LEAD ARTICLE

The Control of Difficult Refinements

By David Watkin

Chemical Crystallography Laboratory, 9 Parks Road, Oxford OX1 3PD, England

Abstract

Each year, many single-crystal structure analyses are reported that show evidence of over- or under-refinement. Often, the refinement strategies have been naïve or over-complex and alternative strategies might have been more effective. The many descriptions of crystallographic and numerical techniques suitable for assisting with the control of difficult refinements are distributed widely in the literature and so are not always easily accessible. Without being a review of these procedures (which would require a substantial book), this article attempts to list readily available procedures, together with a brief outline of their backgrounds and examples of their applications to organic and organometallic compounds. The analysis of extended-lattice materials (usually inorganic materials) often raises problems in addition to those covered here. The particular aim of this article is to remind the reader that X-ray structure analysis is a modelling process and that, while standard models may be adequate for most analyses, more care and imagination must be applied to the treatment of difficult cases. Principles of methods are described without detailed mathematical derivations, although sufficient references to the literature are provided to permit careful study. Future requirements for refinement processes are outlined, including the use of new machine architectures, applications of sparse-matrix methods and the development of expert systems.

David Watkin received his doctorate from Birmingham University, England, in 1968, having studied with Dr Tom Hamor. Since then he has worked in the Chemical Crystallography Laboratory, Oxford. He has long been associated with, and is now responsible for, the widely distributed program system CRYSTALS. In recent years, he has become concerned about the growing gap between mathematical and analytical crystallographers. In 1987, he organized the first of the now biennial British Crystallographic Association Intensive Courses in X-ray Structure Determination, aimed at young chemists doing their own analyses. He is a member of the BCA Council and the IUCr Commission on Crystallographic Computing, and a consultant to the IUCr Commission on Crystallographic Teaching.

Introduction

The determination of small and medium-sized structures (of, say, less than 200 atoms) by single-crystal X-ray methods is on the verge of becoming a routine analytical tool. Commercial diffractometers are robust and reliable and are equipped with software packages that permit a scientist with scant formal training in crystallography to solve 80% of the structures he examines. A little experience and perhaps tries with alternative direct-methods programs may increase the success rate.

Sadly, this very success of semi-automatic systems has carried with it the seeds of a new problem, which has blossomed as a growing number of structures are published that are, if not actually incorrect, not the best that could have been made of the data. In addition, we have no way of knowing how many data sets are abandoned because the structures do not resolve easily and to a publishable conclusion. We now have the delicate situation in which harsh refereeing by one journal may lead to analyses being tucked into footnotes in other journals, where the crystallography is not closely refereed. These substandard analyses still find their way into databases, where they become indistinguishable from reliable work.

There is a risk that an article like this will only be read by scientists already keen to do first-rate work and so not reach those workers beguiled into believing that all the problems of structure refinement have been resolved. Even so, it is probably worth cataloguing some of the techniques that can turn a 'fair' structure analysis into a 'good' one. Often, all that are required are curiosity, patience and experience. Sadly, computers lack two of these attributes and the final responsibility still lies with the analyst.

In structure analysis, there are a number of main phases where user decisions have to be made: choosing the crystal, determining the space group, ascertaining that the direct or Patterson methods have yielded a valid structure and ensuring that the refinement has yielded optimal parameter estimates.

Phase one, choosing the crystals and collecting the diffraction data, usually requires no extraordinary skills

of the analyst, yet is often done badly in the haste to get a data set. This haste can lead to wasted time later.

The software provided by most diffractometer manufacturers (together with the high natural frequencies of occurrence of certain space groups) are a great aid in phase two, though Baur & Tillmanns (1986) give an enjoyable and informative account of published problematic cases. Space-group determination can sometimes, however, be a subtle and uncertain process, requiring cunning and experience to achieve an acceptable solution.

Phase three can often be completed by simple inspection of the join-the-dots diagram produced by most direct-methods programs. One reason for these self-evident solutions may be the thousands of hours of effort put into discovering, rationalizing and programming direct-methods techniques, which have led to programs with remarkable success rates, often with little user intervention. Though rarely discussed in such terms, these programs are *expert systems*, combining the mathematics and physics specific to the problems with careful use of discriminators and figures of merit to permit the user (or the program itself) to resolve otherwise degenerate or conflicting results. In addition, structure solution provides an opportunity for the chemist to use his experience with little recourse to crystallography. The outcome of a phasing calculation that is substantially correct is an image of the structure that the chemist may recognise, even if it is not what he expected. If he fails to recognise a correct solution, the worst that can happen is that the structure appears to remain unsolved. If he accepts as correct a false solution, this will generally show up as an unacceptably large R factor, a discriminator that enjoys a wide intuitive appeal.

It is in phase four that the analyst is likely to make the least of the available data. To the modern young crystal-structure analyst, refinement has become synonymous with least-squares refinement, in contrast with the older and often more informative Fourier techniques. Unlike Fourier refinement, least-squares methods are very amenable to being incorporated into semi-automatic procedures. Unfortunately, simple least squares distances the analyst from the refinement and gives rather poor diagnostics of incipient problems. However, the most dangerous feature of least-squares refinement is that it is concerned with the optimization of numerical values of the parameters in a mathematical expression that is supposed to represent the transformation of the structure amplitudes. The real transformation of these amplitudes is the continuous electron distribution in the crystal. The analyst must remember that replacement of the continuous distribution by atoms with simple form factors is only a mathematical convenience and he should not be unduly surprised if occasionally such a substitution fails to enable him to compute structure amplitudes that accurately mirror the observations. In structures with large thermal motion or serious disorder, the only real numerical description of the electron density may be Fourier-summation figure fields. The problem, of course, is that these fields can only be computed once the analyst has estimates of the phase angles, which usually requires the use of a parameterized model. Because of the cyclical nature of this argument, the analyst is forced back to developing an atomic model that best represents the continuous electron density, if only in order to generate valid phases for a Fourier synthesis, and perhaps in the end does not pay too much attention to the values of coordinates obtained for some atomic sites.

In addition, the analyst must always remember that there is the possibility that the observed data do not contain enough information to permit reliable estimates for the parameters in the postulated model. The estimated standard deviations of atomic and molecular parameters may be seriously underestimated in an absolute sense but, if correctly computed with proper attention to the effect of covariance between highly correlated least-squares variables, they do give a vital measure of the internal consistency of the model.

Until recently, programmers concerned with developing the refinement sections of structure-analysis packages have been concerned largely with the problems of performing least-squares refinements on small and/or slow computing equipment. A wide range of computational tools have been devised for dealing with particular problems but little attention has been paid to automatic selection of the best procedures to be used in individual cases or for producing integral tests for the validity of refinements. The result is that the analyst, having been given good instruments and quasi-automatic procedures for data collection, data reduction and structure solution, is suddenly left to choose his own refinement strategy. Even quite modest packages usually have a selection of refinement tools and utilities for assessing the quality of a refinement. However, refinement programs are designed by experts for use by the initiated. Manuals rarely have the space to explain what the utilities do, how their output is to be interpreted and what tools might solve what problems.

In single-crystal X-ray structure determination, we have available a technique of staggering power. A problem not yet resolved is how much of that power should be turned onto any particular problem. The conflicting temptations are to try to do the very best for every analysis or to be content with the first E map. No doubt the optimal solution is somewhere between these extremes, so there are two pressing needs at the present time. One is to show that a 'good' analysis can be obtained with only a little more thought than is required for a 'fair' analysis, so that all results that are deposited in the databases can be relied upon. The second is to try to encourage those few crystallographers still engaged in developing refinement programs to provide powerful numerical tools in packages that can be used

by scientists essentially ignorant of the principles behind the techniques they are trying to use.

This short article is not intended as a historical review of the many refinement tools found scattered through the literature but rather as a summary with brief explanations and examples. My concern is to bring together some of the advice and warnings already to be found elsewhere. Much of it is in texts concerned with crystallographic computing and so is likely to be missed by the hard-pressed synthetic chemist. The aim is to persuade the analyst that full-matrix refinement of 'anything that moves' is not the only permitted refinement strategy and, until programs are developed to do it for him, to encourage him to think critically about what has been or could be achieved.

It is perhaps useful to end this introduction with a general warning. Modern computer programs, the medium through which the crystallographer has access to refinement techniques, are very robust. That is to say, they rarely actually fail to execute to completion. This success by the program must not be confused with the idea of a successful refinement step. The fact that so many different refinement strategies 'work', yet lead to slightly different 'answers', distresses the beginner. Sadly, crystallographic mathematics is no different from any other, and the answer you get always depends upon the question you ask. The art, then, is to ask the right questions.

Background

The definition of a difficult or incomplete refinement is likely to vary from crystallographer to crystallographer and from structure to structure. The concept of a 'correct' structure is meaningless and must be replaced with something like 'fit for its uses'. The problem facing the practical analyst is that the final purposes may not be evident at the start of the analysis and that new purposes may emerge after a considerable lapse of time. This latter problem is becoming more acute as chemists make increasing use of computer-aided searches of structural databases. The high probability that the analyses are going to be processed mechanically with few checks being made on their validity imposes a duty on analysts to do the work as well as is economically possible. The real cost of performing an analysis properly is relatively small compared with the potential cost of having to resynthesize a product because the original analysis was incomplete or inadequate.

For this paper, a difficult or incomplete refinement could be seen as one in which the model is chemically or physically anomalous, or one in which the R factor [or weighted R factor, minimization function $M = \sum(w\Delta^2)$, goodness of fit etc.] has not reached as low a value as might be expected. Note that the minimization function can only approach its prescribed value, something close to the number of degrees of freedom, if valid weights are assigned. The debate about weights is continuing (Schwarzenbach et al., 1989).

Principles

It is not appropriate to repeat here the mathematics of least-squares refinement but it is necessary to describe some qualitative aspects of the method. The aim of the process is to obtain agreement between some function of the observed X-ray intensities and equivalent quantities calculated from a model. There are several definitions of an acceptable agreement, but that widely used in crystallography is that the weighted sum of the squares of the difference between the observed and calculated quantities is minimized – hence least squares (Rollett, 1984). Other functions are minimized in robust-resistant refinements (Nicholson, Prince, Buchanan & Tucker, 1982) and least absolute deviations (Press, Flannery, Teukolsky & Vetterling, 1986). These latter methods seem to be less sensitive to rogue observations. The reader seriously interested in these topics can find a good overview in Diamond (1984), Prince & Boggs (1992) and Press et al. (1986).

Unlike Fourier methods, which require the summation to be over the whole of reciprocal space, least-squares refinement does not require a 'full' data set. However, if only a subset of the full theoretical data set is used, it must be chosen so that it adequately defines the parameters whose refinements are being attempted. This idea is well expressed by the statistical term leverage (Atkinson, 1985). In outline, its implication is that the observations are not all bunched together in some way in one part of the observational space but rather are well separated and give good coverage of the space (imagine, for example, trying to refine cell parameters for a $20 \times 20 \times 20\,\text{Å}$ cell from setting reflections with h_{max} and k_{max} around 10 but l_{max} only 1). In addition, if a subset of the data is used, it is important that the error distribution of the subset reflects the error distribution of the whole data set. If the error distribution of the subset is biased in a different way to that of the full set, not only will the estimates of the errors in the refined parameters differ from those that would have been obtained from the full set but so also will the actual parameter values.

The distribution of errors in X-ray diffraction data and their influence on the final results is not well understood (Schwarzenbach et al., 1989), but it appears to be more complex than a Poisson distribution (Hong & Robertson, 1985). Working with a subset of the data that is selected by some simple function of the data itself is likely to be hazardous. Thus, while it may be acceptable to work with only the strongest reflections during the initial stages of refinement [see, for example, the initial refinement strategy adopted in SIR92 (Altomare, Cascarano, Giacovazzo & Guagliardi, 1993)], it is important to use much more of the weak data in the final stages. For the highest-quality work,

there is evidence to support working with all ('observed' and 'unobserved') reflections (Hirshfeld & Rabinovich, 1973). For a strongly diffracting crystal measured with Cu $K\alpha$ radiation, 'all' will probably mean all reflections up to the maximum Bragg angle of the diffractometer. With harder radiation, e.g. Ag $K\alpha$, all the very high angle data will be insignificant above background levels for organic materials and so will contain no useful information. The criterion for rejection of reflections in the final stages of refinement should not be the individual $I/\sigma(I)$ values, but the θ value at which the majority of reflections become indistinguishable from noise. Below this θ value, all reflections should be used, since a very weak low-angle reflection has a high information content and so must be included in the refinement.

The analyst needs to hold clearly in the back of his mind that most discussions about the use of very weak reflections or of weighting schemes based largely on $\sigma(I)$ values are made subject to the conditions that the e.s.d.'s of the data really do reflect the dependability of the observations. In fact, e.s.d.'s computed from counting statistics alone are only a measure of the reproducibility of the measurement if all factors except the time at which the observation is made are held constant. They are only very loosely measures of the absolute dependability of the data. Thus, while there is a lobby for including all reflections, the literature contains occasional instances that justify a less mechanistic attitude. For example, a recent analysis was published for a structure in space group $Pnn2$, the space group being chosen on the basis of the systematic absences (Cousson, Nectoux & Rizkalla, 1992). Marsh (1993) noticed (he does not reveal why) that the structure had pseudosymmetry higher than that required by the space group and that, if the 'systematic absences' were ignored, the structure could be very satisfactorily refined in space group $Pnna$. The systematic absences were "obviously close to the 'unobserved' cutoff .... Their presence, if real, can surely be blamed on the Renninger effect". Experiences such as this serve to remind us that reliable estimates of $\sigma(I)$ cannot be made from the routine practice of making only a single observation of each reflection but require multiple observations of the reflection and its equivalents under different but in principle equivalent regimes.

For routine work, a minimum $I/\sigma(I)$ value of 3.0 seems to work well, with little to be gained, other than appearances, by setting the limit to, for example, $F_o > 2.0\sigma(F_o)$ [equivalent to $I > \sigma(I)$] for a weakly diffracting crystal. A better solution may be to reduce the number of variables in the model or add supplementary observations (restraints). The norm of six to ten observations per variable only really applies if the observations are of fair quality and the parameters are properly spanned by the data. If data at very low $I/\sigma(I)$ are to be included in the refinement, then the refinement should be against F_o^2, permitting the use of negative observations and removing any bias. The weights will need to be different from a conventional refinement $[(w')^{1/2} = (w)^{1/2}/2|F_o|$, where w is a weight computed according to the analyst's favourite scheme] if the same minimum is sought (note, however, the problem that emerges as F_o tends towards zero).

