
Hov1	$Pr_{14}Ni_6Si_{11}$	4	$C2/m$	0.78	20(21)/25	25/25	22/25	66	66	66
Bobby	$C_6H_6O_6NaCaN$	4	$P2_13$	0.84	10(9)/25	18/25	7/25	21	21	21
Cds	$Cd_4SO_{11.5}H_9$	4	$P6_3$	0.63	25(24)/25	25/25	25/25	34	34	34
Fina13	$C_{14}H_{19}N_2O_6ZnCl_3$	2	$P1$	0.84	25(25)/25	25/25	25/25	45	45	45
Cuimid	$C_6H_8N_4ClCu$	6	$P3_221$	0.71	7(6)/25	11/25	5/25	47	47	47

References: Jul1: Julià *et al.* (1992); Jul3: Alemán *et al.* (1993); Jul4, Jul5: Carilla *et al.* (1995); Hov1: Hovestreydt *et al.* (1983); Bobby: Barnett & Uchtman (1979); Cds: Loüer *et al.* (2001); Fina13: Pons *et al.* (2006); Cuimid: Clegg *et al.* (1984).

confusion is possible, no extra symbol is introduced for the mask of S_m, *i.e.* depending on the context m will specify either the mask of S_2 or that of S_m. By applying Parseval's theorem to equation (3) and by considering that $M = |M| \exp(i\alpha)$ are the Fourier coefficients of m, the structure factors of ρ_m in terms of Φ are

$$E_{m\mathbf{H}}(\Phi) = |E_{m\mathbf{H}}(\Phi)| \exp i(\psi_{\mathbf{H}}) = V^{-1} \sum_{\mathbf{h}} |E_{\mathbf{h}}||M_{\mathbf{H}-\mathbf{h}}| \exp i(\varphi_{\mathbf{h}} + \alpha_{\mathbf{H}-\mathbf{h}}), \quad (5)$$

where **h** in equation (5) refers to all reflections. In view of equation (5), the origin-free modulus sum function can be defined in terms of the $|E_m(\Phi)|$ moduli in the same way that S or S_2 are defined in terms of $|G(\Phi)|$ or $|G_2(\Phi)|$, respectively, *i.e.*

$$S_m = \sum_{\mathbf{H}} (|E_{\mathbf{H}}| - \langle |E| \rangle) |E_{m(-\mathbf{H})}(\Phi)| \quad (6)$$

$$= \sum_{\mathbf{h}} |E_{-\mathbf{h}}| \exp(i\varphi_{-\mathbf{h}}) [V^{-1} \sum_{\mathbf{H}} (|E_{\mathbf{H}}| - \langle |E| \rangle) \exp(i\psi_{\mathbf{H}}) M_{\mathbf{H}-\mathbf{h}}]. \quad (7)$$

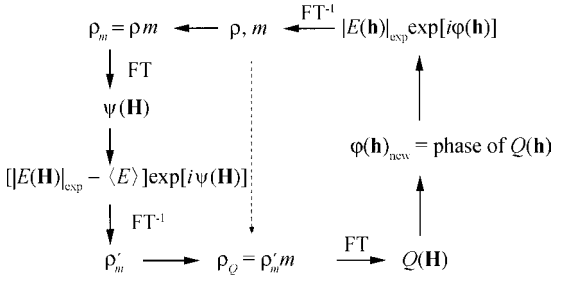

Figure 1
Flow diagram of the S_m-FFT phase-refinement procedure. The initial phases (upper right corner) are combined with the experimental amplitudes to compute the electron density ρ and the associated mask m. The latter is used to compute the phases ψ of ρ_m and also (broken arrow) the product function ρ_Q. The Fourier transform of ρ_Q yields the new structure-factor estimates.

