Sysabs - A program for the visualization of crystal data symmetry in reciprocal space

B. C. Taverner
Centre for Molecular Design, University of the Witwatersrand, Johannesburg, South Africa
Craig@hobbes.gh.wits.ac.za
http://www.gh.wits.ac.za/craig

Abstract

A computer program running under X-windows on any unix system has been developed for the visualization of symmetry in reciprocal space using reduced data from crystal measurements. The program was demonstrated using a test crystal of the inclusion compound of camphor(+) in thiourea, in the moderately high symmetry space group R32. The results indicate the usefulness of the visualization approach to both the determination of space group symmetry, and for the teaching of space group symmetry in reciprocal space to students. The program is freely available from the author, or at ftp://hobbes.gh.wits.ac.za/pub/sysabs.

1 Introduction

Students of crystallography often find the understanding of symmetry in reciprocal space unnecessarily complicated, primarily due to the conceptual gap between the apparent beauty and symmetry of the crystals themselves, and the group theory mathematics behind the symmetry of the unit cells in real and reciprocal space[3]. Currently there are a number of computer packages available for automatic determination of space group symmetry[1] from measured structure factor data[4,5]. These programs work quite well for most routine crystal structure solutions. However, they are not always successful, particularly in cases involving pseudo-symmetry. This has two implications. Firstly, the crystallographer must therefore still understand the concepts of space groups symmetry in reciprocal space in order to correctly determine the correct unit cell symmetry. In addition to this, in the training of crystallography students, a good understanding of reciprocal space symmetry needs to be developed. A common technique used is to teach students the use of photographic single crystal techniques[3]. Often, however, not only is access to such equipment limited, but also the bulk of the crystal structures solved will be done with other single crystal equipment that does not require the physical alignment of the crystal, like 4 circle and area detector diffractometers. For this reason, it is felt that the use of a program to visualize the diffraction data in an ordered way, with emphasis on the symmetry and absence conditions present, is of value to the crystallographic community for both routine work and education.

1.1 Test Crystals

A number of inclusion compounds of thiourea and selenourea have been studied with chiral and non-chiral guest molecules for the purpose of investigating host-guest intermolecular interactions with particular emphasis on chiral interactions and the relationship between the orientational disorder of the guest and the relative strengths of the host-host, host-guest and guest-guest interactions[2]. A number of thiourea and selenourea crystals have been reported in the literature, all with the apparent space group $R\bar{3}c$[2]. This is of relatively high symmetry, with rhombohedral centring, three-fold screw axes and glide planes. This results in a number of absence conditions, and presented an ideal situation for the use of visualization in the study of the details of the crystal symmetry. The use of chiral guest molecules in the inclusion complexes presented a particular challenge with regard to the degree to which the molecular chirality would cause the space group to deviate from centrosymmetric. A particular test data set, that of the crystal of thiourea and chiral camphor(+) was chosen to illustrate the use of the program sysabs.

2 Program Development

The program development started with an initial C program to read a standard shelx[4] hkl file and separate it into six files containing reflections that matched the six common sets used in absence condition identification in the International Tables[1]. These six are the h00, 0k0 and 00l lines and the hk0, 0kl and h0l planes in reciprocal space. Absence conditions can be determined in a primitive way by simply looking carefully through these files. It was obvious, however, that the relative intensities could far easier be interpreted with graphical representation. The initial graphical presentation was in the form of six graphs for the six sets simultaneously displayed on a PC VGA screen using the Turbo C VGA graphics libraries. This presented a dramatic improvement over the text based visualization, and the bar graphs for the line sets and the density graphs for the planes allowed for the visual identification of

pseudosymmetry, and the determination of likely alternative space groups in difficult crystal structures.
The next development involved the rewriting of the graphics code to run under the X windows system. Initially this was done solely to ensure that all our crystallographic software was running on the same platform. However, the X windows system allowed for some dramatic improvements over the DOS VGA version. The emphasis moved from the separation of the data sets into the six mentioned above, to the visualization of any six sets of lines and planes in reciprocal space. Each of the six could be individually displayed in separate control windows with zooming features and well as the ability to step through reciprocal space. Automatic determination of possible absence conditions was added to facilitate the identification of screw axis and glide plane translational symmetry. Online help for both command line mode and the X-windows interface was implemented.

