

reduce to a minimum the number of conditional statements. The result would be the most efficient program of the type used for the particular job, wasting no computer space. This program is of course discarded at the end of the job.

Objections may be raised because the Editor will need to be run on a large computer. However this means that a large centralized computer can be used to produce programs which would then fit into smaller local computers. These would then do the time consuming work. The Editor should be capable of writing the program in any desired dialect of the language, and even of inserting machine coded sections in the few places where the gain is worth the trouble. In least-squares refinement the time consuming part is in setting up the matrix, then in its inversion. Machine coding the former for any computer should not be too difficult, and the latter should be locally available as a fast symmetric matrix inverter is an asset to any computing laboratory.

The proposed system will thus not take all the work from the local computer, but will give much more work to it. This is in contrast to the *X-ray 67* type of system

which runs a job, however small, on a computer with great capacity. The Editor system therefore provides a means of using our computing facilities to the maximum capacity with increased flexibility. The programming effort to develop this system would be very large indeed, but the system does merit much more thought and consideration than has been given to it.

#### References

- DUCKWORTH, J. A. K., WILLIS, B. T. M. & PAWLEY, G. S. (1969). *Acta Cryst.* **A25**, 482.  
 ENEMARK, J. H. & IBERS, J. A. (1967). *Inorg. Chem.* **6**, 1575.  
 GOLDSTEIN, H. (1959). *Classical Mechanics*. Reading, Mass.: Addison-Wesley.  
 HAMILTON, W. C. (1965). *Acta Cryst.* **18**, 502.  
 LA PLACA, S. J. & IBERS, J. A. (1965). *Acta Cryst.* **18**, 511.  
 PAWLEY, G. S. (1966). *Acta Cryst.* **20**, 631.  
 PAWLEY, G. S. (1969). In *Thermal Neutron Diffraction*. Ed. B. T. M. WILLIS. Oxford Univ. Press.  
 PAWLEY, G. S. & YEATS, E. A. (1969). *Acta Cryst.* **B**. In the press.  
 SCHERINGER, C. (1963). *Acta Cryst.* **16**, 546.  
 STEWART, J. M., KUNDELL, F. A. & CHASTAIN, R. V. (1967). *X-ray 67*, Univ. of Maryland.

*Acta Cryst.* (1969). **A25**, 535

## A Simple Method for Drawing Molecules using a Digital Plotter

BY A. J. COLE AND P. G. ADAMSON

*University of St. Andrews Computing Laboratory, St. Andrews, Scotland*

(Received 25 October 1968)

A simple fast method is described for producing three dimensional 'ball-and-stick' representations of molecules, using a digital plotter; any viewpoint at a reasonable distance away is suitable.

### Introduction

The aim is to simplify the computing and drawing processes by reducing the picture to circles (representing 'spherical' atoms), and bonds consisting of tapered bundles of straight lines.

The method has to take into account two possible kinds of overlap in the picture; overlap of circles, and overlap of circles across bonds. Overlap of bonds across bonds is not taken into account, since they are drawn completely black; slight gaps could be left in a bond when it passes behind another; but the accuracy of the plotter (0.01") does not allow a neat enough split to be made. Clarity is not very much impaired by this omission. The bonds themselves present difficulties in balancing plotter limitations (a pen step can only be made in one of eight directions) against a reasonably pleasant result. The bonds are given a forced taper, since perspective alone does not provide sufficient depth of view, unless a very close viewpoint is taken; in which case the sizes of the circles representing

the atoms vary so much as to lose any meaning or identification purpose they may have.

The taper was chosen on an arbitrary basis, subject to plotter considerations; no taper is provided on bonds between atoms whose distances from the viewer are similar (again an arbitrary distinction is made).

The main data required for this method are: the angles between the coordinate axes (if not 90°), the viewpoint coordinates, and for each atom, a name (up to four characters), the coordinates and the radius of the sphere representing the atom and the names of all the other atoms connected to it. All coordinates are given in Å. A facility is provided so that all atoms of one type, say those beginning with C, can be allotted the same radius without repeating this information for each atom individually.

In the interests of simplicity, no labelling of the picture is provided for, since a general label facility becomes complicated when some atoms are partially or even wholly obscured.

### Data manipulation

The main data consist of the names, positions, sizes and connexions to the atoms constituting the molecule. The names are stored in an array, and the atoms are referred to internally by numbers in the order in which the names appear in the list – the same as the order in which the atom data is read.

For any particular atom, its size, coordinates and connections are all read in on one card. If no card with the word 'SIZES' was encountered before the atom data, a radius is also looked for; if there was such a card, the correct radius is found from the SIZES list. The SIZES card consists of a list: atom type, radius, atom type, radius, . . .

