STATISTICAL PROBLEMS ARISING FROM

CRYSTAL STRUCTURE ANALYSIS

by

PAUL WILLIAM REYNOLDS

Thesis submitted for the degree of Doctor of Philosophy

of the University of London

University College London

March 1990

ProQuest Number: 10610054

ProQuest 10610054

ABSTRACT

This thesis is concerned with the application of statistical techniques in the field of crystallography - a branch of science dealing with the structure, classification and properties of crystals - and an analysis of some of the associated statistical problems. We shall concentrate throughout on the estimation of atomic co-ordinates within the unit cells of crystals.

The science of X-ray crystallography will be introduced and a review of some of the existing methodology given. We shall then consider how statistical ideas may be used to improve this methodology.

We shall be particularly concerned with the area of sequential experimentation, in which the data collection process itself is modified as a result of analysing the data already collected. Sequential experimentation for improved efficiency in any particular crystallographic problem requires that decisions be made as to which additional data should be collected in order to achieve the desired objective. Ways of selecting suitable sampling strategies will be described, together with associated stopping rules. We will also describe methods for handling relevant prior information - e.g. structural information available in crystallographic data bases - and nuisance parameters, and procedures for dealing with the

inherent non-linearity of the crystallographic model, matrix updating and the recursive addition of data. The central problem of X-ray crystallography - the 'phase problem' - will also be analysed from a statistical perspective. Practical application of some of our ideas will be given.

Much emphasis is placed on non-linear parameter estimation problems such as those arising in crystallography. A review of relevant statistical work in this general field is undertaken, and geometry-based ideas of our own proposed. We concentrate on either seeking suitable re-parameterisations (in a sense which we define) or on seeking alternatives to the standard tangent plane approximation to the solution surface based on relevant curvature measures.

The thesis ends with a few relevant concluding comments and some ideas for further related statistical work in the area of X-ray crystallography.

INDEX

4

LIST OF FIGURES

7

LIST OF TABLES

ACKNOWLEDGEMENTS

CHAPTER 1

CRYSTALLOGRAPHY: AN INTRODUCTION FOR THE STATISTICIAN

1.1 Introduction to and a Brief History of the Development of X-ray Crystallography

To say something is crystalline means that the atoms or molecules of which it is composed are packed together in a regular (3-dimensional) manner; crystallography is the study of the structure of crystals. The purposes of crystallography can be split into two broad categories, namely

(1)   Identification of solid substances

(2)   Determination of atomic configurations.

It is with (2) that the statistician will be concerned, i.e. the principles underlying the progressive stages in the elucidation of internal structure.

There are 7 types of crystal forms. The division is based upon (3) imaginary axes passing through the centre of a crystal - their relative lengths and relative angles defining a "unit cell" (or unit of repeating pattern): the smallest group of atoms from which the whole crystal may be constructed forms the unit of pattern and may be associated with a lattice point. The statistician is primarily concerned with obtaining estimates of the co-ordinate positions of the atoms inside the unit cell, a

process referred to as structure determination. The determination of the full symmetry of a crystal structure is important (see Section 1.2); we will therefore need to consider point groups and space groups where

(1) Point groups are associated with symmetries of arrangements of atoms around a lattice point

(2) Space groups are associated with symmetries of the complete arrangement in the crystal.

There are 32 possible point groups and 230 possible space groups. Numerous crystallographic text-books cover the derivation of the crystal systems, symmetry, point groups and space groups; see for example [12].

Were we able to use visible light to look inside a crystal (and thus locate the positions of the atoms in the unit cell) the statistician's task would be greatly simplified; however, this is not possible, so an alternative method of structure determination must be found. For this, X-rays are used as discussed in Section 1.3. Before this, however, the structure factor equation - the single most important equation in crystallography - will be derived in Section 1.2; its derivation introduces many important concepts to be used throughout this thesis and will be given here in some detail. We will then be in a position to outline in Section 1.4 some of the associated statistical problems with which this thesis will be

concerned. Some relevant notation will be introduced in Section 1.5.

First, however, a brief history of the development of X-ray crystallography will be given, with emphasis being put on the statistical aspects covered to date. The science of X-ray crystallography can be said to originate with the discovery in 1912 by Laue that X-rays, which had been discovered 17 years previously by Röntgen, have wavelengths comparable to interatomic distances (see Section 1.3). This discovery paved the way for the determination of exact arrangements of atoms in crystals and the first successful structure elucidations soon followed.

Most early structure determinations were carried out by the method of trial and error. This consisted of postulating various atomic arrangements consistent with the known space-group symmetry and comparing the theoretical results derived from such structures with the actual results observed. The configuration that yielded the best agreement was taken to be the final structure - its correctness being confirmed by continuing to check the calculations for further observations if necessary.