3 Results and Discussion

Under normal operation, sysabs will start by loading a specified data set into memory, subject to optional command line parameters, split the data into selected subsets and then start an X11 interface window, into which a variety of graphs are drawn.

The functionality available will be covered be describing the program in sections, in each case referring to the crystal example described above.

3.1 Command line interface

The program can be started from the unix command prompt by typing the line:

```
sysabs datafile.hkl
```

where 'datafile.hkl' is the name of the reduced structure factor file to be loaded. If no filename is provided, sysabs will output the online help information that covers all command line options:

```
sysabs, aid to systematic absences by
        B.C.T. version 2.10 (4/96)

usage: "sysabs 'FileList' <+/-pgsf> <-o##.##>
        <-.*****> <-PABCIFRH> <+PABCIFRH>"

  FileList :   List of reflection data files to
               be scanned.
  <+/-xx>  :   Optional control command (any
               order)
  -p       :   Print h00, 0k0, 00l, hk0, 0kl, h0l
               to files
  +p       :   Do not print files
  -g       :   Draw graphs of selected hkl's
               (default).
  +g       :   Do not draw graphs.
  -s       :   Use Shelx format for output.
  +s       :   Do not use Shelx format for
               output (default).
  -f       :   Use free format for input data.
  +f       :   Use Shelx format for input data
               (default).
  -o##.##  :   Real number m=##.##, where peaks
               with I<m*(sig)I are ommitted.
  -.*****  :   Use the string ".*****" for
               deciding how to
               separate hkl's.  Default is
               .h00.0k0.00l.hk0.0kl.h0l
               The input string is used to
               modify the default.
               eg".h1.2k.11l..2" ->
               ".h10.2k0.11l.hk0.2kl.h0l"
  -PABCIFRH:   Show only reflections that should
               be absent as specified by the
               type of centring.
  +PABCIFRH:   Show only reflections that should
               be present as specified by the
               type of centring.
               eg. '-I' means show only
               reflections that break expected
               body centring.

NOTE :  By default only the graphs are drawn and
        no files are saved.
     :  Data in format "h k l  Int/Fobs
        sigI/sigF" expected.
        i.e. data ready for input to SHELX
        suitable.
     :  The output files will contain data in
        either the shelx format (-s) or a similar
        one with extra spaces (default).
     :  The files and options can be given in any
        order.
     :  Options apply to all files regardless of
        order.
     :  Obviously filenames cannot start with '-'
        or '+'
     :  Options given on the command line can be
        modified during the graph plotting
        section, where files can also be saved by
        clicking on the relevant graph
```

These options provide control over how the program is started and can be grouped into two categories, options affecting the overall state of the program, and options affecting the data separation and visualization. Options `+/-pgsf` fall into the first category, and the text output detailed above is sufficient explanation of their effect on the program. The remaining options, however, warrant examples. The option to control the space group centring condition may not be all that useful for the simple determination of the space group centring condition because most data collection programs will most likely have provided that information already. However, there often exists cases where there are alternative cell axis settings which can be clarified by further inspection. Even more interesting, are cases where there exists pseudo-symmetry, possibly relating to related crystalline phases. The degree to which the 'pseudo-symmetry' deviates from real symmetry can be explored with careful use of the -o (omit) option. An example of where this is relevant occurred with a low temperature study into the structure of the inclusion compound of thiourea with chiral camphor(+). The room temperature crystal had displayed R32 symmetry with pseudo $R\bar{3}c$ symmetry due to very high disorder on the chiral guest masking it's chiral nature. The apparent symmetry of the high quality low temperature data set was exactly the same, but isostructural refinements yielded an unsatisfactory fit (R~20%). The `-R` option displayed only those peaks that should have been absent in the rhombohedral symmetry. Use of the `-o` option with various settings demonstrated that these peaks, although considerably weaker than most peaks, were definitely worth further consideration, often having intensities near 100 times

their sigma's. In this case, the high quality of the data set allowed a more careful look into the pseudo symmetry of the crystal.

