TEACHING AND EDUCATION IN CRYSTALLOGRAPHY

Contributions intended for this section should be submitted to the Teaching and Education Editor, whose address appears on the inside front cover of the journal. They may describe any aspect of the teaching of crystallography and may take the form of short communications, teaching notes or full research publications.

Pepinsky's Machine: an interactive graphics-based Fourier synthesis program with applications in teaching and research

Nicholas M. Glykos *at IMBB, FORTH, PO Box 1527, 71110 Heraklion, Greece. E-mail: glykos@crystal2.imbb.forth.gr*

Abstract

A computer program has been developed which, given a set of structure-factor amplitudes for any centrosymmetric plane group, displays the amplitude-weighted reciprocal-lattice plane and allows the user interactively to assign and modify the phases of the structure factors, while observing the effect of these changes on the corresponding electron density function. The program has the added feature of being able to calculate and interactively display the electron density maps corresponding to all phase combinations of a user-defined subset of structure factors. Applications of the program in both crystallographic teaching and research are discussed.

1. The need for yet another Fourier synthesis program: the permutation syntheses method

Given a set of observed amplitudes for a small number of centrosymmetric structure factors, Fourier syntheses can be calculated for all their unique phase (sign) combinations and the resulting electron density maps can then be examined for the presence of correct features, or, as is usually the case, for the absence of unreasonable features, thus allowing the identification of putatively correct phase combinations and in favourable cases the extraction of useful structural information. This is the basis of the permutation syntheses method, which has found some use in small-molecule crystallography (Woolfson, 1954, Wright, 1958*a,b*) and was responsible for 'the first direct picture of a small protein molecule' (Boyes-Watson & Perutz, 1943). It has been shown recently (Glykos *et al.*, 1998*a,b*) that careful use of the method can give low-resolution protein projection maps which are in excellent agreement with electron microscopical reconstructions of the same projections. Unfortunately, the calculation of permutation maps for more than three or four reflections was a tedious procedure involving manual preparation of numerous ASCII (American standard code for information interchange) files containing h, k, F and $\sigma(F)$ data, their conversion to binary .mtz files [for use with the *CCP*4 suite of programs (Collaborative Computational Project, Number 4, 1994)] using the program *F2MTZ*, Fourier transformation with the program *FFT*, contouring with the programs *PLUTO* or *NPO*, and display with the program *XPLOT84DRIVER* (all programs from the *CCP*4 suite). Our need for an easier, more direct and less error-prone way to carry out these calculations led to the development of a new program, which was named after the creator of X-RAC, the

first electronic analogue computer for X-ray analysis (Pepinsky, 1947). Although the program was written with a research purpose in mind, it turned out to be a useful tool for teaching and illustrating the application of Fourier series in a crystallographic environment. In the following sections, the program and its user interface are described, and its application to crystallographic teaching and research is discussed.

2. Program description

2.1. *Input file*

The input to the program has been kept to the absolute minimum. Only one formatted ASCII file is required, containing the cell parameters, the plane-group number (as given in the *International Tables for X-ray Crystallography*, 1952) and a list of reflections [h, k, F and $\sigma(F)$] in free format. A typical input file would be:

```
CELL   81.40   30.40   90.00
SYMMETRY  7
  2   0   237.10   3.30
  4   0   264.90   3.40
  6   0   346.80   4.90
  · · ·
```

2.2. *The graphical user interface*

Fig. 1 shows the arrangement of items in the working window of the program. The screen area is conceptually divided into three parts: a text area near the top left-hand corner, the weighted reciprocal-lattice area below it, and the Fourier transform part, which covers approximately two-thirds of the total screen area and where the Patterson function or electron density maps are plotted (depending on whether any phase information is available or not).

The program starts by plotting the amplitude-weighted reciprocal lattice, and, since no phase information is available when a new session starts, it also plots the corresponding Patterson function. It then informs the user of the average and maximum observed amplitude, the size of the grid for the fast Fourier transform (FFT) and the number and parities of the origin-fixing reflections.