Crystallographic refinement is carried out by the method of nonlinear least squares. There are a number of complications that arise from this nonlinearity. The method proceeds from an initial model to a refined model by successive adjustments. There is no way of guaranteeing that the model will improve or even that the process is convergent, nor of knowing that, having found a good model, there does not exist an even better one (Gopinathan, Whitehead, Coulson, Carruthers & Rollett, 1974). Equations (1) and (2) (observational equations, see later) show one of the sources of the difficulty:

$$a_1 x_1 + a_2 x_2 + \ldots = y_1, \tag{1}$$

$$\frac{\partial y_c}{\partial x_1}\delta x_1 + \frac{\partial y_c}{\partial x_2}\delta x_2 + \ldots = y_1 - y_c. \tag{2}$$

In linear least-squares refinement [(1)], the terms on the left-hand side are functions of the coordinates in the space at which the observations were made. The choosing of the observations to be made is an important stage in the design of an experiment, which is why the matrix of observations on the left of (3) is called the *design matrix*;

$$Ax = y. \tag{3}$$

In nonlinear least-squares refinement [(2)], these terms are replaced by estimated derivatives of the value of the observation with respect to the model parameters. Since we cannot observe these derivatives, they must be calculated from the current model. This has two unfortunate implications. The first is that we cannot easily choose the terms in the design matrix [see, however, Rollett, McKinlay & Haigh (1976) for a discussion on how to choose the most useful terms and Milledge, Mendelssohn, O'Brien & Webb (1985) for how to choose the most useful observations]. The second is that the actual values in the design matrix, and hence the normal matrix, are calculated from the model itself and have nothing first hand to do with the observations we have made (except that we have somehow - by direct or Patterson methods - obtained the model from them). One evident outcome is that an essentially incorrect model is unlikely to yield derivatives that will lead to improvement of the model. A second is that it is possible to construct models that lead to a matrix that is numerically unstable or singular. Working with such matrices usually leads to meaningless parameter shifts unless special care is taken (Forsythe, Malcolm & Moler, 1977).

It is perhaps worth digressing for a moment to re-state that in crystallographic least-squares refinement we attempt to minimize $\sum(w\Delta^2)$, where Δ is normally $F_o - F_c$, though there is currently a strong lobby for the use of $F_o^2 - F_c^2$. The structure we obtain at a minimum depends only on the values of w and Δ and not in any way on the terms in the normal matrix, the left-hand side of the normal equations

$$A'WAx = A'Wy. \tag{4}$$

However, the terms in the normal matrix determine how, if at all, we arrive at the minima and the use of bad weighting, a bad starting model or an ill-conditioned matrix may slow down or prevent the refinement converging. The corollary is that we can apply any modifications we like to the left-hand side as aids to convergence without fear of prejudicing the outcome. A minimum is a minimum, however achieved. Note, however, that, since the weights also appear in $\sum(w\Delta^2)$, the values of parameters at the minimum may depend critically on the choice of weights. Finally, it must always be remembered that least squares is a method for refining an already fully parameterized model. The technique has no mechanism for warning the user that there are still significant trends in the residuals that could be removed by the introduction of a particular parameter. The analyst must use some other technique to discover which parameters are missing. However, there are cases where the method can inform the analyst that he has introduced too many or redundant parameters (Lawson & Hanson, 1974)

Table 1 lists some potential sources of difficulties during refinements and Tables 2 and 3 give some mathematical and crystallographic tools that may be suitable for various situations. The items in Table 1 are covered by examples throughout this article since several of the crystallographic and mathematical tools may be useful for their diagnosis and, perhaps, their cure.

Mathematical tools

(i) Double-precision mathematics

In crystallography, there are usually very many more observations (one for each reflection) than there are unknown variables (one for each model parameter) so the problem of solving the set of simultaneous equations that define the parameters as a function of the observations should be well overdetermined. Standard texts usually remark that solution of these equations *via* the normal equations is the least-satisfactory procedure, especially by methods that involve inversion of the normal matrix. The preferred method (see, for example, Press, Flannery, Teukolsky & Vetterling, 1986) is *singular-value decomposition* of the observational equations, which can give useful diagnostics if there is something wrong

Table 1. *Potential sources of difficult refinements*

Bad or insufficient data
Wrong space group
Missing atoms
Pseudosymmetry
Disorder
Twinning

Table 2. *Mathematical tools useful in the resolution of difficult refinements*

(i)   Double-precision matrix work
(ii)   Model building
(iii) Matrix of constraint
(iv)   Reparameterization to orthogonal parameters
(v)   Rigid groups
(vi)   Observations of restraint
(vii) Shift limiting restraints
(viii) Refinement without X-ray data
(ix)   Characteristic-value filtering

Table 3. *Crystallographic tools useful for trying to understand the reasons for a difficult refinement*

(i)   Fourier synthesis
(ii) Slant (or generalized) Fouriers
(iii) Thermal parameters, *ORTEP*, principal axes, **TLS**
(iv) Analysis of residuals
(v)   Detection of outliers
(vi) Refinement against F^2

with the model or data. In spite of this, most small-molecule crystallographic programs do use the normal matrix and its inverse and are often perfectly satisfactory. The reason for this is that the structure-factor equation generally involves parameters that are well resolved, so the normal matrix is well conditioned, which implies that its inverse is well defined. In fact, crystallographic refinement can be expected to be so numerically stable that most programs do the matrix work in single precision. This benign nature of the bulk of problems is fortunate, since the size of crystallographic problems would make sophisticated numerical techniques or double-precision representations almost impossible. However, designing refinement programs to a minimal level of numerical sophistication does mean that problems of numerical instability are more likely to occur. The alternatives – rugged and reliable numerical methods – inevitably lead to larger and slower programs (though not necessarily ones requiring any additional interaction with users) that will perform uncompetitively on routine tasks. As the cost of computer memory continues to decrease and processors become faster, we can hope to see more programs perform matrix work in 64-bit precision and the use of more robust numerical methods. For the moment, we should expect mixed-precision programs, using double-length accumulators in vital inner loops.

(ii) *Model building*

Refinement consists of two processes that are usually alternated and iterated: model building and parameter optimization. The initial stage, structure solution, consists of finding a model that makes physical sense and yields structure factors in substantial agreement with the observed values. This model is then refined by optimizing the parameter values and perhaps increasing the complexity of the model. Fourier techniques are described below as a method for improving initial parameter values, particularly positional ones. However, geometric methods are also very valuable for improving a model that is recognizable but either highly distorted or incomplete. Geometric placing of H atoms is one much-used form of modelling. Two others are regularization and refinement against restraints without use of the X-ray data. In regularization, a desired configuration of the atoms is mapped onto and replaces or augments the current distorted model. Thus, if three atoms can be located for a phenyl group, the remaining three can be predicted with a fair degree of confidence. The principle can be extended to larger groups or whole molecules. An example is given later.

The mapping of the ideal group onto the proto-model can be done in several ways (Watkin, 1980) and the user should take care to choose that which is most appropriate. The method involving a curvilinear transformation (Diamond, 1976) does not appear to have had much use, but the general linear transformation is useful if it is suspected that the structure may be a dilated or contracted image of the idealized model, *e.g.* a compressed octahedron. Care must be taken with this method if the proto-model is essentially planar but the template is not, since the general transformation may result in a change of hand and/or substantial dilations. More commonly, the user will really want an orthogonal transformation – one that preserves the bond lengths. Several algorithms exist for this (Kabsch, 1978; Mackay, 1984; Diamond, 1988, 1990), including ones that preserve chirality. It is usually a wise precaution to look at graphical representations of structures that have been extensively geometrically modelled and to check that there are no impossibly short nonbonded distances. Regularization is, of course, an essential prerequisite for rigid-body refinement.

(iii) *Matrix of constraint*

Conditions may arise where the design matrix can hardly distinguish between a number of numerically equivalent but perhaps physically different solutions. Improved mathematical techniques (*e.g.* principal-component analysis) might resolve this ambiguity but such approaches may be rather blunt tools if the physical solution can be estimated by reference to some well established law or by comparison with accumulated scientific experience. In this situation, we may say that the solution, out of all numerically equivalent solutions, must satisfy certain conditions and, having in effect removed some degrees of freedom from the problem (by defining a partial outcome), let the mathematics resolve the remaining parameters. This imposition of a desired outcome on the solution of simultaneous equations may be accomplished by the method of *Legrange multipliers* (Hamilton, 1964) (an approach rarely used in crystallographic practice) or *via* a *matrix of constraint* (Larson, 1980). The effect of this procedure is that the solution is constrained to conform to certain predefined conditions that define either absolute parameter values or strict relationships between them. This procedure reduces the number of independent parameters in the problem so that those remaining may be defined by the data to a greater precision than in a totally free refinement and, as such, is a powerful tool for ensuring that the solutions are physically reasonable.

The catch, of course, is that the analyst must be careful to impose valid constraints. Even if he were to impose nonsensical requirements on the solution, the calculation must inescapably apply them. Now nonsensical constraints, such as might be the result of typing errors in the data for the program, would probably be evident. A much more serious situation exists when the constraints are marginally inappropriate. It may then not be clear that they are not optimal and it will be difficult to analyse the effect they will have had on the determination of the remaining parameters. Nonetheless, constraints provide a convenient way of imposing physically reasonable solutions to the observational equations, with the particular advantage that they reduce the size of the normal matrix.

A common use of the matrix of constraint is in the relating of atomic parameters so that they conform to the space-group symmetry. However, it could be invoked in order to impose symmetry on the shifts of parameters not formally related by the space group. Examples are the imposition of pseudosymmetry operators and the generation of 'riding' models for parameters.

Care should be taken to distinguish between riding models in least-squares refinement, when parameter shifts are related, and riding models for vibration, in which it is postulated that some part, or all, of the molecular framework is inflexible, so that the thermal motions of the component atoms are related (Johnson, 1980, pp. 14.01–14.20). A simple riding model in least-squares refinement might link the shifts in the site occupancies of a number of atoms, say a molecule of solvation (remember that in nonlinear least-squares refinement we are concerned with shifts in parameters rather than the parameters themselves). If the molecule is given full site occupancy, when in fact some solvent has been lost, and this parameter is not refined but the temperature factors are, these may take on large values in an effort to accommodate the reduced electron density (note the implicit constraint that the occupancies are unity). Normally, if the solvent is of no great

interest, the analyst might be happy with unreasonable temperature factors, but a careful examination of a difference electron-density map might show residual density around the solvent sites because changes in temperature factors do not have quite the same effect on the structure factor as changes in site occupancies. If the analyst now tries to refine both individual temperature factors and individual site occupancies, the computation will almost certainly fall into ruin unless the data are of excellent quality and extend to high θ values. One solution to this problem is to use the matrix of constraint to ensure that, even though the individual atoms may have different temperature factors, they all have the same site occupancy. The atoms are all given the same occupation number and this parameter is adjusted synchronously. In the matrix of constraint, this is achieved by defining a least-squares parameter that has no simple physical representation but is a weighted sum of the derivatives of the structure factor with respect to all of the occupancies involved. Thus, all of the atoms contribute to the definition of the shift to be applied to their now common occupancy. A similar strategy can be used to help control temperature factors in an unruly side chain. Remembering that in nonlinear least-squares refifnement we are computing parameter shifts, we can in this case arrange for successive atoms to have progressively larger starting values for the temperature factors. For complete numerical accuracy, the shifts should be scaled as a function of the distance of the atoms from the root of the vibrational disorder, but merely making them ride will normally be adequate.

The most common use for riding refinement is to preserve the local geometry of H atoms bonded to heavier atoms. For a CH_2 group, the three x coordinates are added in to one least-squares variable, as are the y and z coordinates. Thus, the contribution of the H-atom derivatives does have an effect on their joint shifts but the computation is dominated by the carbon contribution. This constraint ensures that the two C–H vectors remain of constant length but also fixes their orientation in the cell, which might not be desirable. It should, of course, be possible to make any groups of parameters ride. In programs that purport to do this, it is worthwhile checking that derivatives are jointly added into the matrix, rather than a simple shift being computed for one parameter and then that shift being applied afterwards to other parameters without their contributions having been added into the normal equations. Apart from slowing down the refinement, this naïve strategy leads to incorrect parameter standard deviations. Note also that a riding H atom will have the same positional e.s.d.'s as the parent C atom, though the bond-length e.s.d. will be zero.

It could be suggested that riding hydrogen constraints might represent computational overkill and that a much more cost-effective process would be to compute the H-atom positions geometrically, refine the C or other

skeleton atoms and then delete the H atoms and recompute their coordinates. In any case, because a positional riding model preserves bond vectors, it can lead to distortion of bond and torsion angles. H atoms must be included in medium- or high-quality refinements because of their contribution to Δ, but their contribution to derivatives is probably small and unreliable.

It must always be remembered that, when a parameter is omitted from the refinement, it is made the subject of an implicit constraint and is given some intrinsic value whose interaction with other parameters cannot be assessed unless it is introduced as a variable. For example, refinement of an isotropic temperature factor implies the constraint that $U_{[11]} = U_{[22]} = U_{[33]}$ and appropriate relationships for the cross terms. Other parameters are often constrained by default (occupancies to unity, and unapplied extinction and absorption corrections). There is no such thing as an unconstrained refinement.

(iv) Reparameterization to orthogonal coordinates

In some cases, the analyst can foresee that parameters may be highly correlated and that a simple transformation of the coordinate system will reduce the problem. For example, most crystallographers know that the high correlations between x and z in $P2_1/c$ with a very large unique angle can often be reduced by a transformation to $P2_1/n$, when they can expect the unique angle to be closer to 90°, so that x and z are now almost orthogonal. In another example, where there is a pseudomirror, the shifts in some parameters on refinement should be approximately equal in magnitude but opposite in sign to the shifts in corresponding parameters. A standard refinement will almost certainly not produce symmetrical shifts. If the original physical parameters are replaced by carefully chosen new least-squares parameters, a symmetric solution will be obtained not far from the trial model (Rae, 1973). This technique is in fact related to characteristic-value filtering, except that we can estimate beforehand the rotation needed to resolve the otherwise degenerate parameters.

The appropriate transformation involves a rotation of the normalized coordinate system by 45°. This is achieved by defining two new parameters:

$$x' = x_1 + x_2$$
$$x'' = x_1 - x_2 \tag{5}$$

where x_1 and x_2 are the pseudorelated parameters (Prince, 1982). The least-squares procedure is then performed on x' and x'', and x_1 and x_2 are extracted after the matrix inversion. The correct computation of the e.s.d.'s of molecular parameters involving x_1 and x_2 becomes more complicated, but not more so than in the case where the atoms are parts of rigid groups. Even after this coordinate transformation, the matrix work is still liable to be sensitive, but very mild shift-limiting restraints (see later), e.g. 1.0 Å, will control things

without slowing down the refinement unnecessarily. The technique has proved to be very versatile in dealing with problems of pseudosymmetry caused by pseudocentres and superlattices. It can conveniently be implemented *via* the matrix of constraint.

