A Filter Technique in Least- Squares Refinement of Crystal Structures

By C. Scheringer

Institut für Kristallographie der Technischen Hochschule, Aachen, Germany

The idea of filtering parameters by using the eigenvalues of the normal matrix, as first expressed by Diamond, is developed for positional parameters in least-squares refinement with X-ray data. With filtering, uncorrelated combinations of the parameter shifts are successively admitted to the refinement procedure in the order of their efficiency in fitting the data. Before the actual filtering is carried out two transformations of the normal equations are necessary: a congruent transformation in order to eliminate the contributions to the eigenvalues which arise only from the chosen description of the structure, and an orthogonal transformation in order to gain an uncorrelated set of parameter shifts. We use filtering as a means of enlarging the range of convergence. The efficiency of the filter technique is demonstrated by computed refinements of some trial structures of the known structures of phenol and 1,3,5-triphenylbenzene. Here filtering is applied to a set of positional rigid-body and molecular parameters. A mean total range of convergence of 1·0–1·2 Å has been found to exist for the two structures, which is about 0·1–0·2 Å larger than without filtering.

1. Introduction

In trying to fit chains of protein molecules to a given set of clearly recognized points of an electron density map, Diamond (1966) proposed a least-squares method of filtering a set of parameter shifts according to the magnitude of the eigenvalues of the normal matrix. The conventional parameters are transformed in order to diagonalize the normal matrix, and those 'eigenshifts' which are most effective in reducing the residual are admitted first, the remaining shifts being set to zero. In the course of the refinement the remaining shifts are successively admitted. Diamond calls this successive admission of parameters a 'sliding filter technique'. Filtering can enlarge the range and the speed of convergence; Diamond has tried to show this theoretically, and confirmed it numerically for his problem.

We raise the question of how much use can be made of the method in the refinement of crystal structures with X-ray data. In the refinement of single-atom parameters a certain type of filtering is commonly applied: scale factors and perhaps an overall temperature factor are refined first, then the positional parameters of the heavy atoms, and then those of the light atoms. The refinement of temperature factors becomes meaningful only when the positional parameters are almost correct, otherwise the respective atoms tend to be 'extinguished' through abnormally high temperature factors. Thus there is a 'natural' order for the admission of parameters.

In trying to apply Diamond's (1966, § 2.2.3) theory of filtering to the refinement with structure-factor data, we found that this theory was not correctly established. Diamond made physical assumptions which do not apply, and at the same time sidestepped a basic dif-

ficulty which is inherent in every filter technique based on the magnitude of the eigenvalues of the normal matrix. In this paper we shall solve the problem for the refinement of positional parameters with X-ray data. We shall deal with filtering only as a means of achieving a maximum range of convergence. Diamond (1958, 1966) also used filtering for other purposes.

2. Basic principles of filtering

A filter procedure, based on the magnitude of the eigenvalues of the normal matrix, is subject to the following difficulty: One may transform the given set of parameters at will, since this only results in another description of the least-squares problem (the parameters of which are not always physically concrete) and does not affect convergence. (For example, the positional parameters may be expressed in lattice units or Å.) The spectrum of the eigenvalues may, however, completely be altered by such a (congruent) transformation. Hence the filter procedure can be influenced at will by choosing a corresponding description of the structure. (It is, for example, always possible to transform a given set of parameters in such a way that the normal matrix becomes equal to the unit matrix. In this case filtering would no longer be possible at all.) Thus the question arises: Which description of the least-squares problem has to be chosen that a filter procedure, based on the magnitudes of the eigenvalues of the normal matrix, becomes a physically meaningful process, and so enlarges the range of convergence?

The answer to the above question would be trivial if the polydimensional parameter space were a metric space, in the sense that a metric fundamental tensor exists. (By 'exist' we mean that the components of the

metric tensor can in principle be determined experimentally, in a similar way to determining, for example, the lattice constants of a crystal. We do not mean that the metric tensor can be formally defined, which is always possible.) For a filter procedure one would have to determine the metric of the parameter space, and then transform to an orthonormal metric. The normal matrix would then be a symmetric tensor, and filtering would mean selecting the principal components of this tensor.

