

$$\Gamma = \begin{pmatrix} -0.037176 & 0.172559 & -0.984297 \\ 0.846866 & 0.528338 & 0.060639 \\ 0.530506 & -0.831313 & -0.165776 \end{pmatrix}.$$

A maximum-likelihood estimate of the mean orientation \mathbf{M} in the matrix Fisher distribution is then $\hat{\mathbf{M}} = \Delta\Gamma$, which expressed in Euler angles is $(\varphi_1, \phi, \varphi_2) = (171.598^\circ, 2.169^\circ, 101.824^\circ)$. This mean orientation is of course different from the result that would have been obtained by simply taking the arithmetic mean of the Euler angles. Such a simple approach to orientation averaging is not recommended. By the approach of Wood (1993) and with an approximation described by Mardia & Zemroch (1977), the concentration parameter is found as $\mathbf{D}_\phi \approx \text{diag}(5.311 \times 10^4, 3.156 \times 10^4, 2.615 \times 10^4)$. The large values for ϕ_i indicate, as expected, a large clustering of the \mathbf{X} 's.

Consider the problem of testing whether the \mathbf{X} 's can be assumed to be uniformly distributed on $\text{SO}(3)$. In this case, the answer is clearly no but a formal test for the hypothesis of uniformity can be obtained using the statistic R defined by (18). For the data of Table 2, $R = 3 \times 10 \times 2.99953 = 89.986$, which should be compared to χ^2_5 . The hypothesis is clearly rejected.

The authors express their appreciation to the referees for their careful reading of the manuscript and helpful suggestions.

References

ALTMANN, S. L. (1986). *Rotations, Quaternions and Double Groups*. Oxford: Clarendon Press.

- BINGHAM, C. (1974). *Ann. Stat.* **2**, 1201–1225.
 BINGHAM, C., CHANG, T. & RICHARDS, D. (1992). *J. Multivar. Anal.* **41**, 314–337.
 BUNGE, H. J. (1993). Editor. *Proc. Microscale Textures of Materials. Textures Microstruct.* **20**, 1–242.
 BUNGE, H. J. (1994). Editor. *Proc. ICOTOM10*. In the press.
 CHANG, T. (1986). *Ann. Stat.* **14**, 907–924.
 CHANG, T. (1987). *J. Geophys. Res.* **B92**, 6319–6329.
 CHANG, T. (1989). *Ann. Stat.* **17**, 293–306.
 CHANG, T. (1993). *Int. Stat. Rev.* **61**, 299–316.
 DOWNS, T. D. (1972). *Biometrika*, **59**, 665–676.
 FISHER, N. I., LEWIS, T. & EMBLETON, B. J. J. (1987). *Statistical Analysis of Spherical Data*. Cambridge Univ. Press.
 FISHER, R. A. (1953). *Proc. R. Soc. London Ser. A*, **217**, 295–305.
 JUPP, P. E. & MARDIA, K. V. (1989). *Int. Stat. Rev.* **57**, 261–294.
 KHATRI, C. G. & MARDIA, K. V. (1977). *J. R. Stat. Soc. B*, **39**, 95–106.
 MACKENZIE, J. K. (1957). *Acta Cryst.* **10**, 61–62.
 MARDIA, K. V. (1972). *Statistics for Directional Data*. London: Academic Press.
 MARDIA, K. V. & ZEMROCH, P. J. (1977). *J. Stat. Comput. Simul.* **6**, 29–34.
 MORAN, P. A. P. (1976). *Perspectives in Probability and Statistics*, edited by J. GANI, pp. 295–301. Orlando: Academic Press.
 PRENTICE, M. J. (1978). *Ann. Stat.* **6**, 169–176.
 PRENTICE, M. J. (1984). *Biometrika*, **71**, 147–154.
 PRENTICE, M. J. (1986). *J. R. Stat. Soc. B*, **48**, 214–222.
 PRENTICE, M. J. (1989). *J. R. Stat. Soc. B*, **51**, 241–248.
 PRESS, W. H., FLANNERY, B. P., TEUKOLSKY, S. A. & VETTERLING, W. T. (1988). *Numerical Recipes: the Art of Scientific Computing (C version)*. Cambridge Univ. Press.
 RANDLE, V. (1992). *Microtexture Determination and its Applications*. London: The Institute of Materials.
 RIVEST, L. (1989). *Ann. Stat.* **17**, 307–317.
 STEPHENS, M. A. (1979). *Biometrika*, **66**, 41–48.
 WATSON, G. S. (1983). *Statistics on Spheres*. New York: Wiley.
 WATSON, G. S. (1989). *Probability Measures on Groups IX, Lecture Notes in Mathematics*, edited by H. HEYER, pp. 398–413. Berlin: Springer-Verlag.
 WOOD, A. T. A. (1993). *Aust. J. Stat.* **35**, 69–79.
 WRIGHT, S. I. & ADAMS, R. L. (1992). *Metall. Trans.* **A23**, 759–767.

