Acta Crystallographica Section A
Foundations of Crystallography
ISSN 0108-7673

Received 21 January 2007
Accepted 23 July 2007

A charge-flipping algorithm incorporating the tangent formula for solving difficult structures

A. A. Coelho

26 Strawberry Road, Manly West, 4179 Brisbane, Australia. Correspondence e-mail: alancoelho@optusnet.com.au

The 'charge-flipping' method proposed by Oszlányi & Sütő [*Acta Cryst.* (2004), A**60**, 134–141] has been extended to include the direct-methods tangent formula within the iterative process. The tangent formula acts as a corrective influence allowing for solutions at resolutions poorer than 1 Å. The resulting algorithm solves difficult structures in minutes rather than days or not at all. Modifications include (i) flipping a percentage of charge rather than charge below a threshold value and critically (ii) dampening the magnitude of charge above the threshold; this impedes tangent-formula solutions comprising one or two very intense peaks in the electron density which is commonly known as the 'uranium atom solution'. For data at poor resolution, an alternate charge-flipping regime avoids uranium atom solutions by truncating electron-density pixels that are greater than half the maximum value.

1. Introduction

The iterative phasing process of 'charge flipping' by Oszlányi & Sütő (2004) solves easy structures in a short time period; this was demonstrated on observed data by Wu *et al.* (2004). Charge flipping has also been used to reconstruct incommensurately modulated structures by Palatinus (2004), demonstrating the fact that charge flipping is not confined to three dimensions. Baerlocher *et al.* (2007) included histogram matching (Zhang & Main, 1990) in a charge-flipping process modified for powder diffraction data.

Charge-flipping procedures solve structures in space group $P1$ as this is the only space group with complex phases for all reflections which allows for incremental phase changes. Charge flipping alone however does have problems solving difficult structures where the ratio of intense to weak reflections is low or when the resolution of the data is poor. Oszlányi & Sütő (2005) reported that for simulated data hundreds of thousands of iterations were required for the two difficult structures of agazud (Alexander *et al.*, 2002) and 1a7y (Schaefer *et al.*, 1998) with rates of success of 50 and 25%, respectively. These two structures as well as four others with space groups other than $P1$ are analyzed in this paper with regard to solving the structures at good resolution and at poor resolution.

The present charge-flipping algorithm includes the tangent formula (Karle & Hauptman, 1956) within the iterations of charge flipping in a manner that perturbs the process in a positive manner. This is demonstrated on data at poor resolution, worse than 1 Å, whereby charge flipping alone operating on phases at their optimum values often results in an increase in R factors and hence divergence. Inclusion of the tangent formula with an optimum number of E_h values and corresponding triplet phase relationships prevents divergence. Both the tangent formula and charge flipping operate on normalized structure factors $E_h = |F_{h,\rm obs}|/(\sum_i f_i^2)^{1/2}$, where $|F_{h,\rm obs}|$ are the observed structure-factor moduli and f_i are the atomic scattering factors and the summation is over the atoms in the unit cell.

Table 1
Charge-flipping algorithm incorporating the tangent formula with control parameter set for the difficult structures of agazud and 1a7y.

1	Randomly choose reflection phases between 0 and 360°.
2†	Set 50% of structure factors with the smaller E_h values to zero and set the moduli of other structure factors to observed values. Set non-observed structure factors including $F000$ to zero.
3	Inverse Fourier transform structure factors to obtain electron densities.
4†	Scale electron densities ρ such that the maximum equals 1.
5†	Determine the threshold charge δ such that 60% of charge lies below δ.
6	For $\rho < \delta$ set $\rho = -\rho$.
7†	For $\rho \geq \delta$ set $\rho = \delta + (\rho - \delta)^{\beta}$ with $\beta = \frac{1}{2}$.
8	Fourier transform to obtain structure factors.
9†	Add to phases with the highest E_h values a fraction of the difference between the values determined by the tangent formula and present values.
10	Continue from step 2.

† Modifications to the Oszlányi & Sütő (2004) algorithm.

2. The algorithm

The basic algorithm is presented in Table 1; it adheres to the charge-flipping process of Oszlányi & Sütő (2004) except for the steps marked as modifications. The modifications are designed to overcome incorrect tangent-formula solutions where high electron densities are obtained around a few grid

points with the rest being of low intensity; this effect is commonly known as the 'uranium atom solution' (Sheldrick, 1997). The control parameter settings in Table 1 are set for the solution of the difficult structures of agazud and 1a7y. Modifications are described where necessary.