The first practical uses [25], [43] of the application of Fourier series to the representation of crystal structures were made in 1925 in the derivation of the electron distribution in the sodium and chlorine atoms in rock

salt. Four years later, W.L. Bragg suggested that Fourier methods could be used in crystal structure determination. The relevant underlying theory is outlined in Section 2.3. Most original applications of the method, however, merely sought to refine structures that had been derived by the trial and error methods above.

Other methods of structure determination in widespread use in the inter-war years included approaches based on Patterson maps and so-called heavy atom methods (see Section 2.7). A landmark in the history of X-ray crystallography was reached in 1941, however, with the first application by Hughes [45] of least squares techniques in the refinement of crystal structures (see Section 2.2), which has since been widely adopted. Not long afterwards, the era of crystallographic statistics was born.

The founding father of crystallographic statistics is Arthur Wilson. Much of his early pioneering work concerned with the application of statistical methods in crystallography considered the statistical properties of the intensities of X-rays diffracted by a crystal (see Section 1.3). His inaugural result in this area is given in [71]. This was subsequently independently confirmed [38] and indeed extended [46]. In 1949, Wilson achieved a major breakthrough by deriving the ideal distribution functions of structure factors for all crystals [72]: as

16

we shall see in the following Sections, the structure factor is that term in the expression for the amplitude of our observable experimental quantities which involves the co-ordinates of the atoms within the unit cell. As a consequence of this work, the development of direct methods (see Section 2.6) of an overtly statistical nature was taken up by Karle and Hauptman [42], [49]. Then, as now, much of the energy channelled into crystallographic statistics sought to consider the properties of groups of related X-ray intensities in an attempt to tackle the phase problem of X-ray crystallography (introduced in Section 1.4).

Although the most active area in crystallographic statistics has been that of intensity statistics, other important areas covered include statistics of recorded counts, Wiener methods for electron density and alternatives to least squares. Useful references for the former two areas are [23] and [18] respectively; in conjunction with the latter area, use of Bayesian ideas [31] merits a special mention. This is a relatively new development in crystallographic circles, with only three relevant such applications [30], [32], [57] prior to 1980. Bayesian ideas will form an integral part of the theory of this thesis, however (see Chapter 3). Finally, we also note that a structure has recently been determined based on the theory of maximum entropy [36], thus indicating another area in crystallography currently rife for

statistical research.

Further comments on the history of the growth of
crystallographic statistics (up to 1976) may be found in
[62].

1.2 Derivation of the Structure Factor Equation

Much of this Section is based on Chapter 2 of [54].

We recall from the previous Section that the (space)
lattice of a crystal is the collection of points, each of
which represents the groups of atoms which, when repeated
at regular intervals, constitute the crystal. It is
assumed that each lattice point has the power to diffract
X-rays, and may consequently be regarded as a (point)
electron. The positions of the electrons can be denoted by
vectors $\underline{r}$ such that

$$\underline{r} = u\,\underline{a} + v\,\underline{b} + w\,\underline{c} \qquad (1.2.1)$$

where $\underline{a}$, $\underline{b}$, and $\underline{c}$ are the primitive translations of the
lattice and u, v and w are integers.

Now consider a parallel beam of X-rays of wavelength λ
falling on the lattice in a direction defined by the
vector $\underline{s}_0$ (of magnitude $1/\lambda$). By referring to Figure 1.2.1
where $\underline{s}$ (also of magnitude $1/\lambda$) is a vector defining an
arbitrary direction and $\underline{r}$ is the vector distance between _displacement_
two lattice points A and B, we see that the path

18

difference between the two scattered waves in direction $\underline{s}$
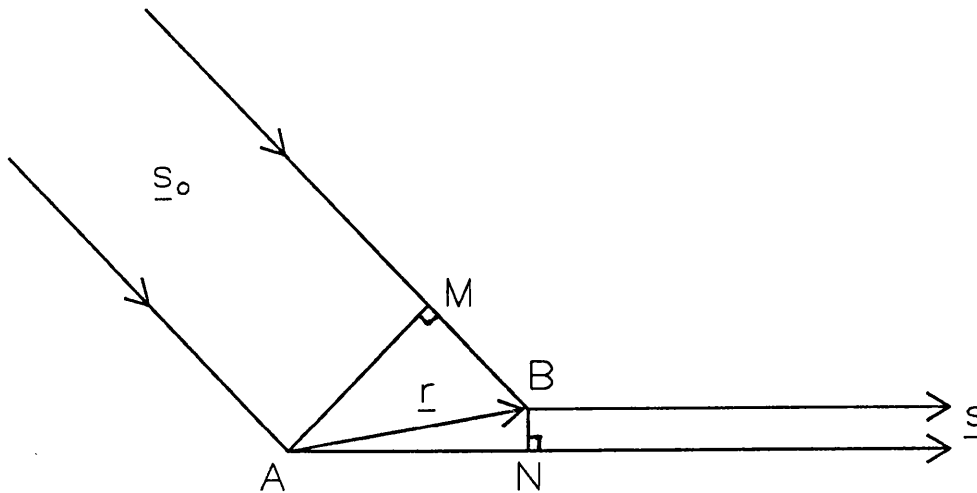
Figure 1.2.1: Scattering from two lattice points ($\underline{s}$ may not be in the same plane as $\underline{s}_0$ and $\underline{r}$)