For example,

SIZES C 0.25 , N 0.5

means that all atoms whose names begin with C are allotted a radius of 0.25 Å, all those beginning with N, 0.5 Å, and so on. This method is useful for a naming system such as C1, C2, . . . , N1, N2, . . . ; but since Cl3 (for instance) also begins with a C, care has to be taken to put the SIZES list in the correct order, with C before Cl.

When the name of a connected atom is encountered, it is searched for in the list of names compiled so far;

if it is in the list, an entry is made in the appropriate place in a triangular connexion matrix; if it is not, then the particular atom has not been met previously, and no entry is made. Although the method of stating all the atoms connected to each one duplicates the information, it simplifies the handling of this data, as does the use of a connexion matrix, most of whose entries are blank.

Information as to the angles between the axes, the viewpoint, whether wide or narrow paper is to be used on the plotter and the required scale factor relative to 1" to 1 Å is also required.

An example of the complete data set for cycloheptane (C<sub>7</sub>H<sub>14</sub>) is given in Table 1. This data is used in the three Figures which show the same molecule from different viewpoints.

Various error conditions, such as two atoms with the same name, as well as syntax errors, are tested for and suitable error messages are printed.

The next stage is to make a transformation from oblique coordinates to rectangular coordinates, and then to move the origin of coordinates to the 'centre' of the molecule which is taken to be the average of the extreme points in the *x*, *y* and *z* directions. The axes are then rotated until the *x* axis passes through the viewpoint chosen. The rotation is still not unique, however, and it is made so by confining the new *y* axis to the old *xy* plane.

Table 1. Data specifications for cycloheptane

Angles	090	090	090					Angles between coordinate axes: assumed to be 90°, 90°, 90° if this card is omitted.	
Scale	1.0							Scale factor, in inches per Å: assumed to be 1 if this card is omitted.	
Sizes	C 0.3	H 0.2						This card specifies radii of atom types, <i>viz.</i> all atoms beginning with C to have a radius of 0.3 Å, and all H to have 0.2 Å radius	
C1	0.0000	0.0000	0.0000	C2	C7	HA1	HB1	These cards specify:	
HA1	-0.6477	0.8989	-0.0001	C1					
HB1	-0.6478	-0.8988	-0.0002	C1					
C2	0.8461	0.0000	-1.2783	C3	C1	HA2	HB2		(i) the name of an atom,
HA2	1.2961	1.0039	-1.4092	C2					(ii) the <i>x, y, z</i> coordinates (in Å) of the atom,
HB2	0.1855	-0.2036	-2.1443	C2					(iii) (not required or used here) the radius of the atom; if a SIZES card has been met a radius here will override the general specification,
C3	1.9671	-1.0453	-1.2546	C4	C2	HA3	HB3		(iv) the names of <i>all</i> the other atoms connected to the atom.
HA3	1.5279	-2.0354	-1.0211	C3					
HB3	2.4303	-1.0957	-2.2599	C3					
C4	3.0568	-0.7318	-0.2229	C5	C3	HA4	HB4		
HA4	2.9035	-1.3753	0.6658	C4					
HB4	4.0451	-0.9713	-0.6628	C4					
C5	3.0558	0.7350	0.2222	C6	C4	HA5	HB5		
HA5	4.0440	0.9760	0.6616	C5					
HB5	2.9011	1.3784	-0.6664	C5					
C6	1.9662	1.0469	1.2545	C7	C5	HA6	HB6		
HA6	2.4299	1.0980	2.2595	C6					
HB6	1.5254	2.0363	1.0212	C6					
C7	0.8468	0.0000	1.2788	C1	C6	HA7	HB7		
HA7	0.1866	0.2024	2.1453	C7					
HB7	1.2985	-1.0034	1.4089	C7					
View	025	000	000					These cards specify the view-point of observation, with <i>x, y, z</i> coordinates in Å. Any number of these can be included.	
View	025	025	025						
View	000	025	000						

This method of rotation breaks down in the case where the viewer looks directly down the  $z$  axis of the rectangular frame, and considerable error occurs when the viewpoint is just off this axis. To avoid this, a test is made to see if the line from the viewpoint to the new origin, which is the proposed new  $x$  axis, is within  $45^\circ$  of the  $z$  axis; if it is, the rotation is modified so that the new  $z$  axis is confined to the old  $xz$  plane instead of the  $y$  axis to the  $xy$  plane. In either case, a conical projection is made onto the new  $yz$  plane as picture-plane, so that we now have a set of plane coordinates of the atom centres, together with the atom 'radii', which are also modified in the projection to take into account varying distance from the viewer, compared with the picture-plane.