Step 2 removes weak reflections from the system by setting 50% of them with smaller E_h values to zero. This perturbation reduces the risk of the system being trapped within a local minimum. Oszlányi & Sütő (2005) reported this modification as being less efficient than offsetting weak reflection phases by $\pi/2$. This offsetting of weak reflection phases increases perturbation which increases the amount of parameter space searched. In the present algorithm, however, increasing or decreasing perturbation is instead performed by varying β of step 7. Setting $F000$ to zero, step 2, was also used by Palatinus (2004) as opposed to a floating $F000$ as used by Oszlányi & Sütő (2004).

Wu *et al.* (2004) favored flipping a fraction of charge rather than charge below a chosen threshold charge, δ. This approach has been adopted in steps 5 and 6; it ensures that the amount of charge flipped is independent of the existence of uranium atom regions. In the present algorithm, 60% of charge is flipped. During the iterative process of the Oszlányi & Sütő (2004) algorithm for the structure of agazud, a δ set at 0.08% of the maximum electron density flips approximately 80% of charge and 88% at convergence. Flipping less charge potentially means that more difficult structures can be solved.

Changing the electron densities in the manner shown in steps 4 and 7 decreases strong intensities more than small intensities and hence reduces the occurrence of uranium atom regions. β corresponds to a real number; a value of $\frac{1}{2}$ seems to produce the best results for difficult structures. At $\beta = \frac{1}{2}$, a Gaussian region of charge is modified to another Gaussian with a full width at half-maximum that is $\sqrt{2}$ times that of the original. This introduces a temperature-like effect into the structure factors with added complications due to flipping; the result is higher R factors. R factors reported in this paper first scale calculated structure factors such that $\sum_{hkl} |F_{\mathrm{calc},hkl}| = \sum_{hkl} |F_{\mathrm{obs},hkl}|$.

Step 7 also introduces a significant amount of perturbation into the system; this added perturbation reduces the need for a high threshold. Tests indicate however that high threshold values where 80% of charge is flipped also lead to the correct solution.

For poor-resolution data or in cases where the amount of perturbation is too much, the charge-flipping regime of steps 5, 6 and 7 are replaced by equation (1).

$$\rho = |\rho| \tag{1a}$$

$$\rho = \begin{cases} |\rho|, & \text{for } |\rho| < 1/2 \\ 1/2, & \text{for } |\rho| \geq 1/2. \end{cases} \tag{1b}$$

Equation (1*b*) truncates high electron-density pixels that are greater than half the maximum electron density. The effect is to reduce the occurrence of uranium atom solutions whilst increasing R factors at convergence marginally. The reason R factors increase by a small amount is because relatively few pixels are greater than half the value of the maximum pixel. Note that data at poor resolution even for simple structures produce uranium atom solutions more often than data at good resolution; this is demonstrated in the *Analysis* section.

In step 9, a certain number of phases produced by charge flipping $\phi_{h,cf}$ are modified to produce $\phi_{h,\mathrm{new}}$ as shown in equation (2). This incorporates the tangent formula and the reliability factor M_h (Burla *et al.*, 1999).

$$\begin{aligned} \phi_{h,\mathrm{new}} &= \phi_{h,cf} + \alpha_h(\phi_{h,tf} - \phi_{h,cf}) \\ \tan(\phi_{h,tf}) &= T_h/B_h \\ T_h &= \sum_k E_h E_k E_{h-k} \sin(\phi_k + \phi_{h-k}) \\ B_h &= \sum_k E_h E_k E_{h-k} \cos(\phi_k + \phi_{h-k}) \\ \alpha_h &= M_h/M_{h,\max}, \quad M_h = \sqrt{T_h^2 + B_h^2}. \end{aligned} \tag{2}$$

The use of E_h implies atomicity for which the number of atoms and their types within the unit cell are required. Owing to the present approximate nature of the use of the tangent formula, it is sufficient to simply know the type of structure, light element for example, and thus $\sum_i f_i^2$ can typically be approximated. $M_{h,\max}$ correspond to the maximum reliability factor produced in the previous iteration resulting in α_h varying between 0 and 1 approximately. Unreliable phases with small α_h are changed by small amounts. $\phi_{h,\mathrm{new}}$ are calculated starting with the highest α_h down to the lowest with ϕ_k and ϕ_{h-k} of the triplet relations being updated.

Phases produced by the tangent formula $\phi_{h,tf}$ are not directly used but only a fraction of the difference $\phi_{h,tf} - \phi_{h,cf}$. Thus the tangent formula is only solved partially. Tangent formulae for approximately one third of the highest E_h values are included each with triplet phase relations corresponding to those with the highest $E_h E_k E_{h-k}$ values. Inclusion of 30 triplets per E_h increases the computational time by approximately 10 to 15% for an electron-density grid comprising $32 \times 32 \times 64$ points. An optimum number of triplets per E_h has not been thoroughly investigated; in cases where the resolution of the data is good, 30 triplets per E_h seems to be adequate. For poor-resolution data, up to 200 seem necessary.