Since there is no metric in parameter space which can be determined experimentally–except for the metric of the unit cell – we have to apply a condition which can only be derived from the particular type of data that have been collected experimentally. The condition is that the physical quantity which is relevant in the data-collection experiment is expressed in a uniform scale (in the space in which it can be measured). Then the parameters which are attached to that physical quantity, and which occur in the same physical space, will also be expressed in a uniform scale. We may call this the 'orthonormal condition'. (For X-ray data the relevant physical quantity is the scattering power of the atoms, and the parameters of the atoms – or groups of atoms – are in the first instance attached to and determined *via* the scattering power.) The 'orthonormal condition' implies that the set of parameters which is to be filtered is physically homogeneous (for example, that all parameters are positional parameters). This is also a formal requirement, as in the filter procedure a linear combination of the parameter shifts is formed. When the 'orthonormal condition' is obeyed the diagonal elements of the actual normal matrix, and the eigenvalues which are to be filtered, tend to become large because of (1) the contributions of the relevant physical quantity (for X-ray data because of the contributions of the atomic scattering factors) and (2) the amount and weights of the data used. Thus the 'orthonormal condition' is sufficient to put the filter technique on a physically sound basis.

From the above considerations we may now criticize Diamond's (1966, § 2.2.3) theory of filtering. Diamond assumes that a metric exists in parameter space, and further, that this metric is orthonormal. This is obvious as Diamond speaks of a 'rotated set of axes' and of 'principal axes of an ellipsoid' in parameter space. Diamond calls the square root of the sum of the squares of the eigenshifts a 'distance' in parameter space, and he relates this 'distance' to the 'strain energy imposed on the structure'. The assumption of a metric in this sense relieves Diamond from considering the nature of his experimental data; the particular type of data has not been taken into account in his theory. Furthermore, the set of parameters is not required to be physically homogeneous. Therefore Diamond's theory appears to be completely general and applicable to every least-squares problem. Diamond himself applied his theory to two different kinds of least-squares problem (*cf.* Diamond, 1958, 1966). In his model building procedure

he uses angular parameters in units of radians. He does not explain what he understands conceptually (as opposed to mathematically) by a 'distance' in parameter space and how, in principle, he would measure it. We conclude that Diamond's theory is not correctly established as his basic assumption of a metric in parameter space does not apply.

Recently Bruton & Woodward (1967) have also described the filter procedure for the determination of force constants. These authors also do not recognize the problem of choosing the appropriate description of the least-squares problem, and in their treatment the normal matrix which is diagonalized and filtered, is set up with the 'natural parameters' of the problem. In Bruton & Woodward's discussion it is clearly assumed that the various directions of parameter space are expressed in a uniform scale, but a suggestion or even a proof that the 'orthonormal condition' does indeed hold with the parameters used is not given. Bruton & Woodward's numerical success, however, implies that in this case it does hold to a satisfactory extent.

As a second point of criticism we refer to Diamond's derivation of the convergence behaviour. Even if Diamond's assumption of an orthonormal metric in parameter space would hold, we could not accept his derivation of the 'dominant eigenvectors' possessing the largest range of convergence. In his proof Diamond assumed that small eigenshifts belong to large eigenvalues (*cf.* also Diamond § 2.2.3 'each s_i is proportional to $\lambda_i^{-1/2}$'). This, however, is not generally true because the right-hand side of the normal equations, whose magnitude cannot be estimated, destroys such a correspondence. We have examined the magnitudes of eigenvalues and eigenshifts in some computed refinements. In one case the largest eigenshift belonged even to the largest eigenvalue. In the mean over 32 cycles an eigenshift of medium magnitude belonged to the largest eigenvalue. With small eigenvalues we often found large eigenshifts, but by no means in every case. We concede that there may exist a slight tendency for small (large) eigenshifts to belong to large (small) eigenvalues. We shall discuss the convergence properties of the 'dominant eigenvectors' in § 3.