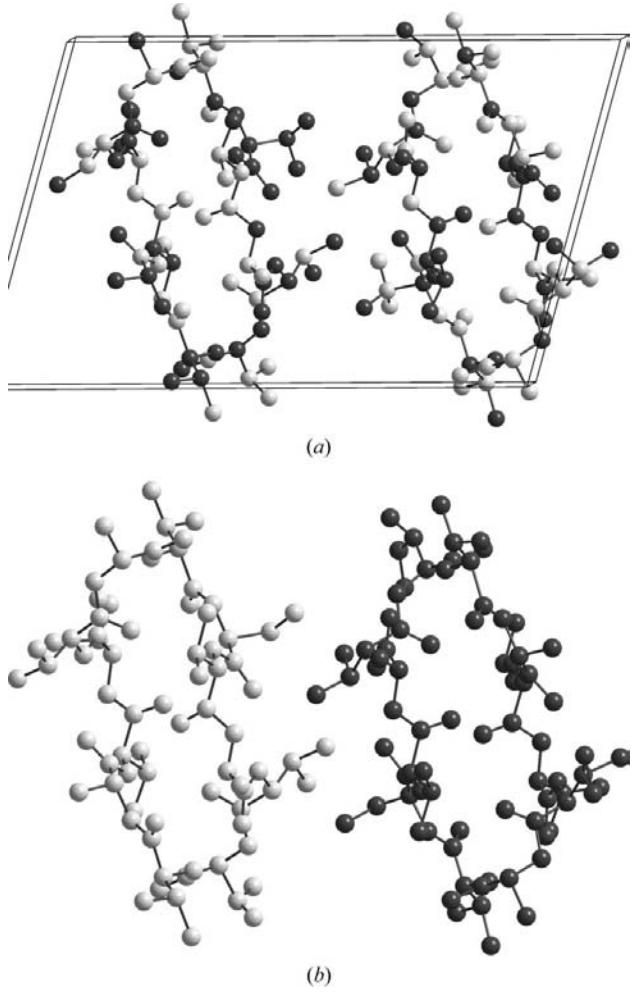

Figure 2
Application of the S_2-FFT phase-refinement algorithm to intensity data of Tval calculated (*a*) with 50% randomly assigned negative atomic scattering factors and (*b*) with all atomic scattering factors of one of the two symmetry-independent molecules made negative. Atoms with negative refined densities are depicted in grey. The peak search was performed on the Fourier map computed with the phase values supplied by S_2-FFT.

Table 2
Application of S_2-, S_m- and S-FFT to X-ray single-crystal data for compounds containing only light scatterers.

nsol, ntotal, ncycle and d_{min} are as in Table 1. The number of correct solutions with the old mask is given in parentheses in the S_2-FFT column.

					nsol/ntotal			ncycle		
Code	Formula	Z	Space group	d_{min}	S_2-FFT	S_m-FFT	S-FFT	S_2	S_m	S
Cortison	$C_{21}H_{28}O_5$	4	$P2_12_12_1$	0.89	40(26)/2500	45/2500	260/2500	150	150	41
Tpala	$C_{28}H_{42}O_7N_4$	2	$P2_1$	0.85	44(26)/400	22/200	120/400	150	150	62
Totc	$C_{33}H_{36}O_6$	6	$P6_1$	1.00	37(16)/200	111/200	140/200	100	100	50
Goldman2	$C_{28}H_{26}$	8	Cc	0.76	44(12)/200	164/200	196/200	120	120	72
Mbh2	$C_{15}H_{24}O_3$	3	$P1$	0.85	95(39)/200	40/200	197/200	120	120	70
Tval	$C_{54}H_{90}N_6O_{18}$	2	$P1$	0.85	77(41)/400	238/400	372/400	108	150	108

References: Cortison: Declercq *et al.* (1972); Tpala: Smith *et al.* (1981); Totc: Williams & Lawton (1975); Goldman2: Irngartinger *et al.* (1981); Mbh2: Poyser *et al.* (1986); Tval: Karle (1975) and Smith *et al.* (1975).

Table 3
Application of S_2- and S_m-FFT to calculated intensity data for crystal structures containing only light scatterers but with different percentages of negative atomic scattering factors in the unit cell.