The quantity α_{sum} of equation (3) provides means of detecting uranium atom solutions.

$$\alpha_{\mathrm{sum}} = \frac{1}{N_h} \sum_{h=1}^{N_h} \alpha_h. \tag{3}$$

With $\beta = 1$ and for difficult structures, α_{sum} rises to high values from a relatively low value in a matter of a few iterations. The rise is approximately two to three times that of its value at convergence. This is in contrast to charge flipping without the tangent formula and $\beta = 1$ which produces an α_{sum} of ~90% of the value at convergence.

2.1. Discussion of control parameters

The control parameters of Table 1 are the percentage of charge to flip (step 2), the value set for β (step 7), the percentage of reflections considered weak, the number of E_h values (step 9) and the number of triplets per E_h value. Note

that both β and the percentage of charge flipped are redundant in the case of using the charge-flipping regime of equation (1).

It is desirable to minimize the number of control parameters for successful use of the algorithm. In the *Analysis* section, a number of investigations have been performed in order to identify optimum values; it is not exhaustive particularly for the case of data at poor resolution where only a limited number of investigations has thus far been performed.

2.2. Discussion of perturbations and *R* factors

Perturbations that increase *R* factors both prior to and at convergence are the flipping of charge (step 2) and the setting of β to values less than 1 (step 7). These perturbations increase randomness in the system and therefore can inhibit the identification of solutions where the difference between *R* factors prior to and at convergence is too small to be recognized. The use of the tangent formula (step 9) is a positive perturbation where randomness is reduced; this is demonstrated in the *Analysis* section below.

The data analyzed in the present work at good resolution all exhibit measurable differences in *R* factors prior to and at convergence with 1a7y exhibiting the least amount of contrast. It is anticipated therefore that the solving of even more difficult structures may require reduction of some of the perturbations by, for example, increasing β to values closer to 1, by flipping less charge, by instead using the charge-flipping regime of equation (1) or by reducing the percentage of reflections considered weak.

3. Analysis

The present algorithm as implemented in the programs *TOPAS-Academic* V4.1 (Coelho, 2007) and *TOPAS* V4.1 (Bruker AXS, 2004) has been used for the following investigations. Structures investigated are shown in Table 2. *E* values are used in all cases for charge flipping and the tangent formula. Where the tangent formula is used, 3000 E_h values are included each with 30 triplet relations unless otherwise stated.

3.1. Range of convergence

The range of convergence (ROC) can be defined as the limit to which phases can randomly be varied from their true values before charge flipping fails to bring them back within a certain terminating criterion. The criterion used in the present analysis is failure to decrease the *R* factor for 20 consecutive iterations; it also sets $\beta = 1$ and 20% of reflections are considered weak. Phases start at true values corresponding to a particular enantiomorph; after the addition to each phase of a random amount with each phase receiving different amounts, the starting enantiomorph should still be favored except for the case of adding random amounts corresponding to a mean phase error of $\pm 90^\circ$. Even if a different enantiomorph is realized, the ROC analysis considers this a success.

Table 2
Structures investigated including references, number of non-H atoms in the asymmetric unit and space group.

No.	Reference	No. of atoms	Space group
1, agazud	Alexander *et al.* (2002)	219	$P1$
2, 1a7y	Schaefer *et al.* (1998)	314	$P1$
3	Li *et al.* (2003*b*)	43	$P\bar{1}$
4	Fukuoka *et al.* (2000)	17	$P4_12_12$
5	Karakurt *et al.* (2003)	23	$C2/c$
6	Li *et al.* (2003*a*)	53	$P\bar{1}$

The success rate is defined as the percentage of times a correct solution is found.

Without the tangent formula and for the 1a7y structure at 0.8 Å resolution and for simulated data using atomic scattering factors and no temperature factors, a 100% success rate is observed when a mean phase error of $\pm 75^\circ$ is introduced. At a mean phase error of $\pm 72.5^\circ$, only a 6% success rate is observed. Thus the ROC here corresponds to $\sim\pm 75^\circ$ where 100% success is observed. This demonstrates that charge flipping has a large range of convergence and that success is diminished sharply within a few degrees of the mean phase error. A 100% success rate at a mean phase error of $\pm 75^\circ$, however, means that phases randomly chosen will have 1 in 6 outside the ROC.

At 1 Å resolution, the ROC drops to $\sim\pm 67.5^\circ$, where a 100% success rate is observed. At $\pm 72.5^\circ$, a success rate of ~20% is observed. Inclusion of the tangent formula as described in equation (2) and again with $\beta = 1$ increases the ROC by $\sim 2.5^\circ$.

At such large ROC's, it is tempting to try and find the lowest *R* factors by starting a multitude of processes, each starting with random phases and each allowed to continue for a relatively small number of iterations with a terminating criterion as described above. Implementing this strategy proved inefficient in finding the lowest *R* factors in comparison to perturbations introduced at the iteration level.