The -.***** option can be particularly useful in the systematic exploration of reciprocal space by displaying particular lines and slices in reciprocal space, and providing a variety of starting points for stepwise searches. Any centring conditions present can be clearly elucidated by stepping through reciprocal space slices to demonstrate their three dimensional nature. This option is available under the X-windows interface to sysabs.

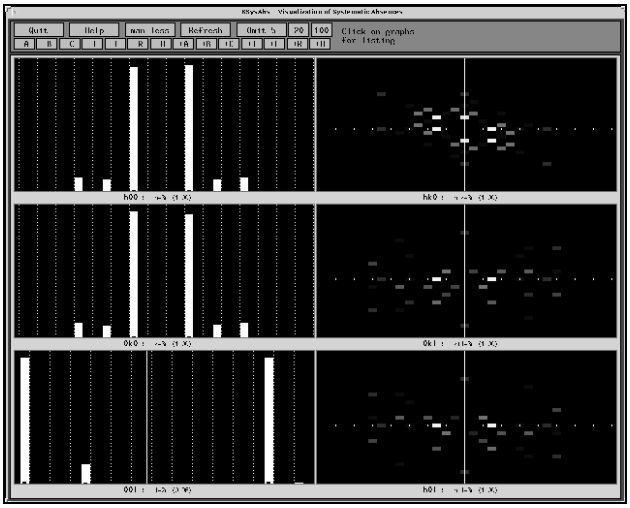


Figure 1 Initial X11 sysabs window. Option buttons and six standard graphs for test inclusion compound data are shown.

3.2 X-windows Interface

The X11 interface consists of a primary window containing two areas, a menu button area and a graph area as shown in figure 1. The graph area contains six sub-windows for the six simultaneous displays possible, each of which can contain either a bar graph for lines through reciprocal space, or a density plot of planes, as shown in figure 2. The menu button area contains the interactive features of sysabs. Help for these options is available by clicking on the button labelled "help". This writes a file called sysabs.hlp into the current directory for later viewing, and then also loads an xterm window displaying the contents of that file for immediate viewing. All interactive options are described in this online help:

```
      XSysAbs version 2.10

      by B.C. Taverner (4/96)

      A program for the visualization assisted
      identification of symmetry in reciprocal
      space.

Use the mouse to select menu options:

      "Quit"          - exit XSysAbs

      "Help"          - this help
      "Refresh"       - refresh window if
                        necessary
      "Omit 5"        - exclude reflections
                        below 5*sigma
      "20"            - exclude reflections
                        below 20*sigma
      "100"           - exclude reflections
                        below 100*sigma
      "-ABCIFRH"      - exclude reflections
                        that should be PRESENT
                        due to requested
                        centring.  For
                        example, in a C
                        centred space group,
                        "-C" should leave a
                        blank screen, while
                        "-A" or another option
                        will change the pattern
                        more dramatically.
      "+ABCIFRH"      - exclude reflections
                        that should be ABSENT
                        due to requested
                        centring.  For
                        example, in a C
                        centred space group,
                        +C should leave the
                        reflections unchanged,
                        while "-A" or another
                        option will change the
                        pattern more
                        dramatically.

      Click on any of the hkl graphs for a
      single window view, and then click on
      that for a listing of the relevant hkl
      values.  These values will be written to
      a file, and then displayed in a window
      using the common unix paginator "less"

      "man less" displays the relevant manual
      pages for the "less" command

Use the "q" command in "less" to exit this
window
```

Many of these options are the same as the command line options, and have been described in that section. The options to automatically save the data used in each graph to separate files are not directly available, but this is implicitly done when the user clicks in the closeup graphs. The paginator used to display both the help file above and the separate data set files defaults to the unix utility "less". Many unix systems, however, do not have this utility available, and so sysabs can be compiled to use the more common, but less functional paginator "more" instead.