3. Filtering in structure refinement

For pure translation parameters the metric of the unit cell already determines the transformation of the initial set of parameters: One has to express the translation parameters in a Cartesian coordinate system. In this case we need not know the scattering power of the individual atoms explicitly, as the metric of the unit cell determines the transformation of both the scattering power and the parameter shifts. However, we need an explicit measure of the scattering power for the transformation of angular parameters.

We shall measure the scattering power in units of Z^2 ($Z=$ number of electrons of the atom). Z^2 is roughly

proportional to the total intensity radiated by an atom, and it is well known to be a rough measure of the efficiency of an atom in the determination of crystal structures (see *e.g.* Lipson & Cochran, 1953). This formal measure of the scattering power is only approximately true, but it is the only one which we are able to give in advance (as long as the structure is unknown). With units of Z^2 we neglect anisotropic scattering factors (thermal vibrations, *etc.*), and different scattering curves for the various types of atom. We first confirm that, neglecting these factors, the normal matrix displays the units of Z^2 for the positional parameters of the single atoms. It has been shown (Scheringer, 1965, 1968) that, with an ideal set of data (isotropic and of infinite density in reciprocal space), the normal matrix is a block-diagonal matrix consisting of 3×3-blocks. For a general metric of the unit cell, being defined by the metric tensor $\mathbf{g}$, the ith block, referring to the ith atom, is equal to $Z_i^2 . \mathbf{g}$. For an orthonormal metric of the unit cell it reduces to three diagonal elements of Z_i^2. (This incidentally shows that the scattering power, as it appears in the elements of the normal matrix, obeys the correct law of transformation: It transforms in the same way as the metric tensor of the unit cell.) For the translation parameters of a rigid group we obtain diagonal elements proportional to ΣZ_i^2 and zero off-diagonal elements, with an orthonormal metric of the unit cell.

For angular parameters, however, we cannot expect that the scattering power will appear in units of Z^2 in the diagonal elements of the normal matrix, because angular units of 2π are usually not comparable with Å units, and because the various atoms of a group are usually located at various distances r_i from the axis of rotation. We obtain the correct scale of the scattering power for an angular parameter, if we represent it as a translation parameter and apply the 'orthonormal condition'. For small angular shifts we treat the motion of the atoms on the arcs as linear. The mean distance (over all atoms of the group) $\bar{r}$ from the rotation axis is then determined from the condition that, for an ideal set of data, the diagonal element becomes proportional to ΣZ_i^2. If we introduce the formal angular measure $\varphi\bar{r}$, the diagonal element turns out to be proportional to $\Sigma r_i^2 Z_i^2/\bar{r}^2$. (This is shown in the Appendix. In the summation all those atoms which are not actually moved by the rotation – since they lie on or very close to the axis – must be excluded.) Thus the 'orthonormal condition' leads to the introduction of a new angular measure $\varphi\bar{r}$ for each angular parameter, where

$$\bar{r} = (\Sigma r_i^2 Z_i^2 / \Sigma Z_i^2)^{1/2} . \tag{1}$$

The transformations for parameters, which are reduced by equations of constraint (Scheringer, 1965), can be derived in a similar manner. These transformations are listed in the second part of the Appendix.

We sum up the steps which have to be carried out in using the filter technique as follows: With conventional parameters (expressed in lattice units and radians) we have the normal equations

$$A\varepsilon = B . \tag{2}$$

We now apply a congruent transformation $\mathbf{H}$ (the superscript T denotes the transposed matrix)

$$(\mathbf{H}^T\mathbf{A}\mathbf{H})\mathbf{H}^{-1}\varepsilon = \mathbf{H}^T\mathbf{B} \tag{3}$$

so that translation parameters are referred to an orthonormal basis and angular parameters are expressed in units of $\varphi\bar{r}$. By (3) we have $\varepsilon = \mathbf{H}\varepsilon_H$ (the subscript H denotes the transformed system). With oblique crystal systems $\mathbf{H}$ possesses a non-diagonal 3×3 block for the three translation parameters of a group. For angular parameters the corresponding section in $\mathbf{H}$ is diagonal and has values of $\bar{r}^{-1}$. Now we diagonalize the system (3) by determining the eigenvalues of $\mathbf{A}_H = \mathbf{H}^T\mathbf{A}\mathbf{H}$ so that

$$\Lambda = \mathbf{T}^T\mathbf{A}_H\mathbf{T} \text{ with } \mathbf{T}^T = \mathbf{T}^{-1} . \mathbf{B}_A = \mathbf{T}^T\mathbf{H}^T\mathbf{B} .$$