(v) Rigid groups

The matrix of constraint has so far only been described with respect to its influence on small numbers of parameters. However, it is not difficult to extend the implementation so that complex shapes can be manipulated as coherent and inflexible units. The inflexibility can be restricted to apply only to the relative atomic positions (molecular shape) or can represent relationships between the atomic temperature factors. For n atoms, the $3n$ atomic positional shifts are replaced by six composite least-squares parameters, three defining rotational shifts and three translational shifts. The $6n$ anisotropic temperature-factor shifts are replaced by 20 shifts to a **TLS** thermal model, in which the whole group is treated as having translational and librational harmonic motion (**T** and **L**) and these motions are coupled through the screw tensor, **S**. Rigid-body refinement is often advised as a technique for reducing computational time (since the normal matrix is often much reduced), but its more important role is in improving the observation-to-parameter ratio in cases where there is a shortage of data (perhaps through difficulties in getting good crystals) and the subunits of the structure are reasonably well characterized. This is often the case with large organometallic compounds and with polypeptides and polysaccharides. A further case relevant to this paper is that in which the refinement becomes unstable owing to an inability to resolve atomic parameters, for example when there are interpenetrating disordered molecular fragments. If the geometry of these fragments is well characterized from other experiments and there is no reason to expect that they are distorted in the material under investigation, then a very productive strategy is to regularize the geometry of the fragments to their accepted values and then refine them as rigid groups. A rather useful rule of thumb for refinement is that, if the atoms are not reasonably well resolved in an F_o map or, better, in a $2F_o - F_c$ map, then there will almost certainly be problems with the refinement and constraints or restraints will be needed. In the event of constrained refinement leading to the phasing of electron-density maps that clearly reveal atoms, it will generally be possible to remove the constraints.

The analyst should note that, broadly, there are two approaches to programming rigid-group refinements. In one scheme, the rigid model is presented in a standard orientation and location and the angles and translations needed to bring it to its best position in the cell are computed and updated at each cycle of refinement. The rotation matrix is linearized for the computation of the

angular corrections but the nonlinear matrix is used to compute the updated coordinates. In the other scheme, the model group is placed at approximately its correct position and orientation in the cell and then the linear rotation matrix is used to compute angular shifts that are applied, through the linear approximation, directly to the model. This second approach, which has the computational advantage that no matrices other than the matrix of constraint need be computed (and thus leads directly to the correct computation of the e.s.d.'s of the molecular parameters), has the disadvantage that for large rotational shifts of large groups there may be some slight distortion of the group. If the program description does not say which method is being used, the user should check that atoms near the periphery of a large group have large positional e.s.d.'s but zero e.s.d.'s on bond lengths.

(vi) Observations of restraint

The matrix (6) and the block diagram in Fig. 1 show how all the observations are filtered through the matrix of constraint, so that its effect cannot be escaped, whether the imposed constraint is reasonable or not. A more sensitive approach to influencing the outcome of a refinement is to suggest to the mathematics a solution that you would like to see, providing it does not come into too great a conflict with the X-ray observations. A mechanism for this gentle approach to fixing the result is to use observations of restraint.

$$\delta x_{\text{applied}} = P(M'A'WAM)^{-1}M'A'W\Delta Y,$$

$$A = \begin{bmatrix} \dfrac{\partial F_o/\partial x}{} \\ \dfrac{\partial R_t/\partial x}{} \\ \hline 1 \end{bmatrix},$$

$$\Delta Y = \begin{bmatrix} \dfrac{F_o - F_c}{} \\ \dfrac{R_t - R_c}{} \\ \hline 0.0 \end{bmatrix}, \tag{6}$$

$$\delta x_{\text{applied}} = P\delta x_{\text{leastsquares}},$$

$$x_{\text{physical}} = Mx_{\text{leastsquares}} + c,$$

where A is the matrix of derivatives, R_t is a restraint target value, ΔY is the vector of residuals, P is the matrix of partial shifts, M is the matrix of constraint and c is a vector of constants.

If some aspect of the hoped-for refinement can be expressed as equations relating the crystallographic parameters to some known and desirable quantity, then the equations can be differentiated and treated in exactly the same way as the structure-factor equations. The differentials of the variables go into the design matrix as usual and the differences between the desired values of the function and those computed from the current model are added to the difference vector in the normal way. The sceptical reader will wonder, very reasonably in view of the controversy about the weighting of ordinary structure-factor observations, how these supplemental

equations should be weighted. As always with least-squares weighting, the numerical values of the weights depend upon what the analyst is trying to achieve. An initial proposal might be that, during the model-building stage, especially if the data are poor and scarce and the known functions of the parameters are indeed very well known, their equations can be given a relatively large weight. As the refinement proceeds, it becomes necessary to discover if the experimental observations are sufficient by themselves so that the restraint weights can be reduced to zero. A practical compromise is for the analyst to weight the X-ray data in his favourite manner, assign relative weights between the equations of restraint reflecting his confidence in these equations, and then scale X-ray and supplemental equations together (Rollett, 1970). This is a form of hypothesis testing, in which the analyst is posing the question: 'Given my estimates of the accuracy of my experimental observations and my confidence in the supplemental functions of the parameters, does my model satisfy all these requirements and, if not, can it be made to?'. A great virtue of such a technique is that, if the experimental data do not really define some composite function of the parameters, the equations of restraint will try to modify the parameters to satisfy the restraint and at the same time not infringe the experimental observations. We remarked above that the matrix of constraint must be obeyed even if the requested constraint is nonsensical. With equations of

restraint, a nonsensical restraint (that is, one in conflict with the experiment) can be outvoted by the experiment, leading to a large (relative to the assigned weight) difference between the requested and computed values of the restraint. The equation of restraint is added into the normal equations just like any other equation and, if its weight is reasonably assigned, its effect will be no more than advisory.

The essence, therefore, is that restraints can be more gentle than constraints and can be applied when the analyst has little more than a hunch about the best solution. They might also appeal to the programmer through their ease of implementation and their delicacy of application. For example, the matrix of constraint, which is commonly used to fix a C-H bond length, will also fix its direction. This is usually undesirable. Suggesting a desirable bond length via an observation of restraint makes no demands on the orientation of the bond. Molecular-symmetry considerations might encourage the application of other restraints, such as common H-C-H, X-C-H or other repeated-fragment angles. The further condition of fixing the orientation of a bond with restraints is not difficult but does not seem to be much in demand.

Equations of restraint are easily incorporated into programs and a package might reasonably be expected to include functions to control distances, angles, planarity, thermal parameters, chirality, sums (and differences) of

Fig. 1. Tableau for damped weighted constrained restrained least-squares refinement, showing that restraints (R) are only added into the matrix with the observations (F) but that both observations and restraints are inescapably passed through the matrix of constraint (M). m = X-ray observations, n = restraints, p = shift-limiting restraints, $t = m + n + p$, r = crystallographic parameters and s = least-squares variables.

parameters, floating origins, absolute parameter values and parameter shifts.

(vii) *Shift-limiting restraints*

Shift-limiting restraints are a common feature of protein and some small-molecule refinement programs and are available in a slightly disguised form in *SHELX76* (Sheldrick, 1976) as a 'damping factor' [which in this case must not be confused with partial shifts, which are also often called damping factors - as in *ORFLS* (Busing, Martin & Levy, 1962) or *XTAL* (Stewart & Hall, 1986)]. Shift-limiting restraints tend to be rather crude implementations of the Levenberg–Marquardt algorithm, which is designed to improve the convergence properties of least-squares processes when only a very rough approximation to the normal matrix is available. (In proteins, this arises because only a band matrix - sometimes only diagonal - is accumulated for the X-ray data, together with a sparse scattering of off-diagonal terms coming from restraints.) The technique consists of augmenting the diagonal terms of the matrix with a constant that is a fraction of the trace of the matrix (Gill, Murray & Wright, 1981).

Choosing the optimal fraction is an iterative procedure that crystallographic programs rarely bother to follow. *SHELX76* adds the same constant to all the diagonal elements. In *CRYSTALS* (Watkin, Carruthers & Beteridge, 1985), the emphasis is shifted slightly and, instead of augmenting every diagonal element, contributions are only applied to parameters or groups of parameters that the analyst suspects will be correlated or ill-defined (Watkin, 1988). Whatever the details of the implementation, the general effect is that of suppressing singularities in the matrix and limiting the solution to one that is not too far from the starting point. If the matrix is well conditioned (that is, contains enough information to define all the parameters that the analyst is hoping to refine) and the augmentation of the diagonal elements is sufficiently delicate, the refinement proceeds more or less normally. However, if the matrix would otherwise be singular (a condition that eventually leads to incipient division by zero through the appearance of a zero on the diagonal of the partially processed normal matrix), the augmentation leads to the computation of a very small shift for the relevant parameter. In other words, if the matrix 'does not know what to do' for a particular parameter, more or less nothing is done and the parameter is only slightly modified. Shift-limiting restraints may well act as origin-fixing conditions for polar space groups, though the parameters involved will usually have large estimated standard deviations unless the restraints are very severe. They will also prevent the refinement of all coordinates of an atom on $(x, -x, z)$ from going singular but will not give a solution with the correct symmetry. That should be achieved either with the matrix of constraint or with characteristic-value filtering (see below).

(viii) *Refinement without X-ray data*

An obvious extension of the schemes described above is to keep the weights of the restraining equations at full value but reduce those of the X-ray data to zero, that is, refine against the restraints only. This technique is often called *distance least-squares* (DLS) refinement, after the well known program *DLS* used for the modelling of inorganic structures (Baerlocher, Hepp & Meier, 1977). However, it is not necessary to restrict the restraints to distances only. Any valid restraints may be used - geometric, shift-limiting, energetic. If the geometric restraints only refer to an isolated molecule and there are no atoms on special positions, then the normal matrix will be singular since it contains no information to locate the molecule in the cell, nor to define its orientation. For an extended (lattice) structure, there are usually known geometric features that span symmetry operators, *e.g.* an O–Si–O' angle (with O' related to O) may be defined as 109°. Sufficient geometric restraints involving symmetry operators will fix the origin and orientation of the structure. In the case of molecular crystals, normal bond and angle restraints only determine the molecular geometry and cannot fix the position in the unit cell. The only restraints spanning symmetry operators are from nonbonded contacts; for these, an asymmetric distance restraint (which inhibits short contacts while tolerating long ones) is required to fix the origin and orientation (Watkin, 1988). The technique of not refining three arbitrary atoms is not recommended for origin and orientation fixing and, if nothing better can be hypothesized, shift-limiting restraints or characteristic-value filtering should be used to control the incipient singularities.

Refinement without X-ray data is an excellent method for regularizing a distorted structure or for developing a new model from a related structure.

(ix) *Characteristic-value filtering*

The simultaneous equations for shifts in the crystallographic parameters obtained by a Taylor expansion of the structure-factor expressions do, in the happiest of situations, have a single unambiguous solution. However, when there are approximate (or unidentified exact) relationships between the apparently independent crystallographic parameters, these equations may have more than one solution, which will produce more or less numerically indistinguishable residuals. When this situation arises, simple numerical processing will select a solution largely on the basis of rounding errors in the arithmetic and may yield solutions that give enormous but complementary shifts. For example, simultaneous refinement of temperature factors and site occupancies often leads to them both taking on small but compensating values. In linear least-squares refinement, when the equations are solved for the unknown parameters themselves, this indeterminacy poses a very serious

problem. In nonlinear least-squares refinement, which is the situation in crystallographic refinement, we are only solving the equations for shifts in parameters. If we have a reasonable starting model, then a potentially useful solution to the degenerate equations is that which minimizes the changes in the parameters. This is the solution that emerges from *singular-value decomposition* of the observational equations. For n observations and m unknown parameters, this technique requires the storage of the nm observational coefficients and subsequently nn and mm orthogonal but not necessarily symmetric matrices. This is currently impractical but an approximation to this solution can be obtained by *characteristic-value filtering*, also called *latent root* or *eigenvalue* filtering or *principal-component analysis*.

To the human eye, the principal characteristic features of the normal matrix are its size and symmetry. However, with suitable numerical processing [(7)] it can be reduced to a diagonal form (D), that is to say, all the values not on the diagonal are zero. Those numbers that remain are the characteristic values, latent roots or eigenvalues of the original matrix and reveal much of the important information concealed in the original normal matrix. A rotation matrix (V), the latent or eigenvectors considered *en masse*, relates the diagonal matrix to the original normal matrix:

$$A = VDV'. \qquad (7)$$

The inverse of D is another diagonal matrix with elements $1/d$ [(8)]. If a d_i is zero, then its inverse is not defined and the corresponding term is set to zero. This may occur when there are linear relationships between the parameter shifts or there is no information in the data to define them. In fact, if any d_i is 'small', the inverse can be set to zero and so filter out small eigenvalues:

$$A^{-1} = V'D^{-1}V. \qquad (8)$$

Since the normal matrix is formed as the sums of squares of terms in the observational equations, some information may have been lost into rounding errors, so that eigenvalue filtering is less effective than singular-value decomposition; even so, this is a useful technique. In effect, the multidimensional space occupied by the model (100-dimensional for a model with 100 unknown parameters) is rotated until all dimensions are orthogonal or independent. The eigenvectors give us this rotation. The characteristic values or eigenvalues tell us about the rate of change of the structure factors with respect to the new independent pseudoparameters. Very small eigenvalues in effect tell us that the structure factors will be virtually unaffected by changes in the corresponding pseudoparameters, which are therefore almost undefined.

If in this new space some pseudoparameters are undefined, then they can be eliminated from the analysis, shifts can be computed for the remaining parameters

and the solution can be rotated back into the original crystallographic space. This yields the solution of minimal Euclidian length, that is to say, least perturbed from the initial trial solution. It requires all parameters to be referred to basis vectors of similar size, that is, similar shifts in all parameters lead to similar changes in the minimization function, but this is only a problem of scaling. In the problem with the polar direction, the floating origin is easily and correctly fixed by this procedure (Rollett, 1970), though it is computationally not as efficient as the method described ·by Flack & Schwarzenbach (1988). The symmetry requirements of an atom on $(x, -x, z)$ will also be taken care of automatically. The technique is widely used by professional statisticians for problems where there is little external evidence to help in the parameterization of the model. It is also used by crystallographers in most **TLS** analyses and in the program *DIFABS* (Walker & Stuart, 1983). It seems to have only rarely been used in structure refinement (Johnson, 1971), perhaps because of the size of the matrices that need to be stored but perhaps also because nonmathematical analysts might be shy to try the method.

Crystallographic tools

If, loosely, a good refinement is one in which the R factor is acceptably low, then all that is necessary to achieve a good refinement is to ensure that the data belong to a single experimental procedure and then propose a model that yields structure factors corresponding to the experiment. The problem, of course, is to be sure that all the data do come from a valid experimental regime and to know what sort of model to propose.

Patterson and direct methods now reliably give good trial models for the bulk of medium-sized organic and organometallic structures so it is not unusual to see an R factor fall below 20% after a few cycles of refinement; for many analyses, it continues to fall to 3 or 4%. When this fails to happen, there are various tools available to help the analyst understand the shortcomings of the data or model.

The following crystallographic tools are available for investigating an unsatisfactory refinement.