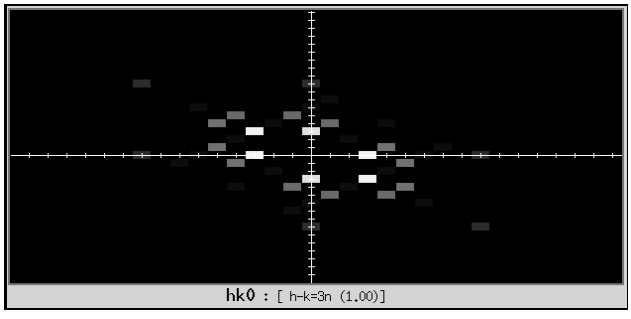


Figure 2 Single subwindow of hk0 plane with glide plane absence condition identified.

The more interesting features of the X11 interface can best be described through a worked example. The room temperature data for the inclusion compound of thioures and

camphor(+) was run through sysabs, providing the initial default graphs shown in figure 1. Figure 2 provides a closer view of the fourth of these default graphs with the automatic absence condition determination result shown at the bottom of the figure. This simple absence condition equation can be used to demonstrate how the glide plane and centring absence conditions effectively result in clear diagonal lines in the pattern. In this example the three-fold absence visible in all three displayed planes is due to the rhombohedral centring condition present. The planes in reciprocal space can be further studied by clicking on the graphs. This results in individual windows, with additional controls, as shown in figure 3.

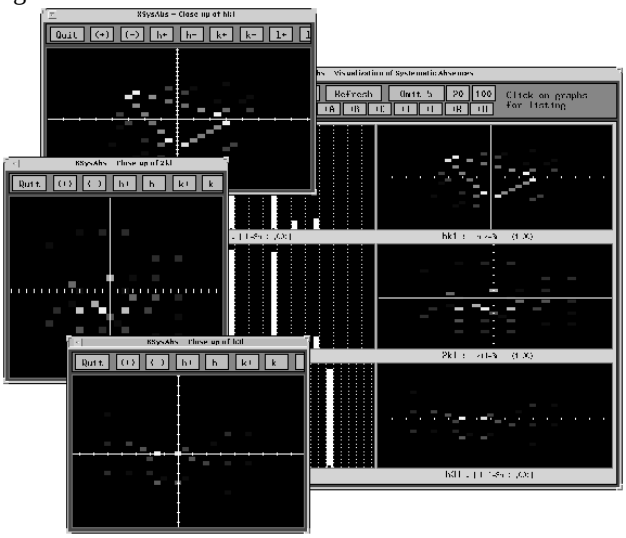


Figure 3 Individual control windows for graphs hk1, 2kl and h3l.

Each new window has a set of additional controls. The `(+)` and `(-)` buttons allow individual zooming in and out on the particular plane, while the `h+/-`, `k+/-` and `l+/-` buttons allow the visualization of parallel planes in reciprocal space. For example, clicking on `h+` in the `0kl` window will move it to `1kl`. The current range, however, is limited to 0 to 9 on any axis. In figure 3 the `hk0` graph has been shifted to `hkl`, the `0kl` to `2kl` and the `h0l` to `h3l`. Clicking in the individual graph windows results in the data represented being written