The solution in the diagonalized system is

$$\varepsilon_A = \Lambda^{-1}\mathbf{B}_A . \tag{4}$$

In (4) we apply a filter matrix $\mathbf{F}$

$$\varepsilon_{AF} = \mathbf{F}\Lambda^{-1}\mathbf{B}_A , \tag{5}$$

and finally obtain

$$\varepsilon_F = \mathbf{H}\mathbf{T}\varepsilon_{AF} \tag{6}$$

for the shifts of the conventional parameters which we shall actually use instead of the solution ε given by (2).

Filtering with respect to the magnitudes of the eigenvalues of the normal matrix means that we select 'eigenparameters' (we can only compute their shifts) according to: (1) the scattering power of an 'eigenatom' (which often does not exist as a physical unit), and (2) to what degree the actual set of data determines the position of that 'eigenatom' (since a summation over the derivatives of the measured structure factors is carried out in the elements of the normal matrix). A large eigenvalue means that, in the mean over all structure factors, the derivatives with respect to the 'eigenparameter' are large. Consequently, a change in a dominant 'eigenparameter' (*i.e.* an 'eigenparameter' which belongs to a large eigenvalue) gives rise to large changes in the structure factors. Therefore, the dominant 'eigenparameters' are those which reduce the sum of the squares of the errors most and are thus most efficient in fitting the data.

The increase in the range of convergence by first admitting the dominant 'eigenparameters' can be understood as follows. We assume a number of equal atoms whose parameters are correlated, and whose positions are determined equally well by the data. The orthogonal transformation, which diagonalizes the normal

matrix, then provides a number of equal 'eigenatoms', whose positions are no longer determined equally well by the data. The spectrum of how well the positions are determined corresponds to the spectrum of the eigenvalues. The geometrical part of the derivatives with respect to the dominant 'eigenparameters' will be large in the mean over all structure factors, and this means that the computed dominant eigenshift has a high probability of being correct with respect to convergence. As the eigenshift produces changes in the position of all the atoms (by virtue of the assumed correlation), the filter technique provides shifts which have a high probability of being correct. It is obvious that the filter technique becomes more effective the larger the spread of the spectrum of the eigenvalues. High correlation always causes a large spread. For two parameters, for example, which have equal diagonal elements $a_{11}=a_{22}$, the two eigenvalues are $\lambda_1=a_{11}+a_{12}$ and $\lambda_2=a_{11}-a_{12}$. (If there is only little correlation present, filtering can be done by hand, and then this technique does not offer a new tool.) This suggests that the filter technique may be applied to

(1) rigid-body, molecular parameters, and parameters which are reduced by equations of constraint (Scheringer, 1965)

(2) parameters of the single atoms which are highly correlated due to a small or unbalanced set of data.

The above considerations on the improvement of convergence are, however, only valid if we assume that a sufficient number of structure-factor functions behave monotonically in the desired range of convergence. If this does not hold, too many incorrect derivatives will be calculated with respect even to the dominant 'eigenparameter', which prevents convergence. Therefore the high-angle reflexions, whose structure-factor functions are most wavy, should be excluded in the initial stage of the refinement. This is particularly important when the filter technique is applied to rigid-body and molecular parameters, since the range of convergence can be much larger (about 1 Å) than the monotonic range of the absolute value of a high-angle structure-factor (e.g. 0·25 Å for Miller index $h=20$ and lattice constant $a=20$ Å).