For oblique coordinates with conventional angles  $\alpha, \beta, \gamma$  between the axes ( $\alpha$  between  $y$  and  $z$  axes, *etc.*), a transformation is made to rectangular coordinates, keeping the same  $x$  axis, and confining the new  $y$  axis to the old  $xy$  plane.

$$\begin{aligned}x' &= x + y \cos \gamma + z \cos \beta, \\y' &= y \sin \gamma + z[(\cos \alpha - \cos \beta \cos \gamma) / \sin \gamma], \\z' &= z / [\sin^2 \beta - \{(\cos \alpha - \cos \beta \cos \gamma) / \sin \gamma\}^2].\end{aligned}$$

For rotation of axes until the  $x$  axis passes through the viewpoint where  $x'_v, y'_v, z'_v$  are in rectangular coordinates, we use

$$x'' = \frac{x'_v \cdot x' + y'_v \cdot y' + z'_v \cdot z'}{\sqrt{[x'_v{}^2 + y'_v{}^2 + z'_v{}^2]}},$$

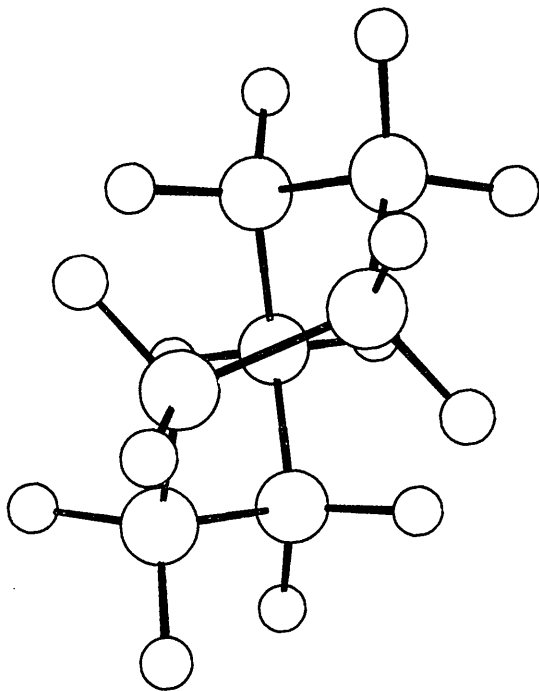


Fig. 1. Cycloheptane molecule with data specified in Table 1 viewed from the point (25,0,0).

$$\begin{aligned}y'' &= \frac{-y'_v \cdot x' + x'_v \cdot y'}{\sqrt{[x'_v{}^2 + y'_v{}^2]}}, \\z'' &= \frac{-x'_v z'_v \cdot x' - y'_v z'_v \cdot y' + (x'_v{}^2 + y'_v{}^2) \cdot z'}{\sqrt{[x'_v{}^2 + y'_v{}^2]} \sqrt{[x'_v{}^2 + y'_v{}^2 + z'_v{}^2]}}.\end{aligned}$$

This is the case where the  $y''$  axis is confined to the  $x'y'$  plane; the other case, with the  $z''$  axis confined to the  $x'z'$  plane, uses the above three equations, but with  $y'$  and  $z'$  interchanged throughout.

The coordinate limits of the projected 'picture' are now found and are tested to see if the proposed scale factor will allow the picture to fit on the plotter paper. If not, a suitable factor is calculated and a message is typed to enquire if the new factor will serve. If no reply is given within twenty seconds the new factor is assumed to be satisfactory.

The distances of the atoms from the viewpoint are now arranged according to magnitude, using a very efficient method (Scowen, 1965). The numbers allotted to the atoms as the data was read in are also arranged at the same time. The arranged list is referred to as NODIST.

#### Drawing the atoms

Starting with the first read in, an atom is taken and the circle representing its plane image is examined for intersection with other atom 'circles'. This operation is greatly speeded up by, firstly, examining only an atom nearer to the viewer than the current one, that is earlier in the NODIST list and secondly, by making a coarse test involving only the circumscribed squares before any fine test is made on the distance between the centres of the circles representing the atoms. Thus, if the circumscribed squares (with sides parallel to the axes) do not intersect, no fine test is made. The squares test involves only simple inequalities on the circle centres and radii.

There are four cases of intersection of a 'nearer' circle with the current one. If no intersection occurs, and the centre of each one is outside the other, then the 'nearer' one is ignored, and the next one in the NODIST list is taken. If no intersection occurs, and the centre of one is inside the other, then either the nearer one appears to be inside the current one, in which case it has no effect on the boundary, or else it completely obscures the current one, when a message is printed pointing out the fact, the current atom is ignored, and the next one is taken. If intersection occurs, the centre and radius of the intersecting circle is put in a short list of 'intersecting circles'.