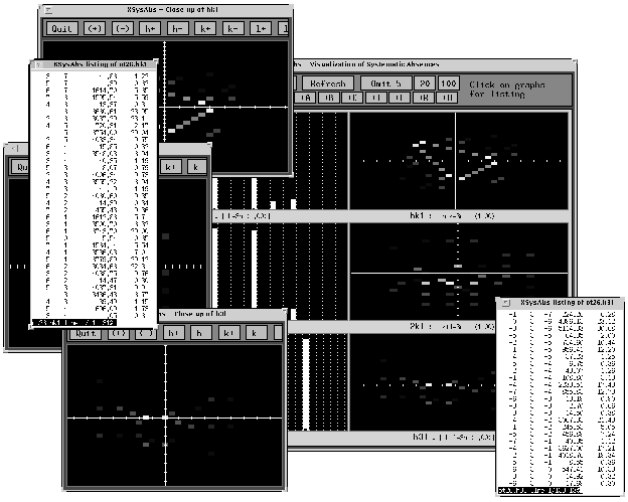
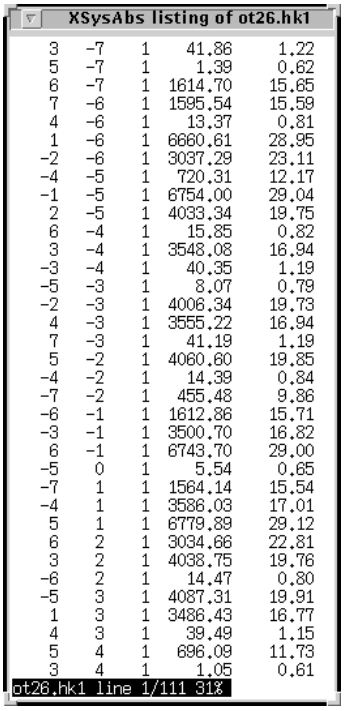


Figure 4 Raw data displays of hk1 and h3l planes

to disk and then displayed with a paginator, as shown in figures 4 and 5.



XSysAbs listing of ot26.hk1

3	-7	1	41.86	1.22
5	-7	1	1.39	0.62
6	-7	1	1614.70	15.65
7	-6	1	1595.54	15.59
4	-6	1	13.37	0.81
1	-6	1	6660.61	28.95
-2	-6	1	3037.29	23.11
-4	-5	1	720.31	12.17
-1	-5	1	6754.00	29.04
2	-5	1	4033.34	19.75
6	-4	1	15.85	0.82
3	-4	1	3548.08	16.94
-3	-4	1	40.35	1.19
-5	-3	1	8.07	0.79
-2	-3	1	4006.34	19.73
4	-3	1	3555.22	16.94
7	-3	1	41.19	1.19
5	-2	1	4060.60	19.85
-4	-2	1	14.39	0.84
-7	-2	1	455.48	9.86
-6	-1	1	1612.86	15.71
-3	-1	1	3500.70	16.82
6	-1	1	6743.70	29.00
-5	0	1	5.54	0.65
-7	1	1	1564.14	15.54
-4	1	1	3586.03	17.01
5	1	1	6779.89	29.12
6	2	1	3034.66	22.81
3	2	1	4038.75	19.76
-6	2	1	14.47	0.80
-5	3	1	4087.31	19.91
1	3	1	3486.43	16.77
4	3	1	39.49	1.15
5	4	1	696.09	11.73
3	4	1	1.05	0.61

ot26.hk1 line 1/111 31%

Figure 5 Data display of plane hk1

Although this data can be useful in determining just how weak the weak reflections are, and whether or not they are absent, the use of the omit option is substatially easier. The X-windows interface, however, only allows the omission of reflections below 5σ, 20σ and 100σ. For finer control, it is necessary to provide the omit parameter on the command line where fractional values of any magnitude are allowed.