4. Numerical results

In order to test how much the range of convergence is enlarged by filtering, we have written filter programs for the refinement of rigid-body and molecular parameters for the IBM 7090/7094 computer. The already existing programs without filtering (Scheringer, 1963a, 1965) were modified in the following way. For the positional parameters a full matrix is calculated, transformed and diagonalized. (A second full matrix which is simply inverted can be calculated for scale factors and an overall temperature factor.) The transformation

of the parameters may either be carried out on the derivatives or on the established matrix. With oblique crystal systems it is simpler to carry out the parameter selection by calculating the transformation matrix $\mathbf{H}$ and then $\mathbf{H}^T\mathbf{A}\mathbf{H}$ and $\mathbf{H}^T\mathbf{B}$; therefore we have adopted this procedure in our program. For the calculation of the eigenvalues and -vectors we have used the FORTRAN program BCHOW by D.W.Matula, in which a non-iterative method is applied. For the filter procedure we have written a special subroutine (see below).

For the successive admission of the eigenshifts we have developed a suitable criterion. The criterion which has been given by Diamond (1966) is not appropriate to structure refinement. (Furthermore, Diamond's criterion (i) is based on a claimed inverse proportionality between the magnitudes of the eigenvalues and eigenshifts which does not exist, cf. § 2.) Our criterion is a function of the R value used in statistical tests (cf. Hamilton, 1965)

$$R_G = \left(\frac{\Sigma\, w(F_{\text{obs}}-|F_{\text{calc}}|)^2}{\Sigma\, wF_{\text{obs}}^2} \right)^{1/2} .$$

Let R_G^i be the initial value of R_G (R_G of the first cycle), and R_G^r the value of R_G at which all parameters should be admitted to the refinement. (Values of R_G^r will be discussed below.) We define $S=R_G/R_G^i$ for the current cycle, and $S^r=R_G^r/R_G$. In any cycle we admit eigenshifts which belong to eigenvalues $\lambda \geq \lambda_F$ with

$$\lambda_F = \frac{3}{n}\, g(S) \times \text{trace}\, , \quad n \geq 4 .$$

The trace is calculated in every cycle. It varies roughly by less than 10% of its value throughout the refinement. n is an integral number, and is steadily increased by the program if necessary. The function $g(S)$ is a simple straight line $g(S)=a+bS$, which has to satisfy $g(1)=1$ and $g(S^r)=0$. Hence $b=1/(1-S^r)$ and $a=1-b$. The admission of eigenshifts is governed in the following way: In the first cycle n is increased by steps of 1 (and thus λ_F decreased) as much as necessary to admit a preset number N_S of eigenshifts. (Values of N_S will be discussed below.) Further eigenshifts are admitted by the decrease of $g(S)$ in the following cycles. Moreover, it is ensured that in any succeeding cycle the number of admitted eigenshifts cannot be decreased. This is achieved by increasing n as much as necessary. (If n becomes equal to 3000 all parameters will be admitted.) In order that the refinement may not come to a halt by deficiency of parameters an additional device is installed: In the 5th, 7th, 9th... cycle a further eigenshift is admitted, if in the 4th or 5th, 6th or 7th, 8th or 9th... cycle the decrease of $g(S)$ does not allow the normal admission of at least one further eigenshift. This is also achieved by increasing n. R_G^r and N_S have to be provided as data, all other quantities are computed. (We have also tried parabolas for $g(S)$ with

Table 1. Refinement of some trial structures of phenol, PH, and 1,3,5-triphenylbenzene, TR

For the trial structures the deviations Δ of the parameters from their values in the minimum are listed. PH: 250 low-angle reflections, 3 molecules, 17 rigid-body parameters. As one z parameter has to be kept constant (in space group $P2_1$), the deviations of the origin for the first molecule are somewhat smaller than those given in column 2. $R_G=0\cdot194$ in the structure minimum. TR: 200 low-angle reflections, 1 molecule, 8 molecular parameters. $R_G=0\cdot306$ in the structure minimum.