(i) *Fourier syntheses*

It is probably useful to spend a few moments in digression and consider the processes involved in crystal structure analysis. X-rays are diffracted by the periodic electron density distribution in the sample and we are able to observe the resultant diffracted intensity. We have equations that enable us to 'correct' the observations for various experimental conditions (*e.g.* Lorentz and polarization corrections, absorption corrections, corrections for crystal decay). Note that these corrections are really part of the model but that in general we

believe we understand them sufficiently well for the errors associated with the correction parameters to be insignificantly small, though this may well not always be the true situation. The intensities can be converted to structure amplitudes, which are related to the electron distribution in the sample by

$$F_{hkl} = \int \rho_{xyz} \exp[2\pi i(hx + ky + lz)] \, \partial x \, \partial y \, \partial z. \quad (9)$$

The complementary transformation to (9),

$$\rho_{xyz} = V^{-1} \sum \sum \sum |F|_{hkl} \exp[-2\pi i \\ \times (hx + ky + lz - \alpha_{hkl})] \quad (10)$$

permits us to compute the electron-density distribution in the crystal if we know the magnitude and phases of the structure factors. It is worth drawing attention to the fact that (9) and (10) are no more than mathematical models for a physical process and their applicability to each new problem should be reviewed.

If we could observe the phases, we could compute electron distributions using only (10), so our electron-density distribution is only one equation away from the experiment. We cannot readily observe phases but approximate ones can be estimated by direct methods and used in Fourier syntheses. The Fourier map can be 'interpreted' and a model proposed that may enable us to compute better phases than those with which we started, so that an improved Fourier synthesis can be computed. Two types of model are in common use. In one, the computed electron density is examined and features that are believed to be unreasonable are modified. For example, regions of negative density are eliminated and density is reduced in regions where it is too high. If structure amplitudes computed from (9) are close to the experimental values and the computed phases return a good approximation to the modified model when used in (10), the model is believed to be an acceptable approximation to reality. Though this method involves a minimum of numerical modelling, it is only occasionally used since the model, the Fourier figure field, is rather unmanageable.

For most structural analyses, a second and more problematic modelling is introduced, in which the continuous electron distribution is replaced by 'atoms', which have their own electron distributions, mean positions in space and, usually, some parameters describing their periodic displacements from the mean position. The computational advantage of such a model is that an atom can be represented by about 10 parameters, whereas the corresponding volume of a Fourier map, sampled at $1/3$ Å resolution, would require over 500 values. The parameterized model is now rarely determined in detail from the corresponding Fourier map (though at one time it was); instead, the structure factor is expressed directly in terms of the model,

$$F_{hkl} = \sum f_j \exp[2\pi i(hx_j + ky_j + lz_j)], \quad (11)$$

and the model parameters are refined by least squares.

Hybrid techniques are sometimes used, as in solvent flattening in protein crystallography (Blake, Pulford & Artymiuk, 1983), disorder modelling by discrete Fourier transforms (van der Sluis & Spek, 1990) and Bessel-function modelling (Bennett, Hutcheon & Foxman, 1975).

When a refinement is failing to proceed as expected, a most cost-effective method for trying to understand why is to examine the corresponding Fourier map in detail. If the model is fundamentally correct, then so also will be the computed phases. As pointed out above, least-squares refinement cannot introduce new parameters so, if the model is lacking some atoms, they must be sought in the Fourier map or be introduced by another technique. In addition, fine features in the map may reveal that a simple model with spherical atoms 'vibrating' harmonically about their mean positions may not really represent the actual electron distribution. A major problem now is that few analysts have time to examine carefully Fourier figure fields; instead, peak searches are used. These tend to be rather crude, giving little more than the positions of local density maxima and conveying almost nothing about topographical details of the density [but see, however, recent work by Fortier *et al.* (1993) and much earlier work by Johnson (1977)]. Even so, careful use of F_o-, $F_o - F_c$- and $2F_o - F_c$-map peak searches can reveal disorder, anharmonic vibration and charge redistribution. Whenever a refinement sticks, Fourier techniques provide the fundamental tools for visualizing the problem.

(ii) *Slant (generalized) Fourier maps*

The complete Fourier map for even quite a small structure is likely to be tedious to produce and examine. Fortunately, structures that are failing to refine are usually doing so because of rather localized problems. A good practical approach is to compute the Fourier map for just the region in question. Since the analyst himself is going to have to interpret the map, his task is often simplified by computing the sections so that they lie parallel or perpendicular to molecular or physical features. The maps need to be carefully contoured and then carefully examined for features that cannot be explained by the current model.

(iii) *Temperature factors*

Though least-squares refinement has no mechanism for introducing new parameters into a model, refinement of given parameters to physically unreasonable values may give clues about inadequate data processing or the need for major changes to the model.

If temperature factors have been refined for the problem structure, they will often reveal the site of trouble and sometimes indicate solutions.

(1) *Isotropic temperature factors (U_{iso}) too low.* Too low means very much smaller than might be expected in comparison with similar materials at similar temperatures. Apart from the possibility of errors in the atomic form factors, isolated low U_{iso}'s indicate incorrect atomic assignments and that an element of higher atomic number should be inserted at the site. If all the temperature factors are low, this may be caused by neglect of the θ-dependent part of an absorption correction. This situation is particularly common with corrections computed by empirical methods based on azimuthal scans. These leave a residual error corresponding to a cylinder or sphere with radius equal to the minimum dimension of the sample. The analyst must check in *International Tables for X-ray Crystallography* (1985) that the ratio of the absorption correction at zero and θ_{max} for the residual body is less than about 2. If it is greater than this, some sort of θ-dependent correction must be made. In materials with intrinsically low temperature factors (*e.g.* minerals and metal oxides), neglect of the θ-dependent correction often causes the temperature factor to go nonpositive definite. Double application of the Lp correction or taking square roots of the observations twice have similar effects.

If there is no evidence for errors in the data processing, the analyst must decide for himself what must be done about small or negative temperature factors. If they are ignored, some attentive referee may seek an explanation. If they are set to 'reasonable' values and then not refined further, there will always be a nagging doubt about the refinement. One solution that may offer comfort to the analyst is to restrain the temperature factor to currently acceptable values. If the minimization function is not seriously perturbed, so that the restraint is satisfied, all is well and the explanation is that the data do not define the temperature factors in question (leaving also the question as to what other parameters are not well defined). If the restraint is contravened, then the diffraction data do contain information relevant to the parameters under scrutiny and the analyst no longer has a problem that can be resolved in a general article such as this.

(2) *Isotropic temperature factors too large.* Again, this could be caused by mis-assignment of atom type. A fairly high temperature factor may mean that the atom is disordered, either positionally or through partial occupancy. A very large factor probably means the atom is not there at all. The proper strategy for analysing this problem is to look carefully at Fourier syntheses, but some clues may be made available by trading computer time for attention to detail and initiating anisotropic refinement.

(3) *Unusual anisotropic temperature factors.* Fortunately, the B and b temperature-factor representations are now going out of fashion and are being replaced by U's. U's have the advantage of having units of $Å^2$ and being independent of cell dimensions. Thus, the diagonal

components can be directly compared and anomalies spotted. However, the ellipsoid is only represented correctly by all six unique elements simultaneously and the effect of large off-diagonal terms may be difficult to visualize. The proper procedure for examining temperature factors is to extract the principal axes by diagonalizing the tensor. Atoms with very small or very large axes or, worse still, one small axis and one large axis should be suspect. In this latter case, the least-squares refinement may be trying to use the available parameters to model a more complex situation in which the atoms are really disordered over two sites. A confirmation that something is wrong may be given by anomalously short bond distances to the atom concerned, though this will be less evident if several adjacent atoms are all disordered.

A good improvement to the starting model for this kind of disorder is to place a 'half-atom' at each end of the longest axis of the anomalous temperature factor. Free refinement of these half-atoms will probably not work - the half-atoms will try to fuse together again. Refinement of this kind of model usually requires restraints or constraints. A frequently successful formula is to make the temperature factors of the two half-atoms ride (*i.e.* constrain them to be identical) and restrain the bond lengths from the disordered atoms to their neighbours.

Thermal-ellipsoid (*ORTEP*; Johnson, 1976) plots are invaluable for displaying the relationships between atomic temperature factors and should always be produced and examined even though some journals are reluctant to publish them. For molecular compounds, there are usually fragments that can be regarded as rigid bodies, that is, any motion of the atoms in the fragment will be correlated since the geometric dispositions of the atoms are fixed by the covalent bonds. In this situation, the ellipsoids of adjacent atoms should have an evident relationship between them. Any large or small axes must be explicable in terms of normal intra- and intermolecular forces.

Patterns of correlated motion identified in diagrams should be quantified by a TLS calculation (Schomaker & Trueblood, 1968). This verifies that the individual atomic motions can be represented by coherent motion of the fragment as a whole. For small groups of atoms (less than about ten), the calculation is rather unstable and some experience is needed in interpreting the results. Even so, it is a useful diagnostic tool for helping to identify atoms whose refinement is anomalous.

The final question raised by the presence of groups of atoms with large but correlated thermal parameters is whether they really represent thermal motion or disorder. If a definitive answer is important, the only effective solution is to examine the same specimen at successively reduced temperatures. Disordered atoms will continue to have large apparent anisotropy (unless a molecular reordering occurs at a phase change), while true thermal motion will slowly subside as the temperature drops.

(iv) *Analysis of residuals*

Difference density Fourier maps give a graphical representation of residual discrepancies between the observed and calculated structure factors and so may reveal some missing feature of the structure. It is always possible that the residuals will not transform into some feature recognizable in these maps. If the transform of the discrepancies is not interpretable, the residuals must be analysed in some other way. Plots of $\langle F_o/F_c \rangle$ and $\langle w\Delta^2 \rangle$ *versus* various criteria may be illuminating. Commonly used criteria are $|F_o|$ and $\sin(\theta)$. The ratio F_o/F_c should be about unity for the interval containing the median of F_o.

Ratios less than unity for strong reflections usually indicate the need for an 'extinction' correction, though this may really be serving to compensate for losses in the quantum counting chain.

Ratios less than unity for weak reflections may indicate too wide a scan during data collection, so that the tails of strong reflections were subtracted from weak reflections. This is a real problem for data collected from crystals with large cell lengths on diffractometers using molybdenum radiation but can sometimes be avoided by using an ω scan rather than an $\omega/2\theta$ scan.

Ratios greater than unity for the weak data may indicate systematic overestimation of these observations, systematic underestimation of strong reflections owing to too small a counter aperture or too small a scan width, or failure to include solvent.

Two methods are commonly used for defining the F_o intervals. They can be chosen so that each interval contains the same number of data, or so that the intervals are regularly spaced, often logarithmically. The second method is to be preferred, even though some of the intervals may contain only a few reflections. Choosing a wider interval in order to include a specified number of reflections usually means, in the high intensity ranges, that reflections differing by orders of magnitude are grouped together. An average property of such disparate items is probably worthless. A ratio less than unity for low-order batches against $\sin(\theta)$ may indicate that there is missing solvent.

It is also useful to see statistics ranked against each index, combinations of indices (*e.g.* $h + k$) and parity group. For a well refined structure with 'correct' weights, $\langle w\Delta^2 \rangle$ should be constant for all rankings. Two-dimensional plots of the ratio F_o/F_c *versus* h and k, h and l, and k and l may reveal problems that are directionally dependent, such as inadequate absorption corrections. If proper experimental corrections cannot be made, the program *DIFABS* will introduce new parameters into the model that can accommodate these discrepancies.

(v) *Detection of outliers*

In many branches of physical science, the investigator has a good idea about the magnitudes and distribution of the data. For example, the observation that a student is 17 m tall is likely to be in error. The distribution of structure amplitudes follows Wilson statistics but individual values can vary by several orders of magnitude. Bad reflections, outliers, can rarely be seen by inspection. Atkinson (1985) devotes a whole book to the problem of their detection but most of the remedies are impractical in crystallography.

Nicholson, Prince, Buchanan & Tucker (1982) draw attention to methods based on comparing $F_o - F_c$ with σF_o and show cases where the method has been very fruitful. However, σF_o is derived from the same observation as F_o so that, if the observation is in gross error, so also is the resultant σ. A complementary strategy, used in *CRYSTALS*, is to assume that, once a refinement is under way, the estimates of an individual F_c are reasonably reliable since each F_c is a complex mean, *via* the model and the structure-factor equation, of all the F_o. With normal data, plots of Δ^2 *versus* F_c usually follow a smooth curve, which can be represented by a simple mathematical function. This curve gives the average residual as a function of F_c and permits one to estimate the probable residual for each reflection. Reflections with residuals lying well away from this estimate are either fundamental to the identification of some as-yet-unparameterized feature in the material or are very much in error. The probable residual can be used much like σ in Nicholson's method and, with care, can be used as the basis for a weighting scheme.

(vi) *Refinement against F^2*

With suitable weighting functions, refinement against F^2 can arrive at parameters very similar to those from a conventional refinement though, with other weights, other minima will be achieved. There seem to be a number of reasons for using F^2 refinements:

(1) The standard deviations of F^2 are obtained from a linear operation on the standard deviations of the intensities so their distribution is not distorted.

(2) Very weak reflections may be observed with a negative net intensity, which can be preserved for the refinement.

(3) The contours of the minimization function are different from those of the conventional function so the path from a trial model to a final structure is not the same (Rollett, McKinlay & Haigh, 1976). These authors suggest that F_o^2 refinement may have a wider range of convergence than F_o refinement and my experience confirms this, providing some care is taken with the weighting. Attempts to start refinement with unit or simple statistical weighting often diverge explosively. This usually seems to occur when some of the strongest computed structure factors are in poor agreement with the observations. Down-weighting these strong reflections, for example using the scheme mentioned above, usually controls the divergence. Harris & Moss (1992) report the use of F^2 refinement for proteins and conclude

there is little to recommend it, and there is some evidence [much unpublished, but see Nardelli (1993)] that for small molecules the differences between an F_o and an F_o^2 refinement are largely insignificant. Even so, for small molecules the sometimes increased radius of convergence, together with the absence of any real disadvantages, suggest that it should be used routinely, though not necessarily (for reasons explained above in the discussion of the treatment of weak reflections) accompanied by the use of all data.

Problem situations

(i) Bad or insufficient data

Sheldrick is reputed to have replied, when asked what was the major cause for structures not solving easily, 'bad data or the wrong space group'. The same answer would probably serve as an explanation for many difficult refinements.

The speed and reliability of automatic search and index routines on diffractometers has convinced many analysts that they do not need to take conventional X-ray photographs of their materials. This may well be true for an experienced worker who knows how to use an optical microscope and knows how to make the diffractometer yield information about crystal quality. However, the procedures used by most diffractometers will permit data to be collected with really quite dreadful crystals. Evidence that the data are of poor quality only emerges as the refinement becomes difficult, by which stage the analyst may be beginning to lose interest anyway. Data-collection strategies will vary for the different classes of analysis and good practices have been set out by Hamor, Steinfink & Willis (1985). Blessing (1987) has drawn attention to extra precautions that need to be taken for first-class analyses. Because of the influence of the data on the subsequent refinement, some comments must be added here.