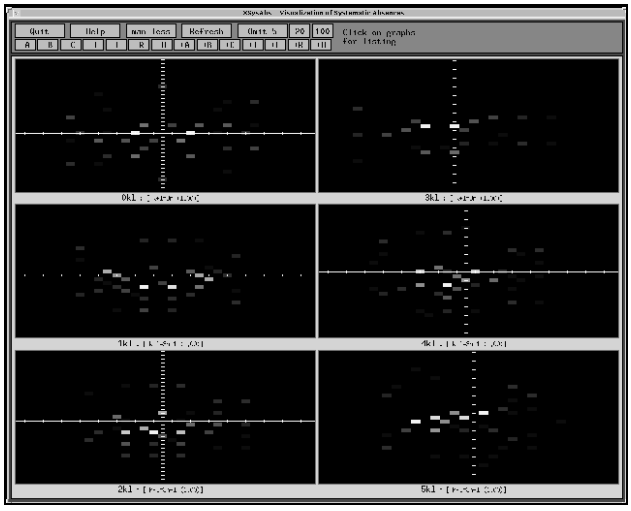


Figure 6 Six slices from 1kl to 5kl showing rhombohedral centring.

One command line feature which has a marked impact in the X-windows interface is the -.***** option described in the

previous section. This option allows the user to specify the starting six data sets to be displayed graphically. Any six sets of lines or planes can be displayed instead of the default:

```
-.h00.0k0.00l.hk0.0kl.h0l
```

If, for example, the rhombohedral centring condition was of particular interest, the line:

```
-.0kl.1kl.2kl.3kl.4kl.5kl
```

places these six parallel planes in reciprocal space into the six graphs in the X-windows window. In each case, the apparent absence conditions have were identified. As can be seen in figures 6 and 7, there were three fold absence conditions in all six planes, with k+l=3n-1 for all but those where h is a multiple of 3, in which case the condition is k+l=3n. It takes only a moment to realise that this is simply the projection of the rhombohedral centring absence condition -h+k+l=3n onto these planes. Although this demonstration is certainly not the quickest way of determining the centring condition, it is definitely a very useful demonstration for educational purposes.

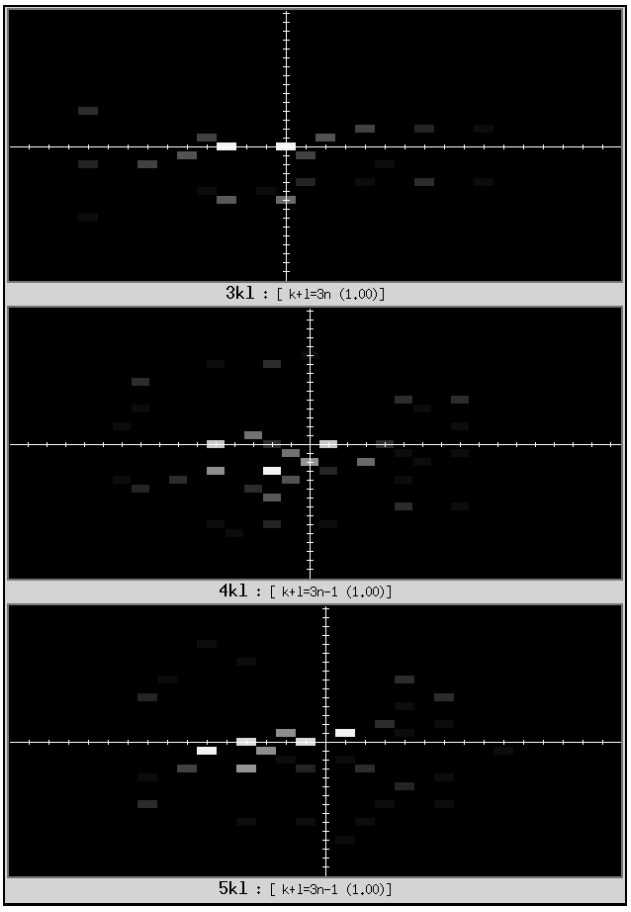


Figure 7 Closeup of planes 3kl, 4kl and 5kl with absence condition matching -h+k+l=3n

A final feature of the X-windows interface is entirely due to the X11 system, and that is the flexibility of use of multiple windows, as shown in figure 8. The ability to have a number of independent windows controlled by the same program and displaying the same, or related data, provides a level of power far beyond that available under DOS. This is demonstrated in figure 3 where the individual windows show the exact same graph as their partner in the set of six. As one zooms the individual window, so does the partner zoom. A further advantage to X11 that can be utilized by sysabs is its network capability. One server can contain all the data and the programs, while the actual display can be done on a number of remote machines. This is very suitable for situations either where central control is required to protect data, or where the main computational crystallographic facilities are centralized, as would be the case where all data is stored on a central supercomputer and users have access to a number of smaller terminals supporting X11.