Trial structure	Δorigin (Å)	$\Delta\varphi,\Delta\theta$ (°)	$\Delta\varrho$ (°)	$\Delta\varrho^1,\Delta\varrho^2,\Delta\varrho^3$ (°)	Initial R_G	Average total shift (Å)	Maximum total shift (Å)	Filter			No filter		
								Number of cycles	Final R_G	Convergence	Number of cycles	Final R_G	Convergence
PH 1	1·10				0·548	1·04	1·10	12	0·199	yes	10	0·197	yes
PH 2	1·38				0·614	1·26	1·38	24	0·196	yes	24	0·196	yes
PH 3	1·56				0·643	1·42	1·56	25	0·198	yes	35	0·201	yes
PH 4	1·26				0·586	1·21	1·26	18	0·197	yes	20	0·532	no
PH 5	1·36				0·583	1·31	1·36	23	0·453	no			
PH 6		16	8		0·587	0·85	1·65	12	0·199	yes	8	0·197	yes
PH 7		18	10		0·610	0·95	1·85	31	0·198	yes	15	0·469	no
PH 8		18	12		0·612	0·95	1·85	28	0·230	yes	30	0·472	no
PH 9		18	16		0·613	0·96	1·91	20	0·505	no			
PH 10	0·51	16	12		0·593	1·11	2·09	11	0·196	yes	11	0·197	yes
PH 11	0·51	16	16		0·598	1·12	2·17	13	0·197	yes	12	0·195	yes
PH 12	0·51	18	10		0·610	1·20	2·21	13	0·197	yes	20	0·449	no
PH 13	0·56	18	15		0·620	1·24	2·38	28	0·197	yes	20	0·450	no
TR 1		12	8	10	0·665	1·10	2·12	19	0·308	yes	24	0·628	no
TR 2		10	10	10	0·650	1·00	1·99	15	0·309	yes	38	0·400	no
TR 3	0·51	8	8	15	0·657	0·98	1·89	21	0·309	yes	24	0·412	no
TR 4	0·26	12	6	10	0·666	1·09	1·98	17	0·307	yes	22	0·553	no
TR 5	0·21	11	8	10	0·654	1·05	1·94	17	0·308	yes	23	0·542	no
TR 6	0·18	10	10	10	0·645	1·02	1·93	20	0·558	no			
TR 7		10	10	20	0·651	1·01	2·19	20	0·558	no			

various slopes at $S=1$, but did not obtain any significant improvement.)

The increase of the range of convergence with filtering has been determined with the structure of phenol (PH) (Scheringer, 1963a,b) and of 1,3,5-triphenylbenzene (TR) (Farag, 1954; Scheringer, 1965). With phenol we have three rigid groups with 17 parameters to be refined, with 1,3,5-triphenylbenzene we have one molecule with 8 parameters (2 translation and 6 angular ones). Scale and temperature factors were kept constant at their previously refined values. For both structures reduced sets of data with 250 and 200 low-angle reflexions respectively were used (Table 1). In columns 2–5 of Table 1 the deviations of the parameter values from their correct values are given. Columns 7 and 8 give the mean and maximum distances in Å through which the atoms have been moved during the refinement. (The values of R_G are only roughly equal to the conventional R values, and for poor trial structures we often have $R_G > R$ by several per cent.)

We found that, with filtering, R_G decreases slowly in the first few cycles, even in case of correct convergence. Without filtering, R_G decreases more rapidly in the first few cycles, by virtue of the larger number of parameters being refined, but then often comes to a halt before the structure minimum is reached. Refinements which converge without difficulty do so faster without filtering (here filtering only acts as a brake.) Therefore it is important to filter only in that range in which convergence is obtained with difficulty or not at all. This largely determines the value of R_G^r, although a sharp numerical limit can rarely be given. In Table 1 we have used $R_G^r = 0.45$ for all refinements (except for PH6 with $R_G^r = 0.35$ and PH1 with $R_G^r = 0.40$). $R_G^r = 0.45$ is reasonable for rigid-body and molecular parameters because, in cases of groups having 6 atoms or more, the range of convergence for these parameters is already large without filtering and the R_G^i-values are usually > 0.45 (Table 1). (If there also are single-atom parameters, it is advisable to choose a smaller value of R_G^r, about $0.40 - 0.35$ depending on the problem.) The refinements PH1–4 and PH6 may illustrate the above (see Table 1): With PH1 ($R_G^r = 0.40$) and PH6 ($R_G^r = 0.35$) filtering acts as a brake; without filtering no false minimum is encountered. With PH2 at $R_G = 0.40$ a weak false minimum occurs and both filtering and no filtering achieve the same speed of convergence (24 cycles). With PH3 the false minimum is more pronounced; without filtering convergence is hardly possible and only obtained in 35 cycles (with filtering convergence is achieved in 25 cycles). With PH4 no convergence is obtained without filtering whereas with filtering the false minimum is ruled out.