When all the nearer atoms have been considered, and a list of intersecting circles, if any, has been made, the plotter pen is instructed to go to the appropriate position for drawing the circle to represent the current atom.

The circle is drawn by moving the pen one step at a time, choosing the step direction, which is limited to one of eight possibilities, so that the required dis-

tance from the centre is maintained as closely as possible; and also at each step, testing whether the pen will be taken within any of the intersecting circles if such exist. If the pen is already down and the next step takes it inside one of the intersecting circles, then it is raised and the fact that it is up is noted so that it is not raised again on subsequent steps. Similarly, if it is already up, and the next step takes it into a visible arc, it is lowered, and a note made. When there is more than one intersecting circle in the list at each new step a distance test is made on the domain of each one until either the list is exhausted or it is found that the next step trespasses inside a particular circle. This unsophisticated method is, in practice, very fast.

When the circle is finished, the next atom in the list as read in is taken as the current one, and the above drawing procedure is repeated.

### Bonds

After all the atom circles have been drawn, the bonds are dealt with: the connexions to be made are found from the triangular connexion matrix constructed when the atom data was read in. The  $(n-1)$ th row of the matrix indicates the connexions to the  $n$ th atom and has  $n-1$  locations, any of which can be 0 or 1, a 1 in the  $i$ th position indicating a bond to the  $i$ th atom.

When any two atoms are to be joined, a test is first made to see which is further away so as to set up a suitable taper for the bond. If they are approximately equidistant from the viewer, no taper is provided. The limits of the bond-ends are defined, for the nearer atom, as the edge of the circle, and for the further atom, as the edge of an internal circle whose radius is determined by the atom radius and the angle at which the bond is sloped away from the viewer. Various examples are shown in the Figures, including cases in which bonds are nearly perpendicular to the line of vision.

The bond line is now tested against other atoms possibly intersecting it; only atom circles which are 'nearer' to the viewpoint than the further of the two to be connected can intersect with the bond so as to hide it. An intersecting circle is always assumed to *obscure* the bond; it could happen that an atom was nearer to the viewer than the further atom to be connected and intersected with the bond but did not obscure it; however, examination of this case shows a molecule containing such a bond to be rather unlikely.

A coarse test is made first, using the rectangle with corners at the bond ends and sides parallel to the axes, and the square with sides parallel to the axes circumscribing the atom-circle; a finer test is made if it is found that the square and the rectangle intersect. If the atom circle is found to intersect with the bond its centre and radius are put in a list similar to that used for the circle-drawing procedure. If the bond is found to be completely obscured by the atoms to be joined, it is ignored.

The basic bond line is then drawn; tests being made at each plotter step for bond end-limits and intersection with other atoms if any.

A taper is introduced by drawing two more lines outside the basic bond, with the taper towards the 'further' atom-circle. The amount of taper is determined somewhat arbitrarily by the slope of the bond away from the viewer, in a similar way to the bond-end determination.

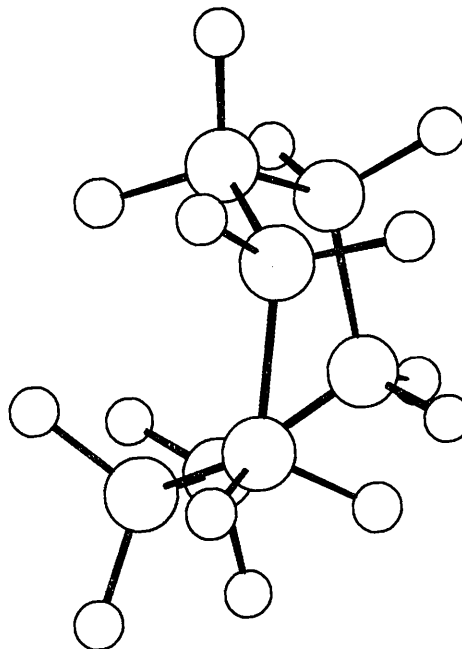


Fig. 2. As for Fig. 1 but with viewpoint (25,25,25).