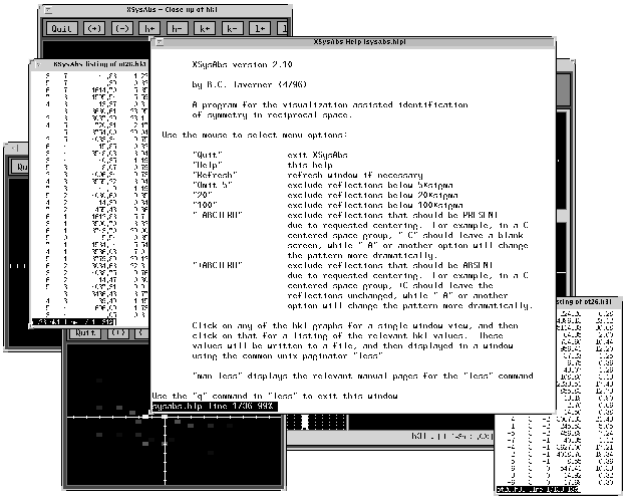


Figure 8 Multi-window features of X11 and sysabs, with on-line help in foreground.

4 Distribution and Installation

Sysabs can be obtained via anonymous ftp from ftp://hobbes.gh.wits.ac.za/pub/sysabs or by emailing the author at Craig@hobbes.gh.wits.ac.za. It is provided as portable ANSI C source code which should compile without much difficulty under most unix platforms and platforms that support X11. Use is made of the base X11 library only, which should be compatible across all X11 systems. Although the system was developed under X11R6 (under Linux), it should also work under releases 4 and 5, but this has not been extensively tested. Binaries are provided for Linux and Irix5.2. The distribution includes installation instructions and porting hints, as well as a copyright and disclaimer.

5 Summary and Conclusions

The program sysabs was successfully developed under Linux for generic unix/X platforms. It can be used to read reduced data in shelx[4], or similar, hkl format and plot a wide range of bar and density graphs of lines and planes in reciprocal space. The additional facilities to match correct centring

conditions, screw axes and glide planes greatly facilitate the visualization of translational symmetry in reciprocal space and the determination of the correct space group symmetry. The command line options allow for considerable configuration of the program behaviour to suite the needs of the crystallographer. On-line help is available in both command line and X-windows interfaces. The program is freely available from the author, and can also be obtained by ftp in source form, or in binary form for Linux and Irix5.2 at ftp://hobbes.gh.wits.ac.za/pub/sysabs.

References

[1] T. Hahn, International Tables for Crystallography, Volume A - Space-Group Symmetry, Fourth Edition, Kluwer Academic Publishers, 1995.

[2] B. C. Taverner, "Disorder in Inclusion Compounds of Thiourea and Selenourea," Masters Dissertation under the supervision of D. C. Levendis, Centre for Molecular Design, University of the Witwatersrand, South Africa, 1994.

[3] B. C. Taverner & L. M. Cook, "Introduction to Crystallography", one day workshop at fourth year level, Chemistry Department, University of the Witwatersrand, South Africa, 1994 & 1995.

[4] Shelx programs:
G. M. Sheldrick, Acta Crystallogr., A46 (1990) 467-473.
G. M. Sheldrick, J. Appl. Cryst., (1993) in preparation.
G. M. Sheldrick, Shelxtl - Siemens, 1994.

[6] NRCVAX - An Interactive Program System for Structure Analysis J. Appl. Crystl 22, 384, 1989.