The number of eigenshifts to be admitted in the first cycle, N_S, should be about $\frac{1}{4} - \frac{1}{3}$ of the number of parameters; perhaps somewhat less, if there are many parameters. The admission of too few eigenshifts retards the refinement, too many destroy the effect of filtering. We have used $N_S = 6$ for phenol (17 param-

eters) and $N_S = 3$ for 1,3,5-triphenylbenzene (8 parameters). $N_S = 3$ was found to be optimum with TR2. $N_S = 2$ decelerated the refinement by 4 cycles; $N_S = 4$ was too poor a filter and decelerated the refinement by 2 cycles. In any case we suggest that, before the actual refinement is started, one cycle should be calculated and the distribution of eigenvalues inspected in order to establish a good choice of N_S.

In order to apply the filter technique successfully the spectrum of eigenvalues should not be too narrow. (With PH we found values of $\lambda_{min}/\lambda_{max}$ ranging from 0.025 to 0.008; with TR values from 0.115 to 0.070.) This always holds with matrices of rigid-body and molecular parameters, since these parameters are always highly correlated. If there are heavy atoms present, filtering automatically provides that the positions of the heavy atoms will usually be determined first.

With the use of the filter technique the computing times are only slightly longer. For phenol, 20 cycles took 12·63 minutes without filtering and 12·75 minutes with filtering on the IBM 7094 (we used magnetic tapes to store structure-factor data). It is possible that with very large matrices (of order > 100) the computation of eigenvalues and eigenvectors takes more time than the simple matrix inversion.

From the results of Table 1 we conclude that with rigid-body and molecular parameters the filter technique provides an increase in the range of convergence of about 0·1–0·2 Å.

APPENDIX

Derivation of equation (1)

We formally introduce an angular measure $b = 2\pi \bar{r}$, for a given angular parameter. We then have for the diagonal element of the normal matrix of this parameter

$$a_{bb} = \sum_h w_h \left(\frac{\partial |F_h|}{\partial b} \right)^2$$

$$= \sum_h \sum_{i,k}^n \sum_{s,t}^3 w_h \frac{\partial |F_h|}{\partial x_{is}} \frac{\partial x_{is}}{\partial b} \frac{\partial |F_h|}{\partial x_{kt}} \frac{\partial x_{kt}}{\partial b}. \quad (A1)$$

$|F_h|$ is the absolute value of the hth structure factor, the coordinates x_{is}, x_{kt} are referred to a Cartesian basis, i and k denote the atom, s and t the directions in space; w_h is the weight of the hth structure factor, and n is the number of atoms which are moved by the parameter. With the assumptions of an ideal set of data and isotropic atomic scattering powers the terms for dif-

ferent atoms ($i \neq k$) are zero because of different geometric structure factors, and we are left with

$$a_{bb} = \sum_{i=1}^{n} \sum_{s,t=1}^{3} \frac{\partial x_{is}}{\partial b} \frac{\partial x_{it}}{\partial b} \sum_{h} w_h \frac{\partial |F_h|}{\partial x_{is}} \frac{\partial |F_h|}{\partial x_{it}} . \quad (A2)$$

For small changes of the angle we have

$$\frac{\partial x_{is}}{\partial b} = \frac{\Delta b_i}{\Delta b} \cos (\Delta b, s) = \frac{r_i}{\bar{r}} \cos (\Delta b, s)$$

Δb_i being the arc by which the ith atom would be moved, and r_i being the distance of the ith atom from the axis of rotation. $\cos (\Delta b, s)$ is the direction cosine of the directions of Δb and s. With the assumptions mentioned above we have

$$\sum_{h} w_h \frac{\partial |F_h|}{\partial x_{is}} \frac{\partial |F_h|}{\partial x_{it}} = kZ_i^2 \delta_{st} \quad (A3)$$

with $\delta_{st} = 1$ for $s = t$, and $\delta_{st} = 0$ for $s \neq t$, k being a constant. We now make use of the relation

$$\sum_{s=1}^{3} \cos^2(\Delta b, s) = 1 ,$$

and obtain

$$a_{bb} = k \sum_{i=1}^{n} r_i^2 Z_i^2 / \bar{r}^2 . \quad (A4)$$

We introduce the postulate $a_{bb} = k \sum Z_i^2$, and obtain equation (1).