The user can then place the graphics cursor on any of the displayed reflections and set its phase to be positive, negative or unknown by pressing the corresponding mouse buttons. The

program updates the phases of all symmetry-related reflections (depending on the plane-group symmetry), re-draws the weighted reciprocal-lattice representation (the reflections are colour coded according to their phase), and then calculates and plots the Fourier transform using only those reflections which have been assigned a trial phase. A series of graphical buttons (located below the weighted reciprocal-lattice area) allow the user to increase or decrease the size (radius) of the reflections, the size of the density map, the number of contour levels, the size of the grid used for the FFT, to clear all phase assignments and go back to the Patterson function, and, finally, to write a *PostScript* (Adobe Systems Inc.) file containing the current map, weighted reciprocal lattice and phase assignments. While zooming is active, the electron density map can be moved with the arrow keys, thus allowing the user (once the origin is fixed) to concentrate on an area of interest.

The program supports a second mode of operation, termed the 'permutation syntheses mode'. In this mode the user selects (by clicking on them) a number of reflections, the phases of which will be systematically permuted; the program calculates all their unique sign combinations, combines these with the fixed phases (if any) and uses the resulting phase sets to calculate and plot the corresponding electron density maps. The permutation maps can either be shown one-by-one (with an option for the user to characterize and save the corresponding phase combination as 'good', 'bad' or 'ignore'), or in a nonstop fashion for quickly estimating the clarity of the resulting maps.

3. Applications

3.1. *Applications in teaching*

Pepinsky's Machine can be a useful teaching tool both at an introductory level and at a more advanced (research student) level. At the introductory level, the program can be used to illustrate the relationship between real space and reciprocal space, to show how a crystal structure can be built up by summing plane (line) waves, and to exemplify the relative importance of the amplitude and phase information in determining a crystal structure. It can also be used as the graphical environment for a 'spot the symmetry elements' exercise, but covering only a subset of the plane groups.

At a more advanced level, the problem-solving architecture of the program can be used to great advantage: the instructor can calculate structure factors for a simple few-atom structure and give the students the resulting amplitudes, together with the unit-cell parameters, plane group and chemical formula of the compound, and ask them to determine the structure. Depending on the skill of the student, the structure can have some regularities which will give rise to a few outstandingly strong reflections, thus making the task easier. The fact that the program automatically displays the Patterson function can be used to set up a heavy-atom exercise: the positional parameters of the heavy atom(s) are determined from the F^2 synthesis and the signs of the strongest reflections are calculated based on the heavy atom alone (the geometrical structure-factor calculation for just one heavy atom and for only a
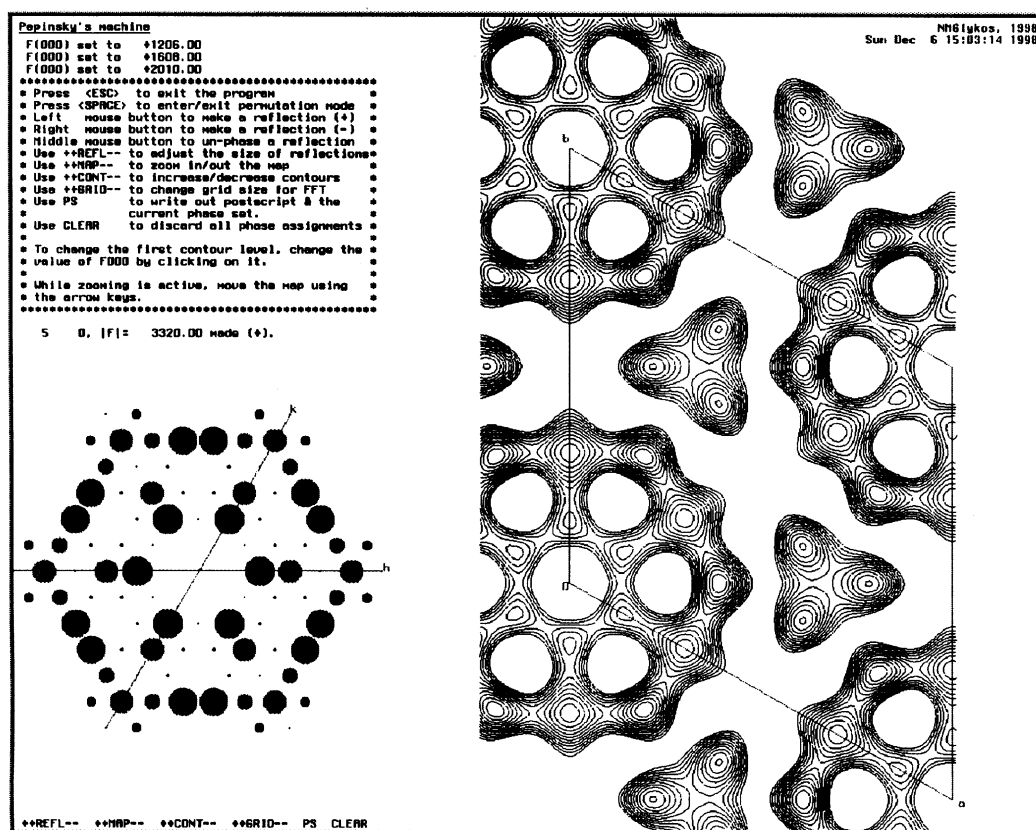