The negative scatterers were assigned either randomly (rnd) or grouped in molecules (grp). The number of correct solutions with the old mask is given in parentheses in the S_2-FFT column. nsol, ntotal, ncycle and d_{min} are as in Table 1.

			nsol/ntotal		ncycle	
Code	Negative scattering factors	d_{min}	S_2-FFT	S_m-FFT	S_2	S_m
Mbh2(n)	50% rnd	0.94	100(83)/100	66/100	120	120
	50% rnd	1.00	80(49)/100	100/100	120	120
Tval(n)	10% rnd	0.82	78(51)/100	0/100	150	150
	50% rnd	0.82	74(44)/100	0/100	200	200
	50% grp	0.82	92(90)/100	99/100	150	150
Tval(n)	10% rnd	1.00	53(33)/100	97/100	150	150
	50% rnd	1.00	32(12)/150	0/200	200	200
	50% grp	1.00	47(31)/100	47/100	150	150

Note that in S_m the observed quantities are the experimental $|E|$'s. According to its definition, S_m makes no use of the equal-sign constraint, so it should be able to treat density functions with positive and negative peaks. A flow diagram of the S_m-FFT phase-refinement algorithm is shown in Fig. 1 and follows the same scheme as for S-FFT and S_2-FFT (Rius *et al.*, 2007; Rius & Frontera, 2008). Some test results showing the viability of this phase-refinement algorithm, at least for simple cases, are given in Tables 1, 2 and 3.

4. Discussion of the test calculations and conclusions

The results for the application of the sign-unconstrained S_2- and S_m-FFT phase-refinement algorithms are summarized in Tables 1, 2 and 3. However, the discussion of these results is limited to the S_2-FFT algorithm, since there is currently no firm indication that S_m-FFT is a significant improvement over S_2-FFT.

Table 1 shows the application to experimental X-ray data for compounds with moderate scatterers in the unit cell. It can be readily shown that the success rates of S_2-FFT with old and new masks and of S-FFT are similar for all data sets. Consequently, the effectiveness of S_2-FFT is preserved in this case.

Tables 2 and 3 include the corresponding results for compounds with only light scatterers.

(1) *Application to experimental intensity data from crystal structures containing only light scatterers*. Analysis of Table 2 indicates that the S_2 function can solve all data sets and that the introduction of the new mask in S_2 approximately doubles the number of correct solutions. However, when the performances of S_2- and S-FFT are compared, it is evident that the number of successful trials is smaller for S_2-FFT and that the convergence rate of the latter is somewhat lower, requiring about 2–3 times more refinement cycles than the S-FFT algorithm. In the first test calculations of S_2-FFT (Rius & Frontera, 2008), which were performed with the direct-methods program *XLENS* (Rius, 1993), the maximum number of allowed refinement cycles (ncycle) corresponded to the ideal values for S-FFT. Consequently, these values were in general too low and prevented S_2 from reaching convergence. This explains, at least in part, the modest effectiveness of S_2-FFT pointed out by the authors. The application of S_2 to Mbh2 data is a clear example of the effect of ncycle on nsol: with S-FFT a 96% success rate is already reached for ncycle equal to 40; however, with S_2-FFT, the evolution of nsol (for a total of 200 trials) as a function of ncycle (in parentheses) is 12.5% (40), 28.5% (70), 66.0% (120) and 64.0% (210).

(2) *Application to calculated intensity data from crystal structures containing only light scatterers but with different amounts of negative atomic scattering factors*. The first tests were performed on data calculated by randomly assigning positive and negative scattering factors to the 54 non-H atoms present in the asymmetric unit of Mbh2. As can be seen in Table 3, the refinements of the different sets of starting random phases by means of the S_2 function were quite satisfactory for the two resolution limits studied. The positive and negative density peaks showed up in the Fourier map at the expected positions.

A second more complex series of test calculations were carried out with Tval. The data sets were calculated by randomly (rnd) assigning negative form factors to 10% and 50% of the 160 symmetry-independent non-H atoms. As shown in Table 3, the S_2-FFT phase-refinement algorithm

produces the correct solutions in all cases. The high success rates of S_2 when the resolution limit is 0.82 Å are remarkable. Fig. 2(*a*) reproduces the positive and negative density peaks in the unit cell of Tval(n). To complement these test calculations and to see whether the non-randomness in the assignment of negative form factors has any influence on the refinement process, the negative form factors were grouped (grp) into one of the two symmetry-independent molecules of Tval. This represents 50% of non-randomly distributed negative scatterers in the unit cell. Table 3 clearly indicates that, for most trials, refinement with S_2-FFT converged to the correct phase values and developed the expected peak distributions in the subsequent Fourier maps (Fig. 2*b*).