Acta Cryst. (1994). **A50**, 748–753

STEP – a Trial-and-Error Procedure for Crystal Structure Determination. I. A Description of the Program System SYSTEM90

BY HOU YONGGENG, GAO MIN, LI LIPU AND HOU PEIXIN

Institute of Chemistry, Academia Sinica, Beijing 10080, People's Republic of China

(Received 19 January 1994; accepted 25 April 1994)

Abstract

This paper describes a new trial-and-error direct-methods procedure called *STEP*. A set of strong reflections, sufficient to solve the structure, is divided into a hierarchy of smaller soluble and connected subsystems. Within each subsystem, the reflections

are required to be well connected with each other, given that the phases of all reflections in the previous subsystems are known. A trial-and-error procedure is then employed to provide an approximate solution to an overdetermined set of equations. Subsequently, phases are refined by one of two available tangent formulae and then assessed for plausibility by figures

of merit. A new overall figure of merit, XDFOM, has been found to be very effective in picking up a correct solution and the procedure is stopped when XDFOM exceeds 4.0. *STEP* is incorporated in a package called *SYSTEM90*, which has been tested with four difficult known structures and one difficult unknown structure. In all cases, a correct phase set was found. *SYSTEM90* is indicated as a powerful technique for the solution of difficult and large structures.

Introduction

The tangent formula, first derived by Karle & Hauptman (1956), has become the basis of many direct-methods computer packages, for example, *MULTAN*, *SHELX*, *SIMPEL*, *MITHRIL* and *SAPI*, which routinely solve the majority of small to medium-size structures and can even be effective for small proteins. Nevertheless, some smaller structures can still be difficult to determine and we describe here another direct-methods approach, the systematic trial-and-error procedure (*STEP*), which has been successful where other methods have failed.

Setting up systems of reflections

A number of strong reflections are chosen, sufficient to solve the structure if their phases are known. These are then subjected to the *CONVERGENCE* procedure in *MULTAN*, which enables them to be ordered in such a way that, from a set of known phases – those defining the origin and enantiomorph and Σ_1 phases, for example – new phases can be determined by means of the tangent formula. However, for *STEP*, the *CONVERGENCE* procedure is modified so that a reflection is only accepted in the *CONVERGENCE* mapping if for that reflection $\langle \alpha(\mathbf{h}) \rangle \geq \alpha_p$, where

$$\begin{aligned} \langle \alpha(\mathbf{h}) \rangle^2 = & \left(\sum_{\mathbf{k}} \kappa(\mathbf{h}, \mathbf{k}) \{ I_1[\kappa(\mathbf{h}, \mathbf{k})] / I_0[\kappa(\mathbf{h}, \mathbf{k})] \} \right)^2 \\ & + \sum_{\mathbf{k}} \kappa(\mathbf{h}, \mathbf{k})^2 (1 - \{ I_1[\kappa(\mathbf{h}, \mathbf{k})]^2 / I_0[\kappa(\mathbf{h}, \mathbf{k})]^2 \}) \end{aligned} \quad (1)$$

and $\kappa(\mathbf{h}, \mathbf{k})$ has its usual meaning. This more restrictive *CONVERGENCE* procedure does not enable a starting set to be found.