Randomly assigning phases and starting the process without the tangent formula produced an average change in phase after matching from one iteration to another for the highest 100 E_h reflections of $\sim 28^\circ$. Monitoring the change in phase that the tangent formula equation (2) would have introduced but not actually enforcing the change sees the 28° reduce to $\sim 12^\circ$. In other words, the tangent formula seems to want to reverse the change of the charge-flipping step. Actually enforcing the change reduces the average phase change down to $\sim 9^\circ$. The tangent formula therefore reduces perturbation in the system. The additional perturbations brought on by setting $\beta = \frac{1}{2}$ and of designating 50% of reflections as weak compensates for the lack of perturbation when the tangent formula is used.

3.2. Using the tangent formula on simulate data – agazud and 1a7y

For comparison with the work of Oszlányi & Sütő (2005), simulated data are analyzed for the agazud and 1a7y structures. Here however simulated data at 1 Å are used rather

than at 0.8 Å. This corresponds to 7754 reflections for agazud and 12304 for 1a7y. This is approximately half the number of reflections at 0.8 Å of 15138 and 24063, respectively. The simulated data included atomic scattering factors and zero temperature factors.

Executing the algorithm of Table 1 100 times for each structure was 100% successful with agazud, taking 478 iterations on average and 1a7y 585; this corresponds to 23 and 28 s respectively on a 3 GHz personal computer. The maximum number of iterations was 4273 and 3414, respectively. Without the modification shown in Table 1, neither structure could be solved at 1 Å resolution in 200000 iterations.

The perturbations introduced in Table 1 increase the R factor, in particular in step 7. The R factor for agazud is ~0.75 prior to convergence and ~0.57 at convergence; these values do not include weak-reflection structure factors that are set to zero. Fig. 1 however shows that there is ample contrast in the R factors. At convergence, the average difference between the phases of the strongest 200 reflections and their true values after matching is 24°. This is small in comparison to the ROC and thus relaxing the perturbations immediately reduces the R factor.

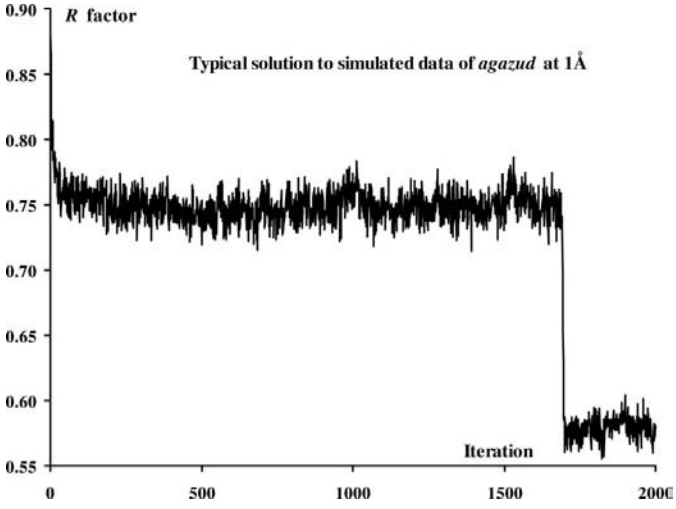

Figure 1
Solution to simulated data for agazud at 1 Å using the algorithm of Table 1.

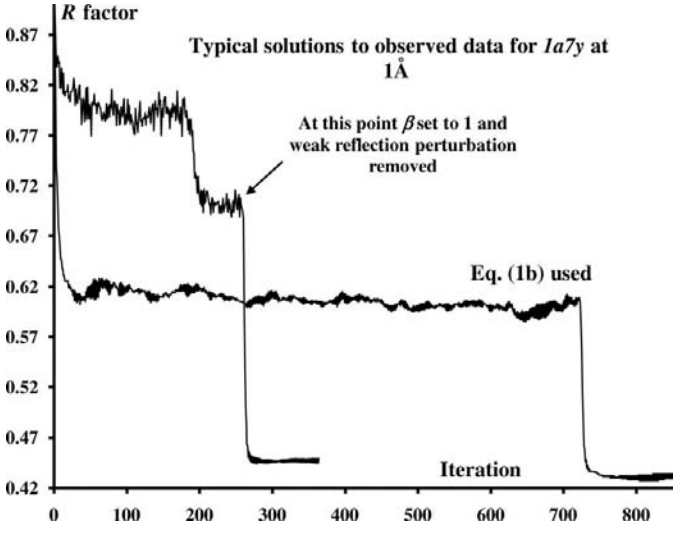

Figure 2
Solution to observed data at 1 Å for 1a7y using the algorithm of Table 1. The plot marked 'Eq. (1b) used' replaces the steps steps 5, 6 and 7 with the flipping regime of equation (1*b*) with 200 triplets per E_h value.