is

$$AN - BM = \lambda(\underline{r} \cdot \underline{s} - \underline{r} \cdot \underline{s}_0) = \lambda \underline{r} \cdot \underline{S} \qquad (1.2.2)$$

where

$$\underline{S} = \underline{s} - \underline{s}_0$$

is called the scattering vector. In order that the waves scattered by A and B in direction $\underline{s}$ shall be in phase, and thus reinforce each other, the path difference should be equal to a whole number of wavelengths, i.e.

$$\underline{r} \cdot \underline{S} = (u\,\underline{a} + v\,\underline{b} + w\,\underline{c}) \cdot \underline{S} = \text{integer.} \qquad (1.2.3)$$

Equation (1.2.3) needs to hold for all integral values of u, v and w; this yields

19

$$\underline{a} \cdot \underline{S} = h$$

$$\underline{b} \cdot \underline{S} = k \qquad\qquad (1.2.4)$$

$$\underline{c} \cdot \underline{S} = l$$

where h, k and l are integers. These equations (1.2.4) are referred to as Laue's equations.

From the first two equations of (1.2.4) we see that

$$\frac{a}{h} - \frac{b}{k} \cdot \underline{S} = 0$$

i.e. the vector $\underline{S}$ is perpendicular to the vector $(\frac{a}{h} - \frac{b}{k})$. The latter vector, however, is in the plane of Miller indices hkl where the Miller indices of a plane are a set of whole numbers (hkl) such that the intercepts of the plane on the axes defined by $\underline{a}$, $\underline{b}$ and $\underline{c}$ are in the ratio

$$\frac{|\underline{a}|}{h} : \frac{|\underline{b}|}{k} : \frac{|\underline{c}|}{l} .$$

Similarly, it can be shown that $\underline{S}$ is perpendicular to the vector $(\frac{a}{h} - \frac{c}{l})$, which is also a vector in the plane hkl: we see that $\underline{S}$ is therefore perpendicular to this plane. But $\underline{S}$ is a vector in the direction of the bisector of the incident and diffracted beam, since the moduli of $\underline{s}$ and $\underline{s}_0$ are equal (see Figure 1.2.2). Thus this bisector is identified with the normal to the hkl plane.

20
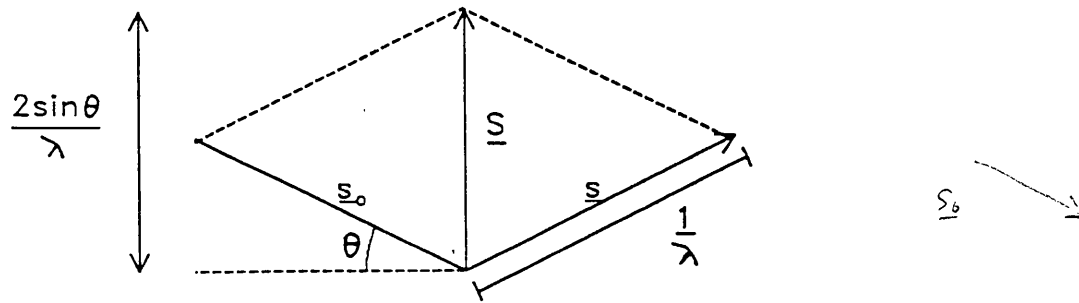
Figure 1.2.2: Relationship between $\underline{s}_0$, $\underline{s}$ and $\underline{S}$

This is the justification for the concept of each diffraction as a "reflection" of the rays from a particular hkl plane.