When the modified *CONVERGENCE* has been run, the reflections are divided into a number of separate but connected systems as follows. The first system contains a 'number of unknown reflections' (NUR1), which should be ≥ 70 ; it must also include reflections that define the origin. However, with the observance of the lower bound of 70, NUR1 is kept

Table 1. *The system of weights used in (3) for the case when reflections can have general phases or be restricted to the pairs (0°, 180°) and (90°, 270°)*

Reflection type	$S(\mathbf{h})$	$C(\mathbf{h})$
General	1	1
Special (0°, 180°)	0	1
Special (90°, 270°)	1	0

as low as possible consistent with the following conditions:

- (i) the number of Σ_2 relationships generated within the system for any reflection must be ≥ 3 ;
- (ii) these Σ_2 relationships should give $\langle \alpha(\mathbf{h}) \rangle \geq \alpha_p$ for all reflections.

For the second system, the reflections of the first system are regarded as known and the minimum 'number of unknown reflections', NUR2, is $\geq (\text{NUR1})/3$. Again, with that constraint, the size of the second system is kept as small as possible consistent with the conditions (i) and (ii) given above. Further systems are constructed until all the large reflections are included. The minimum-size constraints of subsequent systems are governed by

$$\text{NUR3} \geq (\text{NUR2})/2$$

and

$$\text{NUR}n \geq 2 \times (\text{NUR3})/3 \quad \text{for } n \geq 4.$$

This division of the total set of large reflections into well conditioned systems is designed to give better phase estimates and to speed up the computational process.

The phase-estimating equations

The basic equation that is used to estimate phases is of the form

$$\sum_{\mathbf{h}} B(\mathbf{h}) |\alpha(\mathbf{h})_e - \langle \alpha(\mathbf{h}) \rangle| / \sum_{\mathbf{h}} B(\mathbf{h}) \langle \alpha(\mathbf{h}) \rangle = 0, \quad (2)$$

where

$$\begin{aligned} \alpha(\mathbf{h})_e = & \left(S(\mathbf{h}) \left\{ \sum_{\mathbf{k}} \kappa(\mathbf{h}, \mathbf{k}) \sin [\varphi(\mathbf{k}) + \varphi(\mathbf{h} - \mathbf{k})] \right\}^2 \right. \\ & \left. + C(\mathbf{h}) \left\{ \sum_{\mathbf{k}} \kappa(\mathbf{h}, \mathbf{k}) \cos [\varphi(\mathbf{k}) + \varphi(\mathbf{h} - \mathbf{k})] \right\}^2 \right)^{1/2}. \end{aligned} \quad (3)$$

$B(\mathbf{h}) = I_1(\langle \alpha(\mathbf{h}) \rangle) / I_0(\langle \alpha(\mathbf{h}) \rangle)$ is an effective weighting factor in (2) and $S(\mathbf{h})$ and $C(\mathbf{h})$ are weights for the real and imaginary parts of (2) for different reflections, as indicated in Table 1. The number of terms on the left-hand side of (2) equals the number of unknown phases but the number of individual Σ_2 relationships is much greater so what we have is a somewhat overdetermined set of equations.

Solving for the phases

Equations (2) cannot be solved explicitly for the phases but must be solved by a combination of the trial-and-error and least-squares methods. In the first stage, we give random characteristic values to the phases, where a characteristic value is one of 45, 135, 225 or 315° for a general reflection and 0 or 180° for a special reflection, which can only have one of these two values. Depending on the space group, other pairs of characteristic values can occur.

On the basis of these characteristic values, the next stage is to calculate all values of $\alpha(\mathbf{h})_e$ and hence the left-hand side of (2). Every $\varphi(\mathbf{h})$ is then judged against the criterion that, if

$$A = \sum_{\mathbf{k}} \kappa(\mathbf{h}, \mathbf{k}) \sin [\varphi(\mathbf{k}) + \varphi(\mathbf{h} - \mathbf{k})]$$

and

$$B = \sum_{\mathbf{k}} \kappa(\mathbf{h}, \mathbf{k}) \cos [\varphi(\mathbf{k}) + \varphi(\mathbf{h} - \mathbf{k})],$$

then for a good set of phases it would be expected that

$$A \sin \varphi(\mathbf{h}) \geq 0 \quad \text{and} \quad B \cos \varphi(\mathbf{h}) \geq 0. \quad (4)$$

The conditions (4) are based on the signs of $\sin \varphi(\mathbf{h})$ and $\cos \varphi(\mathbf{h})$ being the same as those of the tangent-formula numerator and divisor, respectively.