Traditionally, errors are treated as random or systematic. Random errors can never be removed from an experiment, though good design will reduce them. If they are such that observations are distributed symmetrically about the true value, they will not perturb the resulting model but only affect the parameter e.s.d.'s.

Systematic errors are ones that follow a (perhaps undetected) asymmetric pattern and are therefore biased. They will lead to both bias in the model and incorrect e.s.d.'s. The experimental technique can be improved to remove these errors (to reduce absorption errors, a small sphere should be used), or corrections can be made to the data (Gaussian or Howells-polyhedra absorption corrections), or the model can be extended to reflect the source and pattern of the bias (*DIFABS*).

It is sometimes useful to distinguish a subdivision of this second class – gross changes in experimental conditions during data acquisition. Occasionally, the

model can be modified to accommodate the changes (*e.g.* individual scale factors for each sample used in a multicrystal data collection). Otherwise, data may have to be eliminated (*e.g.* those collected with the shutter partly closed or those suffering from the Renninger effect) so that the data that remain come from a single experimental regime.

If data are being collected in reciprocal layers and the collection is terminated before all layers, to a given resolution, have been collected, some parameters will be poorly defined because there is poor leverage on them owing to the shortage of data perpendicular to the layers. Parameters likely to be affected by this sort of experimental shortcoming are positional parameters in that direction (which will also have large e.s.d.'s) and components of the anisotropic temperature factors. It is useful to see positional e.s.d.'s in ångströms (rather than in fractions) since, for a correctly refined good data set, they are approximately isotropic. At a recent *SIR92* workshop, a participant produced a set of data that had remained unsolved in his laboratory for several years. The structure was solved with *SIR92* and passed to *CRYSTALS* for refinement, which stuck at about $R = 20\%$, with curiously elongated thermal ellipsoids and positional e.s.d.'s. An examination of the data showed that the highest index in one direction had a magnitude of about half the corresponding cell length in ångströms. If the data set is going to be large or there is any suspicion that the crystals may decay, the reflections should be measured in reciprocal-space shells (*i.e.* over consecutive but limited θ ranges) so that, if the data collection ends prematurely, the acquired data will have approximately spherical resolution.

Eliminating strong low-angle data may make it difficult to refine an extinction coefficient. Working only with high-angle data will not permit H atoms to be properly located, since they only diffract well at low angles. The procedure for eliminating some data because they cannot be properly modelled is difficult and dangerous. It assumes that you have correctly identified the deficiency in the model (even though you cannot test this) and that the effect of the deficiency is restricted to a small part of the data.

(ii) Wrong space group

For the majority of structure analyses, the space group is fairly readily determined from the systematic absences and symmetry of the reciprocal lattice. This can lead to the belief that space groups are well defined features of nature, a belief easily encouraged by the enticingly beautiful order and clarity of the diagrams in *International Tables for Crystallography* (1992). A much safer approach to the space groups, however, might be to think of them as mathematically precise points in a continuum; the problem of space-group determination then becomes one of deciding which of these conventional labels best describes the relationship

between the atomic model and the X-ray observations. The space group is part of the model. In the absence of other evidence, analyses should always be started in the most symmetric space group. This principle, of *minimum assumption*, is fundamental to contemporary science and requires the analyst to seek the simplest description that will explain an observed phenomenon. Naturally, careful observations often require detailed explanations though, sadly, careless experiments also invite complex analyses. The role of the space groups is to codify exact relationships between parameters and in effect replace several by a smaller number plus a set of rules for generating the rest. The use of a high-symmetry space group can dramatically reduce the number of variables in the model and so is an economically seductive proposition. During the initial stages of analysis, the use of high symmetry and its ensuing reduction in computation is quite proper and will in general aid in the development of the model. Eventually, however, the identity of the space group must be reviewed. Often the questioning is cursory and the reply intuitive and the assessment of the space group made from inspection of the diffraction pattern is accepted as correct. In this situation, the analyst is in effect saying that he can see no justification in reducing the symmetry and increasing the number of independent variables.

An important article by Schomaker & Marsh (1979) explains the problems underlying attempts to refine a structure in a space group of too low a symmetry. The program *MISSYM* (LePage, 1987) will analyse a completed structure for non-space-group symmetry but, even when this is detected, its rôle will still need evaluating. Baur & Tillmanns (1986) looked again at the problem of deciding if the chosen symmetry is too low and gave a list of recommendations on how to avoid using unnecessarily low symmetry. Their discussion of structures taken from the literature shows that some of the problems arise mainly from the use of inadequate software for the detection of the Laue symmetry. Other problems are of a more delicate and perhaps even philosophical nature. Those concerning the detection of a missing centre of symmetry are of particular interest.

The inverse problem is also important in the treatment of some difficult refinements. What clues, the procedure having started with a space group of high symmetry, suggest that a lower symmetry is really required? The most insistent must be bulk physical effects measured, where possible, on the same crystal as used for the X-ray analysis. Second-harmonic generation and piezoelectric effects are the most convincing, though their failure is no evidence for a centre of symmetry. In fact, knowledge of the physical properties or provenance of the material will often have encouraged the analyst to start the modelling in a low-symmetry space group.

For other cases, the most likely clues come from apparent disorder. This may appear as sought-for atoms being found at less than their expected electron density in Fourier syntheses, as additional unexpected atoms appearing, as unacceptably short nonbonded contacts, as 'novel' (usually short) bond lengths or as high or exceptionally nonsymmetric anisotropic temperature factors. Other symptoms include crystallographic R factors much higher than the merging R factors. We must note in passing that these two quantities are not really related. The merging R factor is a measure of the self-consistency in the data; the crystallographic R factor (and especially the weighted or Hamilton R factor) are measures of the fit of the model to the observations. The *central limit theorem* can come to our aid when the data have large random errors but are copious. Under these conditions, the crystallographic R on merged data will be less than the merging R. Dunitz (1979) has given an analysis showing that the weak reflections may be particularly sensitive to the absence of a centre of symmetry and so has strengthened the argument that weak reflections, even if they are not used in the initial refinement, must be kept available as witnesses for marginal cases.

(iii) Strategy for lowering symmetry

If the physical evidence requires a low-symmetry space group, then there is little to be gained by trying anything else. In most other situations, it is probably most instructive and cost effective to move cautiously towards lower symmetry. For example, if one of the disquieting features of the high-symmetry model is large anisotropic thermal parameters for some atoms, these atoms can be replaced by two (or more) part atoms, as described above. A good plotting program will enable the analyst to see if a 'split' atom is a reasonable hypothesis. Least-squares refinement of any dramatically changed model should recommence gently and, in this case, a single cycle refining only U_{iso} of the two half-atoms would be suitable. In subsequent refinement of the positional parameters, it may be found to be necessary to restrain the distances of the partial atoms from neighbouring atoms if there are well known values for the bonds, or perhaps restrain the inter-partial-atom separation. Often, the restraints can be relaxed or removed once the refinement has stabilized. If the model, R factors and discrepancy distributions now look satisfactory, the refinement can be regarded as complete.

If some or all of the symptoms of lower symmetry persist, then the move must be made. Lowering of symmetry can often be achieved by removing either one of two operators, the second disappearing when the first goes. For example, in $P2/m$, removing either the 2 or the m also removes the centre (though removing the 2 leaves an achiral space group). At the same time, additional atomic coordinates must be generated (to replace those originally generated by the symmetry operator) and sometimes an origin shift or change of axes must be applied if the remaining operators are to

be located as shown in *International Tables for Crystallography* (1992). When more than one operator is available for generating the additional atoms, the one that produces a model of physical or chemical significance should be chosen. This should not of course affect the crystallographic computations but might make it easier for the analyst to understand what is happening if the refinement fails to proceed well. In general, it would be wise not to change the crystal axes even if this means working in a nonstandard space group, since the change will require at least re-indexing of the reflection data, reorientation of the cell parameters and transformation of the cell parameter variance–covariance matrix and orientation matrix. In addition, the relationship with the original model will become obscured. There is, in fact, every justification for working with a nonstandard setting of a space group if it makes human understanding of the situation easier and less error prone. (I remark in passing that, for dramatic changes of space group, say $R3c$ to $C2/c$, great care needs to be taken and the analyst should confirm that the transformed structure gives substantially the same R factor as the original structure and that the data properly cover the unique part of reciprocal space.)

The symmetry having been lowered and the refinement pursued to convergence, it is of course necessary to evaluate the solution. Continuous monitoring of the literature has revealed that a significant number of structures are refined in the wrong space group and most common amongst these are cases where the symmetry is too low. A common characteristic of too low a symmetry is that chemically equivalent bond lengths differ by many standard deviations but their mean value is quite normal. As noted above (Dunitz, 1979), it is the weak reflections that may hold the crucial evidence for marginal deviations from centrosymmetric symmetry.

(iv) Twinning

The major difficulties that arise when working with twinned materials are:

Determination of unit cell and space group. These problems are particularly severe when there are approximately equal amounts of each twin component in the crystal. When this is not the case, it is often possible to pick out a consistent lattice on a set of Weissenberg or good precession photographs.

Data collection. It is generally easiest to work with a set of data in which each observation is either a fully resolved reflection or a total integration over all components falling close to the basis reflection. Partially overlapping reflections lead to uncertainties in the interpretation of the integrated intensity (Rae, 1987).

Structure solution. Unless some estimate of the relative amounts of the twin components can be used to provide an approximate resolution of overlapping reflections, this stage is likely to prove the most difficult.

Least-squares refinement. Once a reasonable trial model has been obtained, there are a number of programs

available for its refinement. All of the problems that can affect a single-crystal refinement can also beset the analysis of twinned crystals.

A particular form of twinning that appears to be quite common and is easily dealt with but sometimes overlooked is inversion twinning of chiral materials obtained by spontaneous partial resolution of racemic mixtures (Flack, 1983). In this case, the material crystallizes in a noncentrosymmetric space group but the crystal contains regions of each isomer. All the reflections exactly overlap and the structure can usually be solved without unusual problems. However, since each reflection contains contributions from F_{o+} and F_{o-}, the anomalous differences may be small or even zero if there are equal fractions of each twin component. This fractional distribution can readily be refined by most modern program systems.

A synthetic example

Raising the symmetry of a refinement rarely poses any serious computational problems (there are sometimes practical ones, *e.g.* origin shifts as well as parameter averaging) but lowering the symmetry is usually much more problematic. Once the symmetry has been reduced, a number of refinement strategies are available. Some of the strategies have catastrophic outcomes and so should be avoided at all costs. Others lead with differing degrees of success to acceptable solutions. The analyst is of course restrained by the programs he has available. However, most modern programs contain some features that can, more or less simply and with more or less ingenuity, lead to satisfactory refinements.

The following example was devised (Watkin, 1986) to demonstrate some features of the different procedures described above. The known structure of *trans*-1,4-dimethylcyclohexane ($P2_1/c$, half a molecule in the asymmetric unit) was remodelled into the *cis*-1,4 isomer in $P2_1$ with a whole molecule in the asymmetric unit and structure factors computed to be used as 'observations' in the subsequent analysis. We now have to pretend not to know what the structure is, to doubt the systematic absences and erroneously to take the space group as $P2_1/c$. $h0l$ reflections with l odd were eliminated.

This pseudostructure was solved with *SHELXS*86 (Sheldrick, 1985) for half a molecule in the asymmetric unit. Isotropic refinement converged at 36%, at which point the methyl group, which had a large U_{iso}, was refined anisotropically. This refinement converged at 19%. Apart from the large R factor, other symptoms of a poor refinement were the short C–methyl bond length (1.41 Å) and the very aspherical methyl temperature factor (Fig. 2).

In accordance with the above suggestions, the methyl carbon was replaced by two half-methyl-carbon atoms, one at each end of the thermal-ellipsoid long axis, and this disordered model was refined isotropically. The final

R was 11%. To complete the example, the converged isotropic model ($R + 36\%$) was recovered and a second half-molecule generated, with all atoms isotropic, using the centre of symmetry, and the space group was reduced to $P2_1$. Several strategies were used to refine this highly symmetric starting model. Table 4 records the R factors and minimum and maximum C–C bond lengths for each refinement.

(i) Full matrix with Choleski inversion

Rollett remarked, 20 years ago, that some analysts were surprised that such a strategy often led to uncontrolled shifts in parameters or singular matrices (Rollett, 1970). Though the reason was described again in detail by Dunitz almost ten years later (Dunitz, 1979), the problem continues to surprise beginners.

The actual behaviour depends upon details of the least-squares program. Commonly, the matrix inversion proceeds *via* the Choleski method. If rounding errors are large, then the inversion may seem to have been successful in that it executes to completion. However, the shifts of parameters that were initially equivalent are likely to be large and are in any case valueless. The old strategy of using partial shift factors to try to contain the disruption is sometimes successful but not necessarily so. 10% of a calculated shift of 100 Å is still a big shift! [Shift factors could have a place in structure refinement in the hands of sensitive operators (Rollett, McKinlay & Haigh, 1976). These authors show that careful use of factors greater than unity can be used to accelerate a well behaved refinement. There seems to be no evidence from the literature that this strategy is in common use.] If the computation is to greater precision, smaller shifts may be computed for some parameters but eventually related parameters become pivots of the method and the latent singularities are discovered, and usually the corresponding shifts are set to zero. Thus, of a pair of originally related parameters, one is modified and the other is not. If the analyst is fortunate and the random shifts thus applied are sufficiently small and more or less in the right direction, further cycles of refinement of the

now nonsymmetric structure may proceed satisfactorily. This is rarely so and the refined structures usually show all sorts of curious anomalous geometries. In fact, the refinement generally 'blows up'. In this example, the R factor rises continuously and bond lengths become worthless.

(ii) Blocked matrix

The analyst, dismayed at discovering singularities in his full-matrix refinement, either refines the related fragments each in its own matrix block or refines alternate fragments in alternate cycles. These techniques differ slightly in detail but suffer from the same problems. In the first method, the structure factors and derivatives are all calculated from the same model and the matrix blocks are accumulated at the same time. Used with care in well behaved refinements, this is probably the most cost-effective method of refining 'medium-sized' structures (the definition of 'medium' depends of course upon the size and speed of the available computer). In the second method, the model is updated after each block of atoms has been refined and so different structure factors are available for subsequent blocks. For structures not showing pseudosymmetry, this latter technique is used as the basis for cascade refinement and has been shown to be very cost effective on computers with limited memory. If there are discrete molecular fragments and the analyst is not too concerned with intermolecular distances, the method can even be used as a crude procedure for fixing origins in polar directions. However, in the current situation, both techniques suffer from the same catastrophic disadvantage.