The global analysis of Table 3 indicates that the introduction of the new mask in S_2-FFT increases the number of solutions by a factor of approximately 1.6.

The principal conclusions to be drawn from the test calculations are:

(i) The S_2-FFT algorithm is capable of refining starting random phases independently of the presence or absence of negative peaks in the density function.

(ii) The success rate of S_2-FFT increases by a factor of between 1.6 and 2 when the new mask is used for structures containing only light scatterers.

(iii) The convergence rate of the S_2-FFT phase-refinement algorithm is lower than for S-FFT (approximately 2–3 times lower for the studied test cases). This is the price that S_2 has to pay for not including the equal-sign constraint, especially when dealing with data from compounds containing only light scatterers.

(iv) The new and simpler S_m phase-refinement function can be used to refine phases. The full possibilities of S_m are still to be explored.

This work was supported by the Spanish 'Ministerio de Eduación y Ciencia' (Consolider NANOSELECT CSD2007-00041) and by the 'Generalitat de Catalunya' (SGR2009).

References

Alemán, C., Brillas, E., Davies, A. G., Fajarí, L. L., Giró, D., Julià, L., Pérez, J. J. & Rius, J. (1993). *J. Org. Chem.* **58**, 3091–3099.
Barnett, B. L. & Uchtman, V. A. (1979). *Inorg. Chem.* **18**, 2674–2678.
Carilla, J., Fajarí, Ll., García, R., Julià, L., Marcos, C., Riera, J., Whitaker, C. R., Rius, J. & Alemán, C. (1995). *J. Org. Chem.* **60**, 2721–2725.
Clegg, W., Acott, S. R. & Garner, C. D. (1984). *J. Chem. Soc. Dalton Trans. II*, pp. 2581–2584.
Declercq, J. P., Germain, G. & Van Meersche, M. (1972). *Cryst. Struct. Commun.* **1**, 13–15.
Hovestreydt, E., Klepp, K. & Parthé, E. (1983). *Acta Cryst.* C**39**, 422–425.
Irngartinger, H., Reibel, W. R. K. & Sheldrick, G. M. (1981). *Acta Cryst.* B**37**, 1768–1771.
Julià, L., Rius, J. & Suschitzky, H. (1992). *Heterocycles*, **34**, 1539–1545.
Karle, I. L. (1975). *J. Am. Chem. Soc.* **97**, 4379–4386.
Loüer, D., Rius, J., Bènard-Rocherullé, M. & Loüer, M. (2001). *Powder Diffr.* **16**, 86–91.
Pons, F., Rius, J. & Ros, J. (2006). *Inorg. Chim. Acta*, **359**, 379–382.
Poyser, J. P., Edwards, P. L., Anderson, J. R., Hursthouse, M. B., Walker, N. P. C., Sheldrick, G. M. & Whalley, A. J. S. (1986). *J. Antibiot.* **39**, 167–169.
Rius, J. (1993). *Acta Cryst.* A**49**, 406–409.
Rius, J., Crespi, A. & Torrelles, X. (2007). *Acta Cryst.* A**63**, 131–134.
Rius, J. & Frontera, C. (2008). *Acta Cryst.* A**64**, 670–674.
Smith, G. D., Duax, W. L., Langs, D. A., DeTitta, G. T., Edmonds, J. W., Rohrer, D. C. & Weeks, C. M. (1975). *J. Am. Chem. Soc.* **97**, 7242–7247.
Smith, G. D., Pletnev, V. Z., Duax, W. L., Balasubramanian, T. M., Bosshard, H. E., Czerwinski, E. W., Kendrick, N. E., Mathews, F. S. & Marshall, G. R. (1981). *J. Am. Chem. Soc.* **103**, 1493–1501.
Williams, D. J. & Lawton, D. (1975). *Tetrahedron Lett.* pp. 111–114.