Fig. 1. A snapshot from *Pepinsky's Machine* in action.

few reflections is by itself a good exercise); these sign indications are then used to complete the determination. The current program distribution includes a few exercises, most of them for the advanced student level.

3.2. Research applications

As discussed in the opening paragraph, this program was originally written with the purpose of simplifying the task of calculating and examining permutation syntheses. Although *Pepinsky's Machine* is probably one of the easiest and most user-friendly programs presently available for calculating and examining a large number of such syntheses, the author feels that modern data collection practices will seriously limit its research applications. This is rather unfortunate given the wealth of structural information that can, under favourable circumstances, be obtained from even a few correctly phased strong reflections. Fig. 2 makes this point: the top panel shows a four-phase permutation map for a projection of a 4-α-helical bundle protein which was calculated before the crystal struc-
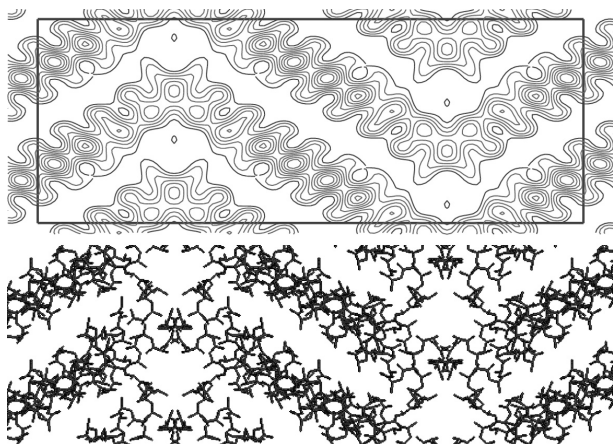

Fig. 2. Comparison of a permutation map (top panel) with the refined crystal structure. See text for details.

ture was determined. The second panel shows the polyalanine model corresponding to the final refined structure. Although one of the reflections turned out to have the wrong sign, the clarity and information content of the permutation map are still surprisingly high.

4. Program availability

The program has been developed and tested on the following Silicon Graphics computers: O$_2$ R5000 Irix 6.3, O$_2$ R10000 Irix 6.3, Indigo2 R10000 Irix 6.2, Origin 200 R10000 Irix 6.4 and Indy R4600 Irix 6.2. (Origin 200, Indigo2, O$_2$ and Irix are trademarks of Silicon Graphics Inc., Mountain View, CA; R5000 and R10000 are trademarks of MIPS Technologies Inc., Mountain View, CA). The program, together with its documentation and some exercises, is distributed free of charge, and is immediately available upon request from the author, or directly from the following World Wide Web address: http://Origin.imbb.forth.gr:8888/~glykos/.

References

Boyes-Watson, J. & Perutz, M. F. (1943). *Nature (London)*, **151**, 714–716.

Collaborative Computational Project, Number 4 (1994). *Acta Cryst.* D**50**, 760–763.

Glykos, N. M., Holzeburg, A. & Phillips, S. E. V. (1998*a*). *Acta Cryst.* D**54**, 215–225.

Glykos, N. M., Holzenburg, A. & Phillips, S. E. V. (1998*b*). *Acta Cryst.* D**54**, 707–707.

International Tables for X-ray Crystallography (1952). Vol. I, edited by N. F. M. Henry & K. Lonsdale. Birmingham: Kynoch Press.

Pepinsky, R. (1947). *J. Appl. Phys.* **18**, 601–607.

Woolfson, M. M. (1954). *Acta Cryst.* **7**, 65–67.

Wright, W. B. (1958*a*). *Acta Cryst.* **11**, 632–642.

Wright, W. B. (1958*b*). *Acta Cryst.* **11**, 642–653.