3.3. Using the tangent formula on real data – 1a7y

There are 14126 observed reflections for the structure 1a7y at a resolution of 0.940 Å. In the present analysis, a resolution of 0.994 Å is chosen with 12085 reflections and a corresponding grid size of $32 \times 32 \times 64$. The contrast between the R factors at convergence and prior to convergence is reduced in the real data and the strategy adopted is to detect a drop in the R factors of more than 0.05. At this stage, the perturbations of steps 2 and 7 are removed whereby weak reflection intensities are set to observed values and β is set to 1. Fig. 2 shows a typical solution with the point of perturbation removal indicated. With the perturbations, the electron density is not sharp; without them, the electron density is clearly recognizable as the correct solution. After matching the highest 100 E_h phases with the true values, an average difference of 26° is obtained.

The algorithm of Table 1 was executed 120 times with the 1a7y structure being solved 100% of the time. Fig. 3 shows a frequency distribution for the number of iterations required showing a peak at around 600. Without the tangent formula and with $\beta = 1$, a solution was not obtained in 100000 iterations.

Fig. 4 shows $\alpha_{\rm sum}$ of equation (3) as a function of iteration for 1a7y for the case of $\beta = 1$ and $\beta = \frac{1}{2}$. For $\beta = \frac{1}{2}$ and with the tangent formula (step 9 of Table 1), approximately the same $\alpha_{\rm sum}$ is seen as for $\beta = 1$ without the tangent formula. At $\beta = 1$ and with the tangent formula, $\alpha_{\rm sum}$ rises rapidly to a high value indicating a uranium atom solution; the same situation is observed for the agazud structure. Thus use of the tangent

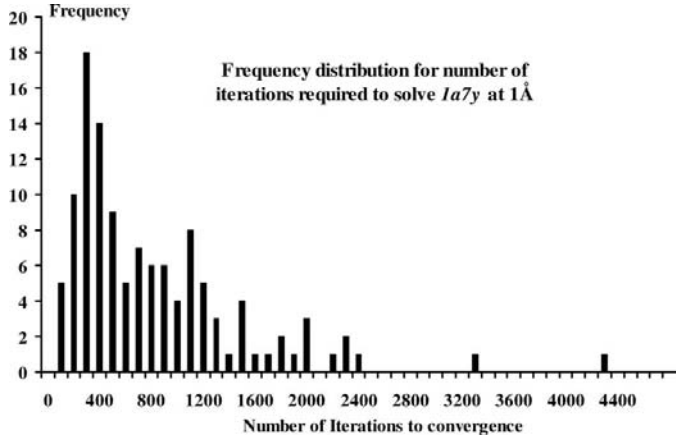

Figure 3
Number of iterations required to solve 1a7y plotted as a distribution comprising 120 executions of the algorithm of Table 1.

formula on these structures requires an additional perturbation, that of setting $\beta = \frac{1}{2}$ to prevent the formation of uranium atom solutions.

For $\beta = 1$, α_{sum} can be kept at reasonable levels for a while by increasing the amount of charge flipped. This increases perturbation due to flipping and decreases perturbation due to setting $\beta = 1$. Often, however, a uranium atom solution is finally realized and the *R*-factor plot resembles a typical convergence. Such a case for 1a7y where 85% of charge is flipped is shown in Fig. 5; the electron density confirmed a uranium atom solution.

Alternatively, the flipping regime of equation (1*b*) can instead be used replacing steps 5, 6 and 7 of Table 1. This increases the contrast in *R* factors as seen in the plot marked 'Eq. (1*b*) used' of Fig. 2. Because of the ability of equation (1*b*) to reduce uranium atom solutions, the number of triplets per E_h have been increased to 200. Equation (1*b*), however, introduces little perturbation and out of 20 different charge-flipping processes 6 became trapped within a local parameter space. Further perturbations may be necessary for difficult structures in order to search more of parameter space, for example, by offsetting the phases of weak reflections by $\pi/2$ and not setting their intensities to zero (Oszlányi & Sütő, 2005). Such a perturbation however decreases the contrast in *R* factor prior to and at convergence.

3.4. Changing the resolution and the number of triplets – 1a7y

Fig. 6 shows *R*-factor results with and without the tangent formula obtained for real data for 1a7y at 1.2 Å resolution with phases starting at their optimum values and with 20% of reflections considered weak; this corresponds to 6958 reflections with 5566 of them considered not weak. Without the tangent formula, the *R* factor diverges for both $\beta = 1$ and for $\beta = \frac{1}{2}$. This result indicates that charge flipping alone cannot solve the structure of 1a7y at 1.2 Å.

Table 3
Resolution used in poor-resolution analysis, number of reflections after converting to space group *P*1, number of E_h values and corresponding number of triplets per E_h; 50% of reflections considered weak.