The reflection is found that does not satisfy these conditions and has the maximum value of $\alpha(\mathbf{h})_e$ and its phase is changed to the characteristic value that satisfies the inequalities (4). If this new estimate of $\varphi(\mathbf{h})$ reduces the value of the left-hand side of (2), then it is accepted.

This process is repeated until no further changes of phase are indicated. If all the NUR1 phases conform to condition (4), then the phases are accepted as the first stage of the solution. On the other hand, if the solution does not completely satisfy condition (4) for all phases, then a new initial set of characteristic phases is generated and the process is repeated. If after NUR1 such trials no solution completely satisfying (4) is found, then, of the NUR sets of phases that have been generated, that set of phases is accepted that gives the least value of the left-hand side of (2). This terminates the first stage of the trial-and-error process.

The accepted set of phases for the first system is now subjected to further refinement. The unknown phases are examined in sequence and the value of $\Delta\varphi(\mathbf{h})$ is found:

$$\Delta\varphi(\mathbf{h}) = |\varphi(\mathbf{h}) - C| \quad [0 \leq \Delta\varphi(\mathbf{h}) \leq 180^\circ],$$

where $C = \arctan(A/B)$. The value of $\varphi(\mathbf{h})$ is changed to C if one of the following three conditions is

satisfied:

$$\Delta\varphi(\mathbf{h}) > 45^\circ$$

$$\Delta\varphi(\mathbf{h}) > 30^\circ \quad \text{and} \quad \langle \alpha(\mathbf{h}) \rangle \geq 5.0$$

$$\Delta\varphi(\mathbf{h}) > 15^\circ \quad \text{and} \quad \langle \alpha(\mathbf{h}) \rangle \geq 10.0.$$

The changed value of $\Delta\varphi(\mathbf{h})$ is then accepted if it leads to a lower value of the left-hand side of (2); otherwise, it is restored to its original value. The rationale here is that some deviation from the tangent-formula indicated value is acceptable but the accepted amount of deviation will depend on the standard deviation of the estimate and will not be allowed to be greater than 45° in any case. This process is repeated in cyclic fashion until there is no further reduction of the left-hand side of (2), at which stage the phases are deemed to be the required solution of the overdetermined system (2).

Once the phases are determined for the first system, then they become known phases for the second system and the process described above can be repeated for the phases of unknown reflections in the second system. In this way, when all the subsystems have been dealt with in a similar way, there is available a set of trial phases that substantially satisfy (2).

Refining the complete phase set

From with the initial estimates of phases, a tangent formula is used for refinement. We have found two forms of the tangent formula to be effective. The first is the Sayre-equation tangent formula described by Debaerdemaeker, Tate & Woolfson (1988) and the second is a variant of the statistically weighted tangent formula given by Hull & Irwin (1978). The weighting scheme we use is

$$W(\mathbf{h}) = \min \left(C_w + \frac{\langle \alpha(\mathbf{h}) \rangle}{\alpha(\mathbf{h})}, 1.0, C_w + \frac{\alpha(\mathbf{h})}{\langle \alpha(\mathbf{h}) \rangle} \right), \quad (5)$$

where

$$\alpha(\mathbf{h}) = \left(\left\{ \sum_{\mathbf{k}} W(\mathbf{k}) W(\mathbf{h} - \mathbf{k}) \kappa(\mathbf{h}, \mathbf{k}) \right. \right. \\ \times \sin [\varphi(\mathbf{k}) + \varphi(\mathbf{h} - \mathbf{k})] \left. \right\}^2 \\ + \left\{ \sum_{\mathbf{k}} W(\mathbf{k}) W(\mathbf{h} - \mathbf{k}) \kappa(\mathbf{h}, \mathbf{k}) \right. \\ \times \cos [\varphi(\mathbf{k}) + \varphi(\mathbf{h} - \mathbf{k})] \left. \right\}^2 \left. \right)^{1/2}. \quad (6)$$

The constant C_w gives control over the phase refinement and we have found values of between 0.2 and 0.3 to be best for giving steady refinement. In the *STEP* program, the default value is 0.25. The two

refinement modes in *STEP* are called *STEP(S)*, which uses the Sayre-equation tangent formula, and *STEP(W)*, which uses the weighting scheme described by (5) and (6). In our experience, *STEP(W)* with the weighting scheme we use is better than that in *MULTAN* and this is the better one to use for large structures. The program selects the refinement procedure automatically but the user may change this choice by an appropriate keyword.