Transformations for reduced parameters

Here we give the necessary transformations for parameters which are reduced by equations of constraint. First an orthonormal basis must be chosen for the unit cell, to which the coordinates of the single atoms are referred. Then the equations of constraint have the form

$$\mathbf{K}\varepsilon^{con}(\text{Å}) + \mathbf{K}^0 = \mathbf{O} , \quad (A5)$$

$$K^0 = r_{tk} - \left\{ \sum_{s=1}^{3} [x_{si}^0(\text{Å}) - x_{ks}^0(\text{Å})]^2 \right\}^{1/2} , \quad (A6)$$

$$K_{is} = -[x_{is}^0(\text{Å}) - x_{ks}^0(\text{Å})]\left\{ \sum_{s=1}^{3} [x_{is}^0(\text{Å}) - x_{ks}^0(\text{Å})]^2 \right\}^{-1/2}, \quad (A7)$$

$$\varepsilon^{con}(\text{Å}) = \mathbf{R}\varepsilon^{red} + \mathbf{R}^0 \quad (A8)$$

[*cf.* Scheringer (1965), equations (4) to (7)]. The elements of $\mathbf{R}$ are now dimensionless. The reduced parameters $\varepsilon^{red}(\text{Å})$ and the rectangular matrix $\mathbf{R}$ have to be transformed for the filter procedure according to

$$\varepsilon_G^{red} = \mathbf{G}\varepsilon^{red}(\text{Å}) , \quad \mathbf{R}_G = \mathbf{R}\mathbf{G}^{-1} . \quad (A9)$$

$\mathbf{G}$ is a diagonal matrix whose order is equal to the number of reduced parameters, and has an element $\bar{r}$, as defined by (1), for each parameter respectively. The sum in (1) has to be taken over all atoms which are moved by the parameter. r_i^2 is calculated according to

$$r_i^2 = (R_{xq}^i)^2 + (R_{yq}^i)^2 + (R_{zq}^i)^2 , \quad (A10)$$

where R_{xq}^i (R_{yq}^i, R_{zq}^i) is an element of $\mathbf{R}$ [$\mathbf{R}$ in equation (A8)], which relates the single-atom parameter $x(y,z)$ of the ith atom to the qth reduced parameter. (Equation (A10) becomes dimensionally correct if one divides r_i by the unit distance.) The matrix to be diagonalized for filtering is $\mathbf{R}_G^T \mathbf{A} \mathbf{R}_G$. The corresponding derivatives can be obtained directly from Scheringer's (1965) equation (13) using the elements of $\mathbf{R}_G$ and $x_{is}(\text{Å})$.

References

BRUTON, M. J. & WOODWARD, L. A. (1967). *Spectrochim. Acta* **23**A, 175.
DIAMOND, R. (1958). *Acta Cryst.* **11**, 129.
DIAMOND, R. (1966). *Acta Cryst.* **21**, 253.
FARAG, M. S. (1954). *Acta Cryst.* **7**, 117.
HAMILTON, W. C. (1965). *Acta Cryst.* **18**, 502.
LIPSON, H. & COCHRAN, W. (1953). *The Determination of Crystal Structures*, p. 207. London: Bell.
SCHERINGER, C. (1963a). *Acta Cryst.* **16**, 546.
SCHERINGER, C. (1963b). *Z. Kristallogr.* **119**, 273.
SCHERINGER, C. (1965). *Acta Cryst.* **19**, 513.
SCHERINGER, C. (1968). *Acta Cryst.* B **24**, 662.