As discussed above, the failure of the full-matrix method (normally the safest method of refinement) is caused by the latent singularities arising out of 100% correlations (except for rounding errors) between the original model and the fragment generated by symmetry. Partitioning the matrix and discarding the off-diagonal elements that relate the two fragments does not cure the problem but only blinds the mathematics to it. As a result, the refinement seems to proceed satisfactorily and no singularities are observed. In fact, in some cases the refinement may appear to be chemically satisfactory, particularly when the two fragments are whole unconnected entities.

However, in the majority of cases, the refinement is unacceptable, with the fragments showing anomalous geometries but with average values close to the accepted ones. This often becomes particularly evident when the two fragments are part of the same molecule and are joined across the former symmetry operator. In blocking the matrix, the analyst has actually thrown out the information that will eventually, once the model has settled down, ensure correct geometries. The serious danger in this procedure is that the program cannot give the user any warning that all is not well so there is real risk of
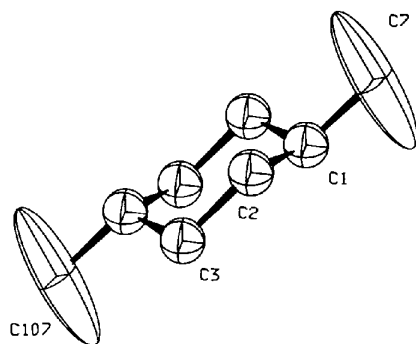

Fig. 2. Structure of pseudo-dimethylcyclohexane at a false minimum.

Table 4. *Course of refinements of a synthetic data set by different techniques, showing minimum and maximum C-C bond lengths*

Process used (see text)	R factor									C-C distances Minimum	Maximum
Full matrix (Choleski)	35.5	49.9	49.8	49.3	48.8	48.4	47.4	47.4	47.4	0.70	1.78
Two blocks	35.5	35.9	35.9	36.2	36.3	36.3	36.3	36.3	36.3	1.39	1.53
Full matrix (limited shifts)	35.5	35.4	35.2	33.4	26.6	14.9	4.6	1.7		1.52	1.52
Orthogonal coordinates	35.5	31.1	16.3	4.9	1.7					1.52	1.52
Antiriding constraints	35.5	34.7	15.6	3.9	1.7					1.52	1.52
Rigid-body constraints	35.5	35.2	26.2	11.9	2.9	2.1				1.52	1.53
Distance restraints	35.5	35.8	25.9	13.4	7.1	1.8				1.52	1.52
Eigenvalue filtering	35.5	35.5	35.5	35.5	35.5	35.5	35.5	35.5	35.5	1.41	1.52
Common sense	9.1	1.7								1.52	1.52

improperly refined structures being published, with little readers can do to recognize the situation. In the synthetic example, minor anomalies appeared in bond lengths but the structure remained approximately pseudosymmetric. It is only the high R factor, still over 30%, which makes us suspicious. If the pseudosymmetric structure had refined to, say, 12%, we might have accepted the model and assumed that there was something wrong with the data. *Blocking the matrix can never be recommended as a cure for singularities unless their source is well understood.*

(iii) Full matrix with shift-limiting restraints

In the case we are concerned with here, in which the original model has higher symmetry than the 'true' structure, the small (but otherwise uncontrolled) shifts permitted by shift-limiting restraints may mean that eventually the model drifts towards a correct one. The matrix then begins to contain terms computed from more or less correct derivatives, the shift-limiting restraints are over-ridden and the refinement proceeds to an acceptable solution. In this case, convergence occurs after eight cycles, at $R = 2\%$, with a model very close to that used to produce the pseudo-observations (Fig. 3).

The data *do* contain enough evidence for the original structure to be recovered and it is merely the unsatisfactory nature of the normal matrix (because of the over-symmetric model) that prevents proper refinement. However, this process, though semi-automatic, may be rather slow to start converging and, since it depends on fortuitous random shifts, cannot really be recommended

Fig. 3. Structure of pseudo-dimethylcyclohexane at a true minimum.

except as a method of last resort. As outlined below, there are generally better methods. If the program being used permits shift-limiting restraints to be expressed explicitly in the same units as the parameter concerned, restrictions of about 0.1 Å seem to be workable values.

(iv) Reparameterization to orthogonal coordinates

Pairs of new coordinates are defined for refinement by least squares, one being the sum and the other the difference of the corresponding positional parameters of the symmetrically related atoms. The matrix for a structure that is exactly pseudosymmetric is still singular but, once the structure is perturbed, the refinement quickly settles down. Shift-limiting restraints are still necessary for the initial computations but they can be fairly slack and limiting the shift to be not more than one unit cell seems to work well. This means that, once the crystallographic derivatives become meaningful, they are not unduly damped by the restraint. Convergence is achieved in five cycles.

(v) Anti-riding constraints

When a structure is refined in a high-symmetry space group, the analyst may be applying all sorts of constraints to the solution. These are implicit in the space-group symmetry operators, which the model for the total contents of the cell must obey. If the analyst believes that the bulk of his structure conforms (at the resolution of his data) to the higher symmetry and only some atoms are in more general positions, he is at liberty to refine the structure in the lower-symmetry space group and then re-impose selectively the relationships previously defined by the space-group operators. These relationships can be set up as constraints. In this case, we might believe that atoms C(1) to C(3) are very close to being centrosymmetrically related to atoms C(101) to C(103) and we could impose this belief as anti-riding constraints by setting the shift in $C(1, x)$ to be of the same size but opposite sign as that in $C(101, x)$ and so on for all the pairs of positional parameters of the atoms in the central ring. Thus, only one least-squares parameter is computed for the shift in $C(1, x)$ and $C(101, x)$,
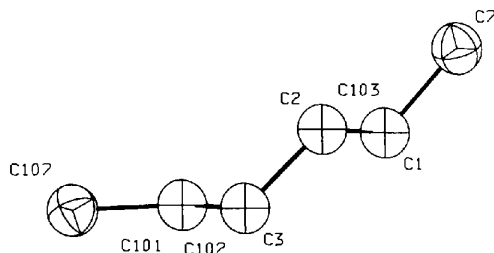
Table 5. *Eigenvalues and selected eigenvectors of the normal matrix for the centrosymmetric starting model*

(a) The 12 largest eigenvalues. The remaining 12 have values close to zero

2.68	2.53	1.37	2.45	1.52	1.57	1.72	2.15	2.01	1.93	1.97	2.02

(b) Components of the eigenvector corresponding to the first eigenvalue. Parameters are ordered x, y, z for each atom. Entries in the second row are centrosymmetrically related to the corresponding entries in the first row

0.32	−0.02	0.24	0.17	0.02	0.16	0.38	0.01	0.35	−0.10	0.03	−0.03
−0.32	0.02	−0.24	−0.17	−0.02	−0.16	−0.38	−0.01	−0.35	0.10	−0.03	0.03

(c) Components of the eigenvector corresponding to the 13th eigenvalue

0.20	0.63	0.04	0.15	0.03	0.02	0.04	−0.06	−0.06	0.14	0.11	−0.03
0.20	0.63	0.04	0.15	0.03	0.02	0.04	−0.06	−0.06	0.14	0.11	−0.03

another for C(1, y) and C(101, y) and so on for all the other pairs of parameters.

Because these are constraints, only three least-squares parameters are refined for each atom pair and after the matrix work appropriate shifts are applied to the related atoms. Formally, this is the same as working in the higher-symmetry space group, with the exception that the structure factors and derivatives have to be computed for all six atoms. The other atoms, the methyl C atoms, which through their temperature factors or anomalous bond lengths made us suspect lower symmetry, will of course be refined without this sort of constraint but may be the subject of reparameterization or shift-limiting restraints.

Once the refinement shows signs of stabilizing, the anti-riding constraints can be removed and, with mild shift-limiting restraints, full-matrix refinement can be used to finish off the task in a total of five cycles.

(vi) *Rigid-body constraints*

The centrosymmetric refinement yielded a core structure, the cyclohexyl ring, that made chemical sense. Another way to proceed to the lower-symmetry structure would be to refine the cyclohexyl ring as a rigid body with its current geometry and only the methyl groups as independent atoms. Replacing the 18 degrees of freedom of the core by only six rigid-body parameters reduces the number of ways in which the refinement can fall into ruin and ensures that the solution makes some chemical sense. As with the constrained refinement above, the rigid-body constraint should eventually be relaxed. Convergence was achieved in six cycles.

(vii) *Distance restraints*

The major problems with the centrosymmetric refinement were the anomalous temperature factor of the methyl group and its bond length from the ring C atom. The full-matrix refinement revealed its failure by both the R factor rising and the quite unacceptable C–C bond lengths. This suggests that another approach to a stable refinement might be to use bond-length restraints, both

for the C–methyl bond and also for the bonds in the ring. In this example, we can make a well informed guess at suitable values. In more general cases, theoretical arguments may not provide actual bond lengths but may indicate that, by symmetry, bonds should have similar lengths. This similarity may be applied as a restraint. Convergence was achieved in six cycles.

(viii) *Eigenvalue filtering*

This method provides excellent diagnostics as to why the full-matrix refinement failed. Table 5 lists the eigenvalues of the scaled normal equations and the eigenvectors corresponding to eigenvalues 1 and 13. If one remembers that the first four atoms are centrosymmetrically related to the second four, it is instructive to note that the signs of the second 12 components of eigenvector 1 are the opposite of the first 12, while they are the same for eigenvector 13. This reveals straight away that, while the sums of corresponding parameters are well defined, the differences are not, and explains why the standard refinement is unstable. It also explains why the re-parameterization described above is a useful technique. In that case, a rotation was applied in which the components of the relevant eigenvectors were exactly zero or $2^{1/2}/2$.

Fig. 4 (*ad hoc* plotting program) represents the variation of the minimization function, $M = \sum(w\Delta^2)$, as a function of the value of C(1, x) and C(101, x). (Note that this is a two-dimensional section through a 24-dimensional space. Changing any other parameter in the model requires us to look at the section parallel to the given section but displaced in the direction of the perturbed parameter.) The dotted line lies in the plane C(1, x) = −C(101, x), so that the minimum for the centrosymmetric structure must lie on this line. In this case, this minimum is also the local absolute minimum in the C(1, x)C(101, x) plane and lies at the bottom of a shallow bowl. Movement away from the dotted line causes the minimization function to rise, so that a centrosymmetric solution for these two parameters is best, even when noncentrosymmetric positions are available.

Fig. 5 is the corresponding contour map and shows that the minimum is well defined in the $C(1, x) + C(101, x) = 0$ direction (centrosymmetric) but not at right angles. Small perturbations along this direction will not affect the minimization function greatly and so are more or less equally acceptable.

Fig. 6 is a similar representation for $C(7, x)$ and $C(107, x)$. Again, the dotted line contains the minimum for the centrosymmetric structure. However, this is at a saddle point if the two coordinates are not required to vary synchronously; lower minima lie to either side of the line. The gradient of the surface perpendicular to the symmetry line should be zero for points immediately adjacent to the line, so there is no information in the normal matrix to tell the calculation to move the parameters off one way or the other. In the presence of rounding errors, a small gradient may be seen, indicating some distant minima, and large spurious shifts are computed. Eigenvalue filtering eliminates these spurious shifts. It differs from Choleski inversion, which can also trap large shifts, in that it recognizes special relationships between parameters and preserves these relationships.

In this example, the normal matrix contains no information at all about what shifts should be applied and the structure remains essentially unchanged after ten cycles of refinement. This is mathematically correct, though
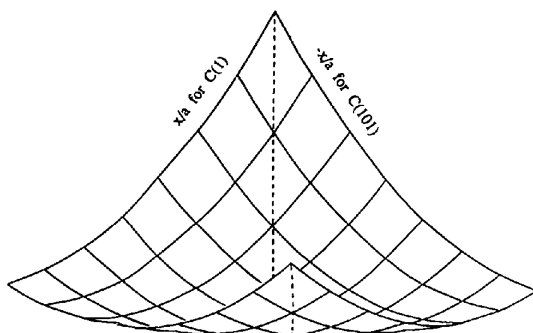
Fig. 4. Representation of a section of the minimization function for well resolved parameters in pseudo-dimethylcyclohexane.
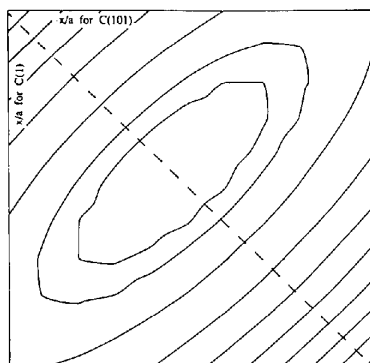
Fig. 5. Contour map corresponding to Fig. 4.

naturally disappointing for the hard-pressed analyst hoping for miracles. The standard Taylor expansion of the structure-factor equation and subsequent building of the normal matrix involves only the first-order derivatives. This helps to increase the range of convergence of the method, avoids saddle points and saves the expense of computing second derivatives, at the cost of a possibly nonquadratic convergence. However, we believe that inclusion of second derivatives would give eigenvalue filtering the information it needs to determine which parameters need to be perturbed and the correlation between these perturbations.

Once the model has been perturbed, the minimization surface (which is computed from the model) ceases to be symmetric and, if the perturbations are in the correct directions, the true minimum appears in the surface and the refinement proceeds correctly. We have not seen an example of the use of second derivatives in structure refinement.

(ix) Common sense

The split-atom refinement ($R = 11\%$) could have given us a clue about a possible model for the noncentrosymmetric space group. As with the ordered model, a second half-molecule could be generated using the pseudocentre, giving four half-methyl-carbon atoms (two at each end of the molecule). Taking a nonequivalent one from each pair and restoring it to full occupancy gives a model with asymmetric methyl groups. The R factor for this structure produced by trivial modelling techniques has a value of 9% and refines by any valid method in two cycles. The game, therefore, in all cases of near pseudosymmetry, is to provide the mathematics with as much evidence as possible drawn from sources external to the X-ray experiment and bearing upon noncontroversial issues in the analysis, and so permit concentration of the information contained in the X-ray data onto the real issues under investigation. The risk, of course, is of feeding in erroneous or prejudiced opinions.

Case histories

The following examples may serve to put the preceding discussion into practical perspective. They all have the
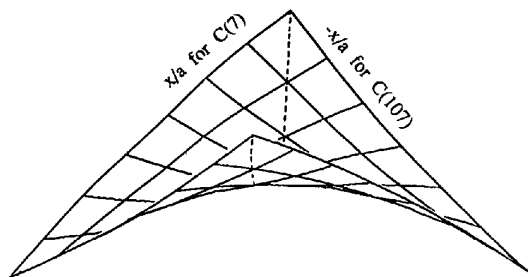

Fig. 6. Representation of a section of the minimization function for pseudosymmetric parameters in pseudo-dimethylcyclohexane.

same underlying theme – that least-squares refinement takes the model to the nearest local minimum and it is therefore necessary to ensure that the starting model is close to the global minimum. There is no mathematical technique to disclose whether we have discovered the global minimum (*i.e.* have a 'correct' structure already) so external information, such as chemical or physical respectability, is needed to guide us to the correct neighbourhood of the minimization space. In all the cases below, simple least-squares refinement converged at an incorrect model until additional methods were used to improve that model.