The figure of merit XDFOM

The process that has been described above is carried out from many initial trial phase sets, as happens with other multisolution methods, which then have to be ranked in order of plausibility by the use of figures of merit. In *MULTAN*, there are three separate figures of merit, ABSFOM, PSI0 and RESID, which are combined to give a combined figure of merit CFOM. For small to medium-size structures, CFOM can usually distinguish the correct solution but we have found that another combination of three individual figures of merit is even more reliable. This is

$$\text{XDFOM} = (1 + c)[\text{ABSFOM}/(R_h \times \text{PSI0})], \quad (7)$$

where R_h is similar to RESID and is given by

$$R_h = \frac{\sum_{\mathbf{h}} |\alpha(\mathbf{h}) - \langle \alpha(\mathbf{h}) \rangle|}{\sum_{\mathbf{h}} \langle \alpha(\mathbf{h}) \rangle}. \quad (8)$$

The constant $c = 1.25$ if there are heavy atoms in the structure and zero otherwise. Generally, it is found that a good structure will be found for XDFOM greater than 4.0 and when this value is found it is assumed that the solution has been found and the phase-determining process is stopped. *MULTAN* also has an automatic stopping procedure based on the values of RESID and PSI0 but XDFOM is found to be more reliable.

The SYSTEM90 package

The *STEP* process is incorporated in the package *SYSTEM90*, which also includes variants of the *MULTAN* components *NOF188*, *INK88* and *MAPS88*, modified to be compatible with the needs of *STEP*. The enhanced *INK* program controls *STEP* by means of six keywords. The whole programme is as automatic as *MULTAN* and *MULTAN* derivatives.

Applications of SYSTEM90

We have used *SYSTEM90* to solve four known structures, so that its performance can be compared with alternative methods, and one unknown struc-

Table 2. Figures of merit for all generated sets of phases for PTS6

Set 23, with XDFOM \geq 4.0, shows the structure.

Set	ABSFOM	PSI0	R_h	XDFOM
1	0.690	1.997	0.368	0.940
2	0.605	2.253	0.419	0.641
3	0.651	2.269	0.432	0.664
4	-0.355	1.146	0.892	-0.347
5	0.642	2.157	0.417	0.714
6	0.787	2.244	0.392	0.894
7	-0.176	2.111	0.790	-0.106
8	-0.090	2.080	0.741	-0.059
9	-0.016	2.215	0.693	-0.010
10	-0.195	1.558	0.802	-0.156
11	0.881	1.145	0.298	2.587
12	0.755	2.390	0.399	0.793
13	-0.279	1.239	0.858	-0.263
14	0.646	2.259	0.421	0.679
15	-0.328	1.041	0.886	-0.356
16	0.914	1.111	0.289	2.842
17	0.394	2.220	0.468	0.379
18	-0.394	1.007	0.908	-0.431
19	0.614	2.006	0.399	0.768
20	-0.455	0.961	0.927	-0.511
21	-0.334	1.093	0.883	-0.346
22	-0.322	1.197	0.872	-0.308
23	1.272	0.963	0.272	4.862

Table 3. Figures of merit for all generated sets of phases for TIN

Set 28, with XDFOM \geq 4.0, shows the structure.

Set	ABSFOM	PSI0	R_h	XDFOM
1	0.021	1.148	0.791	0.053
2	0.078	1.069	0.745	0.221
3	0.033	1.022	0.781	0.094
4	0.097	1.209	0.730	0.248
5	0.009	1.051	0.801	0.023
6	0.004	0.999	0.805	0.011
7	0.056	0.983	0.763	0.169
8	0.066	1.161	0.755	0.170
9	0.087	1.262	0.738	0.211
10	0.049	1.117	0.768	0.129
11	0.078	1.106	0.746	0.212
12	0.055	0.963	0.764	0.168
13	0.030	1.088	0.784	0.078
14	0.009	1.041	0.801	0.026
15	0.012	1.055	0.798	0.032
16	0.021	0.991	0.791	0.060
17	0.040	1.015	0.776	0.113
18	0.017	1.002	0.795	0.047
19	0.031	1.233	0.783	0.072
20	0.018	0.959	0.793	0.054
21	0.055	1.032	0.763	0.158
22	0.045	1.144	0.772	0.114
23	0.039	1.120	0.776	0.101
24	0.091	1.127	0.735	0.249
25	0.024	1.022	0.789	0.067
26	0.019	1.036	0.793	0.051
27	0.032	1.121	0.783	0.081
28	0.886	1.840	0.244	4.438

ture. We now describe the results and comment upon them.