No.	Resolution (Å)	No. of reflections	No. of E_h values	No. of triplets per E_h	Grid size (pixels)
3	1.4	1122	300	50	32×64×64
4	1.2	1743	500	200	128×32×64
5	1.3	4261	1500	50	64×64×64
6	1.6	981	500	200	32×64×64

With the tangent formula and 3000 E_h values each with 30 triplet relations, the *R* factor also increases for both $\beta = 1$ and $\beta = \frac{1}{2}$. In the case of $\beta = 1$, a uranium atom solution is quickly realized. In the case of $\beta = \frac{1}{2}$, a uranium atom solution is not realized but the solution is quickly lost with the *R* factor rising to 0.78. Increasing the number of triplets per E_h to 200 produces the plots marked as 'Tangent' in Fig. 6. A uranium atom solution is realized for $\beta = 1$ but not for $\beta = \frac{1}{2}$. In fact, at $\beta = \frac{1}{2}$, the solution is maintained as a function of iteration, and visual inspection clearly shows the solution with almost all atoms correctly picked.

Using the flipping regime of equation (1*b*) with 200 triplets per E_h and performing a ROC analysis produced results shown in Fig. 7. Here five convergences are shown: at the start of each the phases are offset from their optimum values by an average amount of ±60°. At ±60°, convergence is observed 100% of the time and, at ±65°, 30% of the time. Thus, the ROC is significantly smaller than at the high resolution of 0.94 Å, which produces a ROC between ±72.5 and ±75°. Fig. 7 also shows a superimposed plot marked 'Random start', which shows an equation (1*b*) process initiated with phases randomly chosen. The gap between this plot and the minima of the ROC plot is the contrast in the *R* factors which is appreciable at ~0.18. This ROC analysis is significant as it demonstrates that with enough computing power the flipping regime of equation

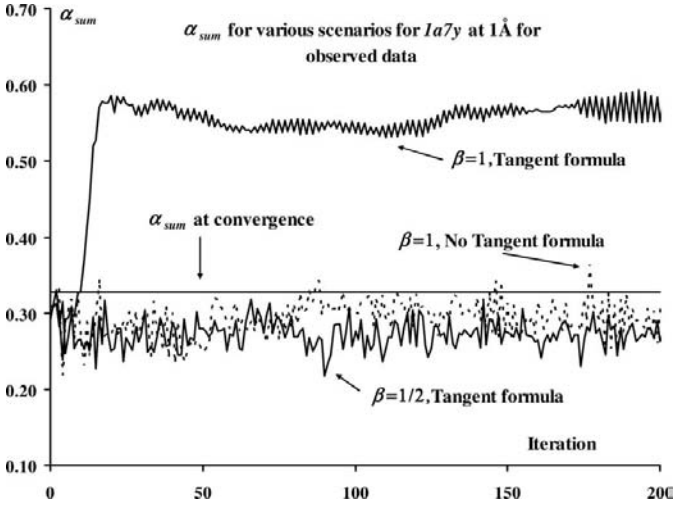

Figure 4
α_{sum} as a function of iteration for various cases of β with and without the use of the tangent formula of step 7, Table 1.

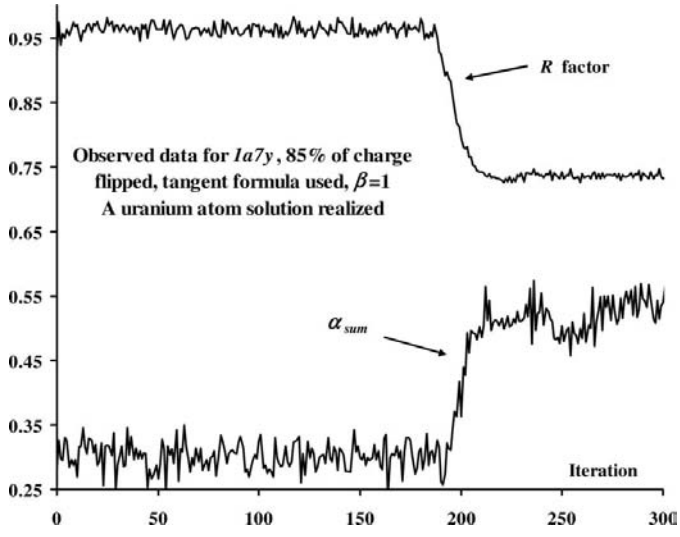

Figure 5
R factors, displaying what looks like a solution but is in fact a uranium atom solution.

Table 4
Lowest *R* factors recorded for poor-resolution analysis after 2000 iterations for two charge-flipping regimes each without (columns marked No TF) and with the tangent formula (columns marked TF).