(i) Careful use of Fourier syntheses

This example illustrates the dangers of not looking carefully at Fourier syntheses, or ignoring maxima found by peak searches of Fourier maps.

The material $(\eta\text{-}C_5H_4CH_3)_2Mo_2S_4$ is in *Cmcm*, *C2cm* or *Cmc2$_1$*. The Mo and S atoms were located from a Patterson synthesis. A peak search on a Fourier synthesis yielded a phenyl ligand, in contrast to the expected methylcyclopentadienyl group. The peak search had shown two small peaks close to the phenyl group but, since these made poor bond angles with the central skeleton, they had been rejected. Careful consideration of the chemistry and discussions with the synthetic chemists indicated that, although the *R* factor stood at 11%, the ligand could not be as postulated, and the proposed phenyl-group atoms were removed from the trial structure and a new difference synthesis was computed in the plane of the ligand and carefully contoured (Fig. 7). The peaks formerly regarded as spurious were now treated as crucial and, using them as anchor points, it was easy to fit a doubly disordered methylcyclopentadienyl group to the density function (Fig. 8). This model was introduced for refinement, using riding temperature factors for corresponding atoms, tied occupancies for each partial molecule and restraints on the C–C bond lengths. Refinement of this chemically acceptable model was quite satisfactory and yielded a model that was stable even with the restraints removed. The final *R* was 3.5% (Prout & Daran, 1978).

(ii) Computer graphics and fragment mapping

This example was chosen to illustrate that the human eye can often recognize molecular fragments in an *E* map that will not develop by simple Fourier or least-squares techniques. As fully computerized methods develop, they will naturally become the first choice for computation but pattern recognition remains a process for which the human brain is highly optimized.

The chemist who made this key 23-atom intermediate in a complex organic synthesis had great difficulty in obtaining crystals of any sort, in spite of trying many different solvents. Eventually, poor crystals were obtained and, because of the importance of the material,

we were coaxed into performing an analysis. The poor crystals yielded poor data and we were distressed to see that the material crystallized in $P2_1/c$, with something between six and eight molecules in the asymmetric unit (from an estimated density). The small problem was now much larger. *SHELXS*86, *MULTAN*84 (Main, Germain & Woolfson, 1984) and *SIR*88 all failed, not unexpectedly, to give clear solutions, though all yielded
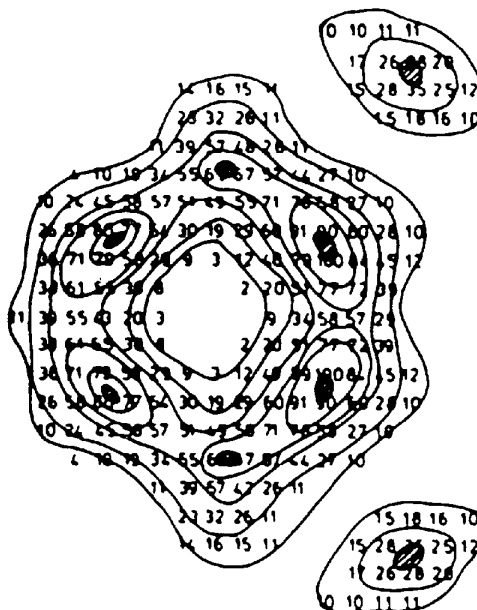

Fig. 7. Electron-density map in the plane of the disordered methylcyclopentadienyl ligand.
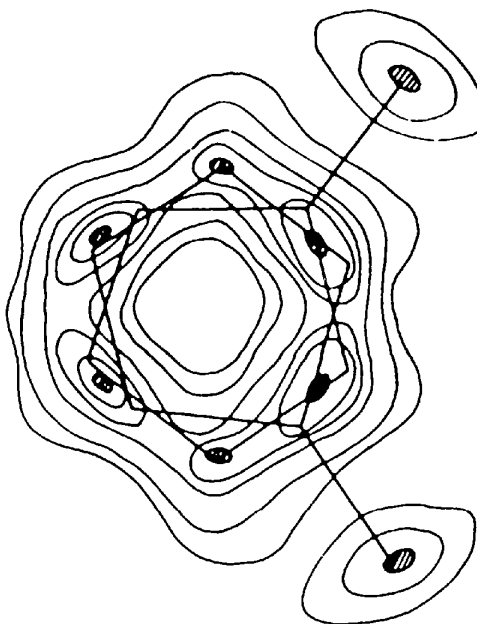

Fig. 8. Skeletons of two disordered ligands fitted to Fig. 7.

maps containing plausible fragments. None of these models could be developed by either simple Fourier methods or least-squares refinement. Fig. 9, the model from SIR88 (though probably other direct-methods models would have been equally useful), was introduced into a modelling package, CHEMX (Chemical Design Ltd, 1990), and the molecular fragments were cleaned up by deleting peaks that the user did not like. This yielded six partial molecules, which were transferred to CRYSTALS and mapped together to form a single chemically reasonable structure. This structure was then mapped back onto the sites of each of the partial structures and, by the use of rigid-body refinement and subsequently similarity restraints, the six molecules were satisfactorily refined, together with two molecules of dichloromethane of solvation. The high degree of
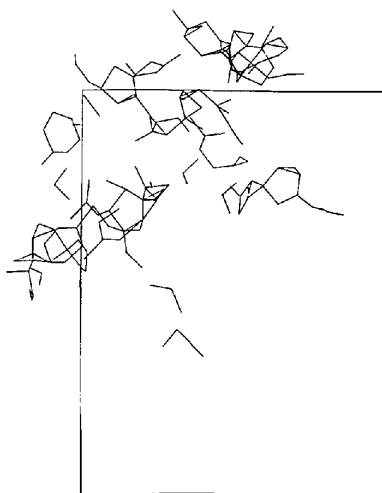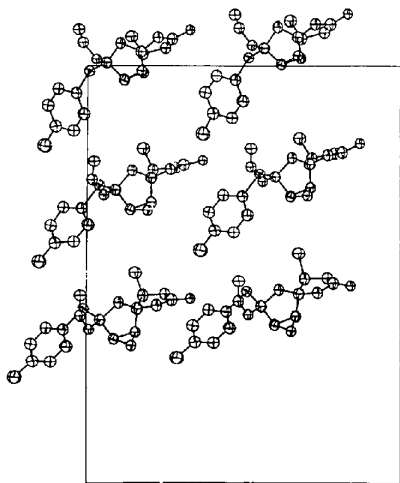
pseudosymmetry is evident in Fig. 10 and may explain the difficulties in solving and starting the refinement of the structure.

Even at convergence, the R factor was still 17%, but, in view of the large size of the structure and the poor quality of the data, the refinement was concluded. Some months later, a visitor, intrigued by the problem, carefully picked over the difference maps and was able to find various other solvent molecules (Einstein, 1988). The final R factor is 5%.

(iii) Regularization and restraints

This example was chosen to illustrate that, if it is evident what the solution should be, it is pointless and sometimes fruitless to let blind mathematics try and resolve uncertainties. Build the best model as quickly as you can, by whatever techniques are available. Do not start every structure assuming that there is no scientific precedent.

The material, a boron cluster, had three molecules in the asymmetric unit in $P\bar{1}$. The structure had been more or less solved by a colleague in another university and was of such a size that it could not be handled conveniently by his programs. Fig. 11 is the trial model sent to us. We were assured that, despite appearances, the ligands should be triphenylphosphine groups.

Automatic Fourier methods failed to develop the rest of the structure and least-squares refinement came to nothing. The phenyl groups were regularized and the poor boron cluster remodelled using coordinates from one of the good molecules. The structure was initially refined using rigid groups for each ring and cluster and then using bond-length, similarity and rigid-body thermal restraints. Most of the restraints were removed at the end of the refinement, which had converged to 8% with two molecules of solvation (Fig. 12).

Fig. 9. E-map peaks for the 138-non-H-atom structure.

(iv) Thermal-parameter model building and reparameterized least-squares refinement

This example was chosen to illustrate that, just as positional parameters need to conform to a realistic model, so also do thermal parameters. There is little to justify the publication of thermal parameters that imply quite implausible bond stretching if the X-ray data are equally compatible with a more traditional model.

Fig. 10. Final structure corresponding to Fig. 9, showing six independent but pseudotranslationally related molecules. Solvent molecules are omitted.

Fig. 11. Ill-resolved trial model of the boron cluster compound.

It also demonstrates that careful examination of the unreasonable results of a 'free' refinement may indicate which parameters need substantial revision.

This material (Fig. 13) diffracted magnificently, giving 90% of the data observable with copper radiation with $I > 3\sigma(I)$. It was expected to be chiral, so the structure was solved in $P1$ using SHELXS86. The initial peaks are shown in Fig. 14. The obvious interpretation gave us much glee, since yet again the chemists' predictions seemed to be wrong. It appeared that the expected Cp* ligand had been replaced by a second alkaloid group. The structure refined to 17%, at which point the ruthenium was confirmed as being at a noncrystallographic centre of inversion. The space group was changed to $P\bar{1}$ and the structure displaced so that ruthenium was at the origin. This new model, with only half a molecule in the asymmetric unit, still gave an R factor of 17% but would not refine. Bond lengths were unacceptable and some temperature factors were very unreasonable.
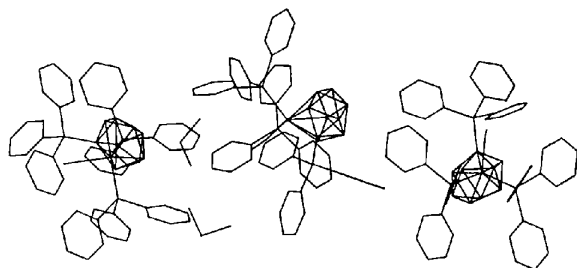
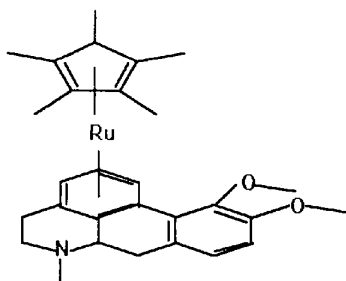

Fig. 12. Final structure of the boron cluster compound.

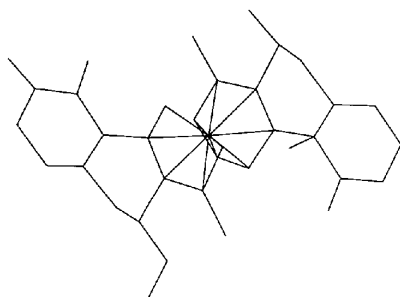
Fig. 13. Structural formula for the ruthenium compound.

Fig. 14. E-map peaks for the ruthenium complex.

The space group was returned to $P1$ and attempts to resolve the structure by Patterson methods and using the direct-methods program SIR88 yielded much the same pseudosymmetric structure. Even when one ligand was deleted and Fourier and difference syntheses computed, these consistently revealed the second alkaloid ligand. Least-squares refinement was restarted with bond-length, rigid-body thermal and shift-limiting restraints, building the normal matrix from parameters corresponding to sums and differences of pseudocentrosymmetrically related parameters. The R factor fell only a little but the SNOOPI (Davies & Prout, 1981) plots (Fig. 15) showed differences in the temperature factors of the two ligands. Note that the bond-length restraints, by holding the atoms at chemically reasonable positions, encourage the temperature factors of atoms far from correct positions to rise, rather than letting the positions and thermal parameters adjust to a mutual compromise. Which of the two ligands decayed was only a matter of chance but was not important at this stage, as no attempt was yet being made to determine absolute configuration.

Atoms with large temperature factors were deleted from the model, together with any that were then disconnected and remote from the Ru atom. This left the whole of one alkaloid molecule, together with atoms C(2), C(3), C(13), C(15), C(17), C(18), C(19), C(20) and C(21) as part of a distorted Cp* group (Fig. 15b).

Structure factors were computed for the now asymmetric model and Fourier syntheses computed. The alkaloid ligand developed well but the Cp* group was still poor and would not refine. An idealized Cp* was built and refined isotropically first as a rigid group and then with bond-similarity restraints. Subsequent refinement of the anisotropic temperature factors for the Cp* failed, with two methyl and one ring atom going non-positive definite. A TLS computation for the remaining seven atoms gave a rigid thermal model from which temperature factors of the three poor atoms could be predicted. This hand-crafted structure was finally refined (Fig. 16) to an R factor of 3%. The only restraints retained were bond-stretching and temperature-factor-similarity restraints on the Cp* group and the PF_6^- counterion.

The future

For small molecular structures (less than 100 atoms), there is little incentive to develop alternatives to normal matrix least-squares methods for parameter optimization of fundamentally correct structures, since the speed of computers (in a given price range) increases yearly. For larger structures, there are some potential improvements and the conjugate-gradient methods being applied to proteins may be useful for large 'small' structures. The main problem with conjugate-gradient methods is that simple implementations cannot take account of high correlations between parameters so refinement of

'difficult structures' may be slow or may even fail. An algorithm recently published for phase refinement and giving orders-of-magnitude increases in speed could be applied to parameter refinement and may change this situation (Main, 1990). Alternatively, direct methods for the solution of sparse matrices may become effective. The nature of crystallographic observational or normal matrices means that in general the distribution of significant terms can be predicted and controlled so computationally efficient sparse equations can be constructed and solved. (Note that in general the inverse of a sparse matrix is a dense matrix so solution by simple inversion

is not efficient.) By careful inclusion of important off-diagonal terms, good convergence properties might be achieved without undue computational expense (Duff, Erisman & Reid, 1989).

The developments in parallel and array processors and in transputers will speed up least-squares calculations. Existing programs will probably have to be re-organized to make maximum use of these machine architectures, though the best of modern optimizing compilers will do excellent work on well written code.

However, what are most urgently needed are automatic mechanisms for increasing the range of convergence of refinement processes so that poor (but essentially correct) initial trial models will not be discarded but be brought within the range of traditional methods. Molecular mechanics and simulated annealing have done this for proteins (X-PLOR; Brünger, 1990). Perhaps some alternative to molecular mechanics will be developed as a mechanism for generating the pseudorandom perturbations needed for simulated annealing to be applicable to nonproteins.