(i) *CORT*, $C_{21}H_{28}O_5$, $P2_12_12_1$, $Z = 4$

This is the example structure of the *MULTAN* method and was determined by a default run of

Table 4. *The figures of merit for the correct solutions for all five structures*

The second column gives the number of the correct set and the default number of sets derived by *SYSTEM90*, which is 1.5 times the number of independent non-H atoms. The final column gives the number of atoms found in the *E* map compared with the number of non-H atoms in the asymmetric unit.

Structure	Solution set/ default sets	Figures of merit				Atoms found/ total number
		ABSFOM	PS10	R_h	XDFOM	
CORT	10/39	0.965	1.006	0.119	4.828	26/26
BHAT	16/63	0.873	1.102	0.171	4.628	40/42
PTS6	23/111	1.272	0.963	0.272	4.862	74/74
AFCMT	1/66	1.350	2.322	0.303	4.319	40/44
TIN	28/138	0.886	1.840	0.244	4.438	87/92

SYSTEM90. Set 10 satisfied the XDFOM criterion so the procedure then stopped. The *E* map showed all 26 non-H atoms.

(ii) *BHAT*, $C_{10}H_8F_4N_{10}O_{18}$, *Pc*, $Z = 2$

This is a difficult structure that was reported by Bhat (1990) to have resisted solution by *MULTAN80* and *MITHRIL83*. It could, however, be solved by *SHELXS85*. A default run of *SYSTEM90* gave an acceptable value of XDFOM for set 16 and the *E* map gave 40 of the 42 non-H atoms.

(iii) *PTS6*, $C_{27}H_{31}NO_9$, $P2_1/c$, $Z = 8$

This structure has pseudosymmetry and presents problems for that reason. It was originally solved by the program system *SAPI* (Fan, Yao & Qian, 1990), which has special provision for dealing with this kind of problem. In a default-mode run of *SYSTEM90*, set 23 gave a satisfactory value of XDFOM and the resultant *E* map showed all the non-H atoms. The individual figures of merit and of XDFOM are shown in Table 2.

(iv) *AFCMT*, $(C_{10}H_{12}S_8)_2Cl_6Cu_2$, *P1*, $Z = 1$

This material is an organic conductor that was solved by *SHELXS86* after attempts to solve it with *MULTAN80* and *SHELXS76* had failed. Set 1 of a default run of *SYSTEM90* gave a satisfactory solution according to the value of XDFOM and 41 of the 44 non-H atoms were found in the *E* map.

(v) *TIN*, $C_{36}H_{33}Cl_3O_7$, *Cc*, $Z = 8$

This is a new calixarene whose structure was previously unknown. Very little of the material was available so an analysis had not been done and its composition was assumed to be $C_{34}H_{33}ClO_5$. It was determined by a default run of *SYSTEM90* and gave a satisfactory solution according to XDFOM for set 28. Table 3 shows the complete list of figures of merit leading up to the complete solution. In the *E* map, the two molecular fragments containing two 28-atom rings in the asymmetric unit were clearly seen. Each

large ring was composed of four phenyl rings and four methylenes.

The *E* map gave 87 non-H atoms out of the 92 that are present. One cycle of Fourier refinement led to the complete structure. Interpretation of the map reveals that each TIN molecule contains 46 non-H atoms and there are three chloroacetyls, two methoxys and two hydroxys in each molecule ring. The structure is different from that assumed so in this case *SYSTEM90* has not only solved the structure but also determined the composition.

A summary of the characteristics of the correct solutions for each of the five structures is given in Table 4.

Concluding remarks

SYSTEM90 has three distinctive features that are also features of some other direct-methods packages.

(1) It automatically decides on the number of strong reflections required to define the structure according to chemical considerations. If the structure contains heavy atoms, the program will increase the number of strong reflections on the basis of the element's position in the Periodic Table. In this way, *SYSTEM90* deals equally well with equal-atom and heavy-atom structures.