	Comparison details, no tangent formula; all data used				*R* factors from resolutions as given in Table 3			
	All data		*R* factors calculated at resolutions given in Table 3		$\rho = \lvert\rho\rvert$		$\rho = \min(\lvert\rho\rvert, 0.5)$	
No.	Resolution (Å)	*R* factor	$\rho = \lvert\rho\rvert$	$\rho = \min(\lvert\rho\rvert, 0.5)$	No TF	TF	No TF	TF
3	0.78	0.32	0.35	0.34	0.46	0.72†	0.51	0.44
4	0.84	0.39	0.38	0.39	0.59	0.73†	0.63	0.54
5	0.66	0.46	0.45	0.47	0.62	0.50		
6	0.84	0.35	0.33	0.34	0.54	0.73†	0.58	0.43

† Uranium atom solutions realized.

(1*b*) together with the tangent formula can solve 1a7y at 1.2 Å. It is worth noting that a solution was not observed after continuing the 'Random start' plot for an additional 50000 iterations. This could be due to the system being caught in a local region of parameter space. A better strategy, not investigated, may be a periodic restart with phases randomly chosen.

3.5. Poor-resolution analysis using real data

Charge flipping when applied to simple structures achieves convergence in a few iterations (Oszlányi & Sütő, 2005). Structure 3 of Table 2 at 0.8 Å is solved with $\beta = 1$ and without the use of the tangent formula in 23.5 iterations on average. With the tangent formula and for $\beta = 1$ and $\beta = \frac{1}{2}$, the average drops to 14.6 and 16.3 iterations, respectively. Interestingly, there is no need to set $\beta = \frac{1}{2}$ as $\alpha_{\rm sum}$ remains unchanged with $\beta = 1$. Thus the behavior of $\alpha_{\rm sum}$ can determine whether a structure is difficult and whether β should be reduced from 1.

At 0.8 Å resolution, structures 3 to 6 of Table 2 are solved with $\beta = 1$ and without the tangent formula and in a matter of minutes or seconds. In reducing the resolutions to those shown in Table 3, charge flipping alone does not produce electron densities that are discernible to a high or even moderate degree. The electron densities can however be sharpened significantly by using the tangent formula with the charge-flipping regime of equation (1). In Table 4, the last four columns show the results with the control parameters as described in Table 3 for the regime of equation (1) and with and without the tangent formula. The number of triplets per E_h was determined by increasing the number of triplets in steps of 50 until the lowest *R* factor was obtained. The grid sizes shown in Table 3 are determined such that they are a power of 2 whilst keeping the grid spacing less than 0.4 Å.

The second and third *R*-factor columns (columns 4 and 5) show the similarity in *R* factors between the regimes of equation (1) demonstrating the fact that equation (1*b*) increases *R* factors marginally. With the tangent formula and the regime of equation (1*a*), structures 3, 4 and 6 produce uranium atom solutions within a small number of iterations. With the regime of equation (1*b*), the uranium atom solutions are suppressed; in addition, *R* factors are significantly better than for the case without the tangent formula. No uranium atom solutions were observed for structure 5 and thus equation (1*a*) is used with the *R* factor significantly better for the case with the tangent formula.

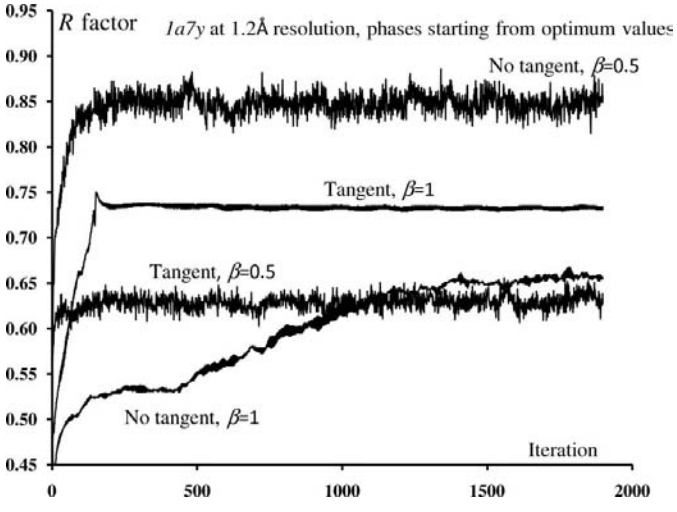

Figure 6
R-factor plots for real data for 1a7y at 1.2 Å resolution with phases starting from their optimum values for cases of $\beta = 1$, $\beta = \frac{1}{2}$ and with and without the tangent formula. 20% of reflections considered weak. 3000 E_h values used each with 200 triplet relations.