Our experience with the structure solution of cimetidene from powder diffraction data showed how easy it can be to discard a useful starting model because the molecule assembly routines do not yield a diagram recognizable as a chemical structure (Cernik et al., 1991). The problem is that, if the E-map peaks list is expanded to fill a complete unit cell, the diagram is likely to be too confused for visual analysis, especially if there are substantial numbers of spurious peaks. Equally, attempts to assemble a unique molecule automatically from the peaks list and using symmetry operators can often fail because spurious peaks generate false connectivity or absent peaks mean that the chance to reposition a fragment is lost. The explosion in the availability of inexpensive graphics systems will help manual examination of trial structures but a concerted effort is needed for the machine analysis of Fourier syntheses (Johnson, 1980, pp. 28.01–28.16).
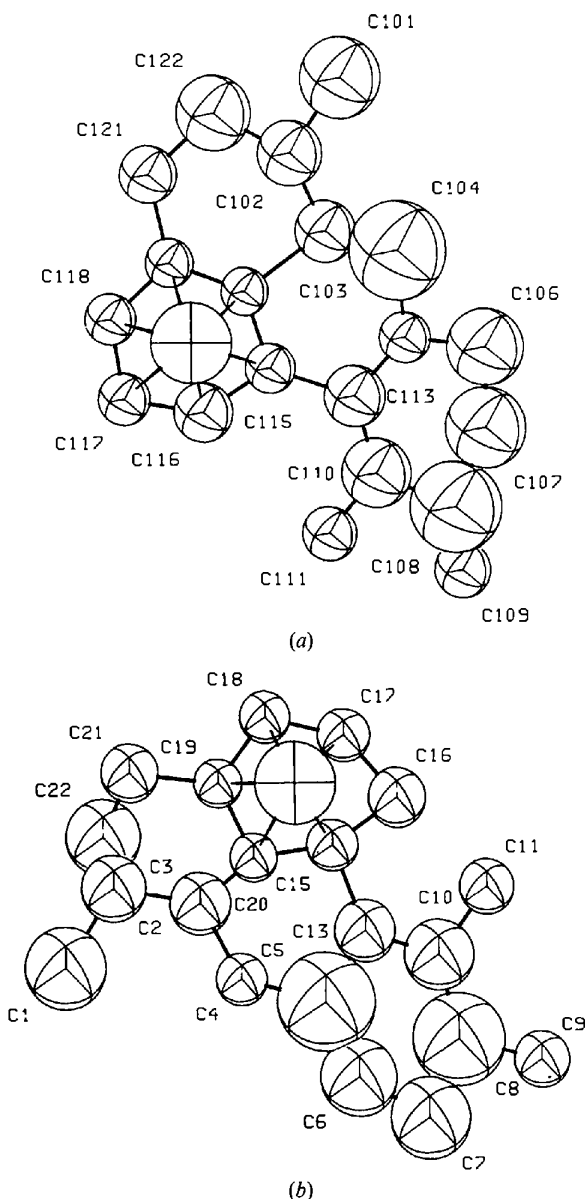
(a)

(b)

Fig. 15. The two tentative ligands after positional and isotropic temperature-factor refinement. Most of the atoms in (b) were eliminated to leave a partial Cp* ligand.
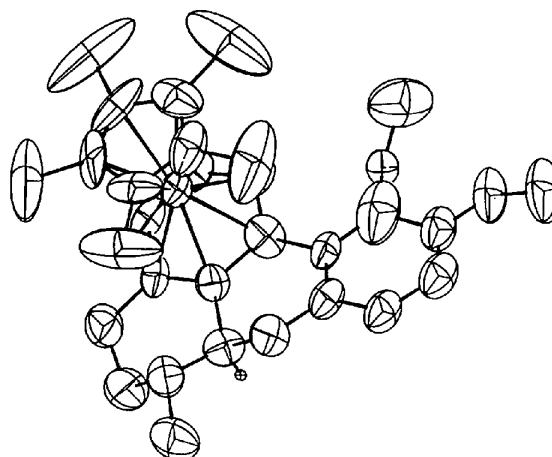

Fig. 16. Final structure of the ruthenium compound.

The really important issue is to develop computer programs that can evaluate complete or partial structures, examine them for inconsistencies and features in conflict with accepted physics or chemistry, and assess parameter values and e.s.d.'s with respect to the quality and quantity of the observed data. To replace, in fact, the experienced crystallographer.

Finally, this article has concerned itself almost wholly with molecular structures. There are very exciting challenges to be met in the field of inorganic and extended-lattice structures, where the problems of pseudosymmetry, superlattices and long-range order can be serious, space-group uncertainties are almost commonplace and there is not too much that can be implied beforehand about interatomic distances and angles.

Concluding remarks

In many ways, crystallographers are very privileged scientists, in that their experiments yield large amounts of (generally) good-quality data. Additionally, the numerical processing of this data often yields unambiguous solutions. Even so, the analyst must remain vigilant and sceptical of all calculations. If the residual discrepancies, $|F_o| - |F_c|$, are large with respect to the estimated errors in the data, then the model must be re-assessed, or the nature of the errors must be more closely investigated.

If the final model is at variance with accepted ideas, if it is 'novel', it becomes important to verify that a less-unusual model is not equally compatible with the data (i.e. gives the same weighted residual). This is perhaps most conveniently done by using the expected features of the model as target values for restraints and verifying that the X-ray residual does not increase significantly on refinement with these restraints. There are currently no programs that automatically generate acceptable models, especially in cases of disorder, so the analyst is still required to use imagination and experience in the interpretation of difficult structures.

We can expect some imminent generation of computer programs to make these checks for us and to have some knowledge of physics and chemistry. If the Crystallographic Information File (CIF) does become widely accepted as an archiving medium and the full list of required information is filled in reliably by the users and their programs, there is a real potential for machine screening of deposited results. For the moment, we have to depend on care, professionalism and, ultimately, good refereeing.

Most of the procedures described in this article can be found in a variety of current program systems. Almost all (except simulated annealing) are in the program CRYSTALS, which is available from the author.

References

ALTOMARE, A., CASCARANO, G., GIACOVAZZO, G. & GUAGLIARDI, A. (1993). J. Appl. Cryst. 26, 343-350.
ATKINSON, A. C. (1985). Plots, Transformations and Regression. Oxford: Oxford Science Publications.
BAERLOCHER, C., HEPP, A. & MEIER, W. M. (1977). DLS. ETH, Zurich, Switzerland.
BAUR, W. H. & TILLMANNS, E. (1986). Acta Cryst. B42, 95-111.
BENNETT, M. J., HUTCHEON, W. L. & FOXMAN, B. M. (1975). Acta Cryst. A31, 488-491.
BLAKE, C. C. F., PULFORD, W. C. A. & ARTYMIUK, P. J. (1983). J. Mol. Biol. 167, 693-723.
BLESSING, R. H. (1987). Crystallogr. Rev. 1, 3-55. ·
BRÜNGER, A. T. (1990) X-PLOR Manual. Yale Univ., USA.
BUSING, W. R., MARTIN, K. O. & LEVY, H. A. (1962). ORFLS. Report ORNL-TM-305. Oak Ridge National Laboratory, Tennessee, USA.
CERNIK, R. J., CHEETHAM, A. K., PROUT, C. K., WATKIN, D. J., WILKINSON, A. P. & WILLIS, B. T. M. (1991). J. Appl. Cryst. 24, 222-226.
Chemical Design Ltd (1990). CHEMX. Chemical Design Ltd, Oxford, England.
COUSSON, A., NECTOUX, F. & RIZKALLA, E. N. (1992). Acta Cryst. C48, 1354-1357.
DAVIES, E. K. & PROUT, C. K. (1991). SNOOPI User's Manual. Chemical Crystallography Laboratory, Oxford, England.
DIAMOND, R. (1976). Acta Cryst. A32, 1-10.
DIAMOND, R. (1984). In Methods and Applications in Crystallographic Computing, edited by S. R. HALL & T. ASHIDA. Oxford Science Publications.
DIAMOND, R. (1988). Acta Cryst. A44, 211-216.
DIAMOND, R. (1990). Acta Cryst. A46, 423.
DUFF, I. S., ERISMAN, A. M. & REID, J. K. (1989). Direct Methods for Sparse Matrices. Oxford Science Publications.
DUNITZ, J. D. (1979). X-ray Analysis and the Structure of Organic Molecules. Ithaca: Cornell Univ. Press.
EINSTEIN, F. (1988). Unpublished work.
FLACK, H. D. (1983). Acta Cryst. A39, 876-881.
FLACK, H. D. & SCHWARZENBACH, D. (1988). Acta Cryst. A44, 499-506.
FORSYTHE, G. E., MALCOLM, M. A. & MOLER, C. B. (1977). Computer Methods for Mathematical Computations. New Jersey: Prentice Hall.
FORTIER, S., CASTLEDEN, I., GLASGOW, J. I., CONKLIN, D., WALMSLEY, C., LEHERTE, L. & ALLEN, F. H. (1993). Acta Cryst. D49, 168-178.
GILL, P. E., MURRAY, W. & WRIGHT M. H. (1981). Practical Optimisation. London: Academic Press.
GOPINATHAN, M. S., WHITEHEAD, M. A., COULSON, C. A., CARRUTHERS, J. R. & ROLLETT, J. S. (1974). Acta Cryst. B30, 731-737.
HAMILTON, W. C. (1964). Statistics in Physical Science. New York: Ronald Press.
HAMOR, T. A., STEINFINK, H. & WILLIS, B. T. M. (1985). Acta Cryst. C41, 301-303.
HARRIS, G. W. & MOSS, D. S. (1992). Acta Cryst. A48, 42-45.
HIRSHFELD, F. L. & RABINOVICH, D. (1973). Acta Cryst. A29, 510-513.
HONG, W. & ROBERTSON, B. E. (1985). In Structure and Statistics in Crystallography, edited by A. J. C. WILSON. New York: Adenine Press.
International Tables for Crystallography (1992). Vol. A. Dordrecht: Kluwer Academic Publishers.
International Tables for X-ray Crystallography (1985). Vol. III. Birmingham: Kynoch Press. (Present distributor Kluwer Academic Publishers, Dordrecht.)
JOHNSON, C. J. (1971). ORJFLS, a Modified Version of ORFLS. Oak Ridge National Laboratory, Tennessee, USA.
JOHNSON, C. J. (1976). ORTEP. Report ORNL-5138. Oak Ridge National Laboratory, Tennessee, USA.
JOHNSON, C. J. (1977). ORCRIT. Oak Ridge National Laboratory, Tennessee, USA.
JOHNSON, C. J. (1980). In Computing in Crystallography, edited by R. DIAMOND, S. RAMASESHAN & K. VENKATESAN. Bangalore: Indian Academy of Sciences.
KABSCH, W. (1978). Acta Cryst. A34, 827-828.
LARSON, A. C. (1980). Computing in Crystallography, edited by R. DIAMOND, S. RAMASESHAN & K. VENKATESAN, pp. 11.01-11.07. Bangalore: Indian Academy of Sciences.

LAWSON, C. L. & HANSON, R. J. (1974). *Solving Least Squares Problems*. New Jersey: Prentice Hall.

LEPAGE, Y. (1987). *J. Appl. Cryst.* **20**, 264-269.

MACKAY, A. L. (1984). *Acta Cryst.* A**40**, 165-166.

MAIN, P. (1990). *Acta Cryst.* A**46**, 372-377.

MAIN, P., GERMAIN, G. & WOOLFSON, M. M. (1984). *MULTAN84. A System of Computer Programs for the Automatic Solution of Crystal Structures from X-ray Diffraction Data*. Univs. of York, England, and Louvain, Belgium.

MARSH, R. E. (1993). *Acta Cryst.* C**49**, 643.

MILLEDGE, H. J., MENDELSSOHN, M. J., O'BRIEN, C. M. & WEBB, G. I. (1985). In *Structure and Statistics in Crystallography*, edited by A. J. C. WILSON. New York: Adenine Press.

NARDELLI, M. (1993). *Acta Cryst.* A**49**, C54.

NICHOLSON, W. L., PRINCE, E., BUCHANAN, J. & TUCKER, P. (1982). In *Crystallographic Statistics*. Bangalore: Indian Academy of Sciences.

PRESS, W. H., FLANNERY, B. P., TEUKOLSKY, S. A. & VETTERLING, W. T. (1986). *Numerical Recipes*. Cambridge Univ. Press.

PRINCE, E. (1982). *Mathematical Techniques in Crystallography and Material Science*. New York: Springer-Verlag.

PRINCE, E. & BOGGS, P. T. (1992). In *International Tables for Crystallography*, Vol. C, edited by A. J. C. WILSON. Dordrecht: Kluwer Academic Publishers.

PROUT, C. K. & DARAN, J.-C. (1978). *Acta Cryst.* B**34**, 3586-3591.

RAE, A. D. (1973). *Acta Cryst.* A**29**, 74-77.

RAE, A. D. (1987). *RAELS87: a Comprehensive Constrained-Least-Squares Program*. Univ. of New South Wales, Kensington, Australia.

ROLLETT, J. S. (1970). *Crystallographic Computing*, edited by F. R. AHMED, S. R. HALL & C. P. HUBER, pp. 167-181. Copenhagen: Munksgaard.

ROLLETT, J. S. (1984). *Methods and Applications in Crystallographic Computing*, edited by S. R. HALL & T. ASHIDA, pp. 161-173. Oxford Science Publications.

ROLLETT, J. S., McKINLAY, T. G. & HAIGH, N. P. H. (1976). *Crystallographic Computing Techniques*, edited by F. R. AHMED, K. HUML & B. SEDLÁČEK, pp. 413-419. Copenhagen: Munksgaard.

SCHOMAKER, V. & MARSH, R. E. (1979). *Acta Cryst.* B**35**, 1933-1934.

SCHOMAKER, V. & TRUEBLOOD, K. N. (1968). *Acta Cryst.* B**24**, 63.

SCHWARZENBACH, D., ABRAHAMS, S. C., FLACK, H. D., GONSCHOREK, W., HAHN, T., HUML, K., MARSH, R. E., PRINCE, E., ROBERTSON, B. E., ROLLETT, J. S., WILSON, A. J. C. (1989). *Acta Cryst.* A**45**, 63-75.

SHELDRICK, G. M. (1976). *SHELX76. Program for Crystal Structure Determination*. Univ. of Cambridge, England.

SHELDRICK, G. M. (1985). *SHELXS86. Crystallographic Computing* 3, edited by G. M. SHELDRICK, C. KRÜGER & R. GODDARD, pp. 175-189. Oxford Univ. Press.

SLUIS, P. VAN DER & SPEK, A. L. (1990). *Acta Cryst.* A**46**, 194-201.

STEWART, J. M. & HALL, S. R. (1983). Editors. *The XTAL System of Crystallographic Programs*. Technical Report TR-1364.2. Computer Science Center, Univ. of Maryland, College Park, Maryland, USA.

WALKER, N. & STUART, D. (1983). *Acta Cryst.* A**39**, 158-166.

WATKIN, D. J. (1980). *Acta Cryst.* A**36**, 975-978.

WATKIN, D. J. (1986). *Crystallographic Computing*, edited by P. PAUFLER, V. GEIST & D. KLIMM, pp. 171-180. Karl-Marx Univ., Leipzig, Germany

WATKIN, D. J. (1988). *Crystallographic Computing* 4, edited by N. W. ISAACS & M. R. TAYLOR, pp. 111-125. Oxford Univ. Press.

WATKIN, D. J., CARRUTHERS, P. W. & BETTERIDGE, P. W. (1985). *CRYSTALS User Guide*. Chemical Crystallography Laboratory, Oxford, England.