(2) Solving for phases by stages gives phase estimates that have smaller errors than those derived by expansion from a small starting set. It also gives a system that is easier to solve than one that begins with a very large starting set. This approach seems to give *SYSTEM90* the ability to solve very difficult and large structures.

(3) An important feature is the ability of XDFOM to recognise the correct solution. Since the program then ceases automatically, it leads to very economical structure solution in terms of computer time while at the same time giving good-quality solutions.

It is well known by the users of direct-methods packages that no single package on its own will solve all structures. For one particular structure, *MULTAN* may succeed where *SHELX* fails; for

another structure, the situation may be reversed. It is therefore useful to have available a variety of methods incorporating as many different approaches as possible. The *SYSTEM90* approach is quite distinctive and we have shown that it stands comparison with other existing methods in its effectiveness.

Anyone wishing to obtain *SYSTEM90*, together with full information on its use, can do so by applying to HY.

The authors are grateful to Professor M. M. Woolfson for his interest in this project and his encouragement. The *SYSTEM90* program system was tested by Professor Fan Hai-fu and Professor Gu Yuan-xin, who provided data and were most helpful. HY also expresses his gratitude to his

advisor Professor Tang You-qi and to Professor Fu Heng and the Director and other senior scientists in the Institute of Chemistry. All of the above helped in the success of this project, which was supported by the NSFC, the Academia Sinica and the Laboratory for Structural Chemistry of Unstable and Stable Species. The authors also thank the referees for advice on the better presentation of this work.

References

- BHAT, T. N. (1990). *Acta Cryst.* **C46**, 112–116.
 DEBAERDEMAEKER, T., TATE, C. & WOOLFSON, M. M. (1988). *Acta Cryst.* **A44**, 353–357.
 FAN, H., YAO, J. & QIAN, J. (1990). *Acta Cryst.* **A46**, 99–103.
 HULL, S. E. & IRWIN, M. J. (1978). *Acta Cryst.* **A34**, 863–870.
 KARLE, J. & HAUPTMAN, H. (1956). *Acta Cryst.* **9**, 635–651.

Acta Cryst. (1994). **A50**, 753–759

The Density of Three-Dimensional Nets

BY S. T. HYDE

Applied Mathematics Department, Research School of Physical Sciences, Australian National University, Canberra 0200, Australia

(Received 3 January 1994; accepted 10 May 1994)

Abstract

Estimates for the density of periodic three-dimensional nets in Euclidean three-dimensional space (\mathbb{R}^3) are derived. The analysis assumes that the nets tile triply periodic hyperbolic surfaces that are free of self-intersections (embedded in \mathbb{R}^3). Upper and lower bounds of the net density as a function of the average ring size on the surfaces are given. These geometrical relations are compared with framework densities of a range of silicon-rich zeolites, silica clathrasils and dense four-connected silicates in order to separate the roles of geometry and chemistry in setting silicate densities. The data suggest that silica frameworks are constrained by an approximate requirement of constant area per framework vertex in addition to the impositions of Euclidean three-space and are thus hyperbolic two-dimensional (layer) structures.

Introduction

Although nets have intrinsic mathematical interest (as 'graphs'), their geometrical characteristics are of relevance also to the solid state. In particular, the bonding topology of covalent frameworks – such as

silicates – can be represented by a three-dimensional network. This has led to a number of theoretical studies of three-dimensional nets and their possible realizations as chemical frameworks (Wells, 1977; Smith, 1988; O'Keeffe, 1991). Despite the universal use of nets to describe structures in the solid state (any chemistry text is replete with examples), little fundamental work has been done. O'Keeffe has conjectured a number of challenging results and conjectures about three-dimensional nets, which suggest that the variety of three-dimensional nets realizable in Euclidean three-dimensional space is more limited than intuition would suggest. As yet, no procedure has been found for a systematic enumeration of three-dimensional networks (hereafter referred to as 'nets'), so it is difficult to establish their general characteristics.

Some intriguing relations between the bulk density of periodic nets and their topology have been reported. The relation between silicate densities and ring sizes has been reported and analysed to a limited extent by Stixrude & Bukowinski (1990). Nets of low density (number of vertices per unit volume), called 'rare' nets, are of interest as possible structure for zeolites. In this context, approximate relations between the net density and the size of the smallest