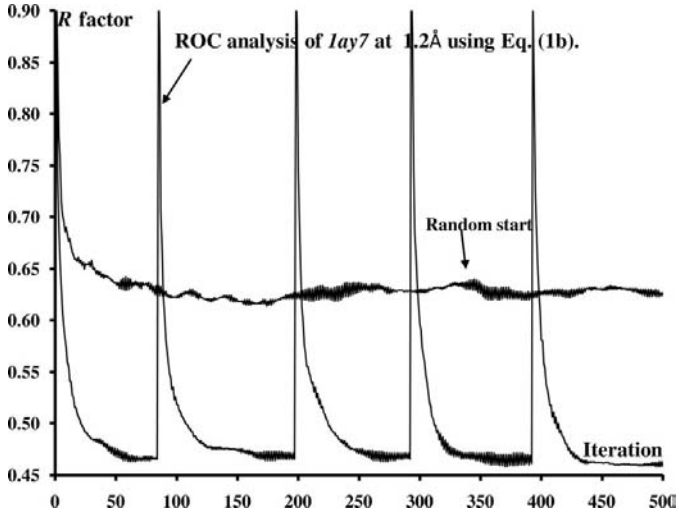

Figure 7
R-factor plots for real data for 1a7y at 1.2 Å for the flipping regime of equation (1*b*). The peaks on the plot marked 'ROC analysis' correspond to offsetting the phases a random amount of ±60° on average. The plot marked 'Random start' was initiated with random phases.

4. Conclusions

Techniques developed for detecting and avoiding uranium atom solutions have allowed the use of the tangent formula within a charge-flipping iteration. This extension to charge flipping seems to solve difficult structures at good resolution in a short time period. With the number of triplets per E_h optimized and further adjustments to control parameters, the use of the tangent formula can prevent divergence of R factors for data at poor resolution. A number of additional control parameters have been introduced in comparison to the original charge-flipping algorithm of Oszlányi & Sütő (2004) and even though many of them have been investigated there remain a number of unsolved matters. The most important being the determination of the optimum amount of perturbation and the type of perturbation whereby parameter space is searched efficiently whilst maintaining contrast in R factors sufficient enough to identify correct solutions. Additionally, further analyses regarding the use of negative quartets may prove worthwhile.

Note added in proof: Since submission of the manuscript, another paper published by Oszlányi & Sütő (2007) has described large increases in the performance of charge flipping by using normalized structure factors and accurate threshold values.

References

Alexander, J. M., Clark, J. L., Brett, T. J. & Stezowski, J. J. (2002). *Proc. Natl Acad. Sci. USA*, **99**, 5115–5120.

Baerlocher, C., McCusker, L. B. & Palatinus, L. (2007). *Z. Kristallogr.* **222**, 47–53.

Bruker AXS (2004). *TOPAS*, V4.1-beta. Bruker AXS, Karlsruhe, Germany.

Burla, M. C., Giacovazzo, C., Lamba, D., Polidori, G. & Ughetto, G. (1999). *Croatica Chem. Acta*, **72**, 519–529.

Coelho, A. A. (2007). *TOPAS-Academic* V4.1-beta. http://members.optusnet.com.au/alancoelho.

Fukuoka, H., Iwai, K., Yamanaka, S., Abe, H., Yoza, K. & Häming, L. (2000). *J. Solid State Chem.* **151**, 117–121.

Karakurt, T., Dinçer, M., Kahveci, B., Agar, E., Agar, A. & Sasmaz, S. (2003). *Acta Cryst.* E**59**, o1616–o1617.

Karle, J. & Hauptman, H. (1956). *Acta Cryst.* **9**, 635–651.

Li, X.-F., Feng, Y.-Q., Gao, B. & Li, N. (2003*a*). *Acta Cryst.* E**59**, o1626–o1627.

Li, X.-F., Feng, Y.-Q., Gao, B. & Li, N. (2003*b*). *Acta Cryst.* E**59**, o1638–o1639.

Oszlányi, G. & Sütő, A. (2004). *Acta Cryst.* A**60**, 134–141.

Oszlányi, G. & Sütő, A. (2005). *Acta Cryst.* A**61**, 147–152.

Oszlányi, G. & Sütő, A. (2007). *Acta Cryst.* A**63**, 156–163.

Palatinus, L. (2004). *Acta Cryst.* A**60**, 604–610.

Schaefer, M., Sheldrick, G. M., Bahner, I. & Lackner, H. (1998). *Angew. Chem. Int. Ed. Engl.* **37**, 2381–2384.

Sheldrick, G. M. (1997). *SHELXL97*, *SHELXS97* and *SHELX97-2*. University of Göttingen, Germany.

Wu, J. S., Spence, J. C. H., O'Keeffe, M. & Groy, T. L. (2004). *Acta Cryst.* A**60**, 326–330.

Zhang, K. Y. J. & Main, P. (1990). *Acta Cryst.* A**46,** 41–46.