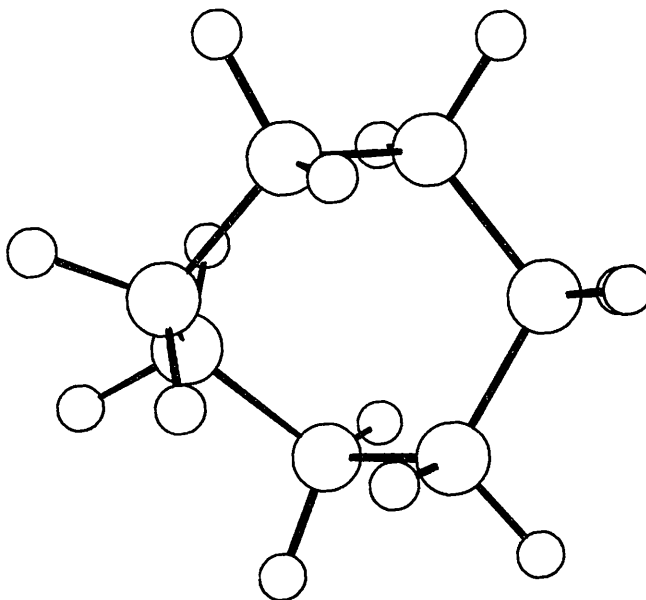


Fig. 3. As for Fig. 1 but with viewpoint (0,25,0).

The bond 'framework' is then filled in with two more lines. The visible ends of the bond are improved slightly by letting the middle three lines forming it run over the determined end limit by one plotter step thus providing an approximation to a cylindrical end to the bond where it cuts the spherical atom 'surface'.

The next bond is then found from the connexion matrix and so on. Figs. 1 to 3 show completed diagrams of the molecule cycloheptane ( $C_7H_{14}$ ) from different viewpoints.

A program has been written in the symbolic programming language for the IBM 1620 Model II computer with 60K store using a 1627 Model II plotter (which operates at up to 200 steps per second with a step length of 0.01"). Since there is a fair amount of calculation to perform per step when drawing circles and bonds, the plotter does not work at full speed. However, an average molecule of 20 or 30 atoms, drawn 1" to 1 Å takes about eight or nine minutes to draw, which is still a considerable saving on a day's manual labour. Note that computation time is equivalent to between 3 and 4 seconds on an Atlas. It is

intended to rewrite the program for the IBM 360/44. Flow diagrams can be obtained from the authors at the St. Andrews University Computing Laboratory.

A fairly free format can be used for the atom data input and up to 200 atoms can be accommodated in one molecule. Several views can be drawn successively without reloading the atom data and stereoscopic pairs can be drawn by choosing suitable viewpoints. Any particular bond-line can be drawn as just a single-line (e.g. for 'connections' between molecules) by following a connected atom name (on the atom data cards) by an asterisk.

We are indebted to several members of staff in St. Andrews and Dundee Universities, who provided *real* molecules for testing the program.

#### Reference

SCOWEN, R. S. (1965). *Collected Algorithms from CACM: Algorithm 271 Quicksort*. National Physical Laboratory, Teddington, Middlesex, England.

*Acta Cryst.* (1969). A25, 539

## On the Application of the Symbolic Addition Procedure in Neutron Diffraction Structure Determination

BY S. K. SIKKA\*

*Atomic Energy Research Establishment, Harwell, England*

(Received 31 October 1968)

The conditions have been examined for the application of the symbolic addition procedure for the evaluation of signs of structure factors from neutron diffraction data for centrosymmetric crystals containing atoms with both positive and negative scattering factors. It is shown that this procedure can be used directly, without first deriving the squared structure factors, when the contribution of the negative scatterers to the total neutron scattering is less than 25%. Several examples of the use of this method are given. The results show that for  $E_n \geq 1.5$ , the signs of about 95% of the reflexions are determined correctly.

### 1. Introduction

It is well known that direct methods of solving the phase problem by the symbolic addition procedure (Karle & Karle, 1966) cannot generally be applied to neutron diffraction if the crystal under examination contains some negative scatterers of neutrons, such as hydrogen atoms, together with other positive scatterers. This is because the basic assumption for the use of direct methods, that the scattering density is always a positive quantity, is violated. To circumvent this difficulty, Karle (1966) proposed the squared-structure

approach. The squared structure is defined as one having atoms with scattering amplitudes equal to the square of the normal amplitudes, but situated at the original positions in the unit cell. Such a hypothetical crystal has positive neutron scattering density everywhere and thus the phase problems can be solved by the symbolic addition procedure. The formula for calculating the magnitudes of squared structure factors from the observed neutron intensities was also given by Karle (1966). The success of this approach was demonstrated by Ellison & Levy (1967), who used it to solve the structure of glycolic acid. Hamilton (1966) has pointed out that this approach does not utilize all the information in the data and also that the method is really not direct since one has still to use chemical

\* On leave from Nuclear Physics Division, Bhabha Atomic Research Centre, Trombay, Bombay 85, India. (Now returned.)