We now introduce the spacing d of the planes hkl; this is the perpendicular distance from the origin to the nearest plane and consequently can be expressed as

$$d = \frac{\frac{\underline{a}}{h} \cdot \underline{S}}{|\underline{S}|} \quad . \qquad (1.2.5)$$

By noting from Laue's equations that

$$\frac{\underline{a}}{h} \cdot \underline{S} = 1$$

and from Figure 1.2.2 that

$$|\underline{S}| = 2\frac{\sin \theta}{\lambda}$$

equation (1.2.5) reduces to give

$$d = \frac{\lambda}{2 \sin \theta}$$
(1.2.6)

which is Bragg's equation (the quantity n which sometimes appears in Bragg's equation is now absorbed in the integers hkl). Bragg's law says that reflections will be obtained when equation (1.2.6) is satisfied, the corresponding θ being called the Bragg angle of reflection.

In atoms the electrons occupy a finite volume and consequently the phase differences between rays scattered from different points in this volume have to be taken into account. For small angles of diffraction these phase differences are small, and the amplitude of scattering by an atom can be taken to be the sum of the amplitudes of the scattering by its individual electrons. This means that if the electrons at the lattice points are replaced by an atom of atomic number Z, then the expression for the amplitude of the scattered beam must be multiplied by the factor Z. As the angle of diffraction increases, however, the phase differences become larger, and thus the scattered beam becomes weaker, i.e. the factor becomes less than Z. The factor is called the atomic scattering factor f and is constant for a given angle of diffraction (assuming spherical symmetry for the atom). The curve of scattering factor against $(\sin \theta)/\lambda$ is called the

scattering-factor curve. Extensive calculation of these curves has been carried out for different atoms and the results can be found in the International Tables for X-ray Crystallography.

We now consider the temperature factor. At all temperatures atoms have a finite amplitude of oscillation (about 10^{13} per second) which is much smaller than the frequency of X-rays (about 10^{18} per second); consequently, to a train of X-ray waves, the atoms would appear to be stationary, but displaced from their true positions in the lattice. Thus, in producing a given X-ray reflection, atoms in neighboring unit cells which should scatter in phase, will scatter slightly out of phase, the total effect being apparently to reduce the scattering factor of the atom by an amount which increases with angle. If the thermal waves have a random phase relationship, the form of the variation of the scattering factor with angle can be calculated (with certain assumptions about the nature of the atomic vibrations). The result is that, if the scattering factor discussed above is called f_0, the factor f to be used in practice is

$$ f = f_0 \exp \left(-B \frac{\sin^2\theta}{\lambda^2} \right) \qquad (1.2.7) $$

where θ is the Bragg angle and B is the Debye-Waller factor (constant), if the mean square displacements are the same for all atoms and are isotropic. More generally,

23

the factor used (for atom r) is of the form $f_0 \exp(-T_r)$

where

$$T_r = h^2 B_{11}^r + k^2 B_{22}^r + l^2 B_{33}^r + hk B_{12}^r + hl B_{13}^r + kl B_{23}^r;$$

(1.2.8)

the six B_{jk}^r's are the temperature parameters for the rth atom which is assumed to be vibrating ellipsoidally.

Suppose now that the unit cell of a crystal contains R atoms, situated at points (x_r, y_r, z_r) $(r = 1..R)$ expressed as fractions of the lattice dimensions. The position of the rth atom in the unit cell can thus be represented by the vector

$$\underline{r}_r = x_r \underline{a} + y_r \underline{b} + z_r \underline{c} \quad .$$

(1.2.9)

The path difference between the electromagnetic waves scattered by these atoms and those that would be scattered by a set of atoms at the points of the lattice that define the origin of the unit cells is, by analogy with equation (1.2.2), $\lambda \underline{r}_r \cdot \underline{S}$. Thus the expression for the individual scattered wave from the rth atom contains a term

$$f_r \exp\left(\frac{2\pi i}{\lambda} \lambda \underline{r}_r \cdot \underline{S}\right) = f_r \exp\left(2\pi i \, \underline{r}_r \cdot \underline{S}\right)$$

where f_r is the scattering factor of the rth atom. Thus the expression for the complete wave scattered by the crystal would contain a term

$$FC = \sum_{r=1}^{R} f_r \exp{(2\pi i \; \underline{r}_r \cdot \underline{S})} \; . \qquad (1.2.10)$$

For S satisfying

Using equations (1.2.4) and (1.2.9), equation (1.2.10) may

be re-written

$$FC = \sum_{r=1}^{R} f_r \exp{[2\pi i(hx_r + ky_r + lz_r)]} \; . \qquad (1.2.11)$$

The quantity FC (a function of h,k and l) is called the structure factor, and equation (1.2.11) the structure factor equation; its importance in the field of crystallography cannot be over-emphasised.

It is reasonable to assume that a statistician may start his analysis conditional upon knowing the appropriate symmetries, space group and number of atoms, R, in the unit cell of the crystal under examination. Indeed, it is regarded as a necessary preliminary to a full structure determination. In the theory outlined above the atomic co-ordinate positions to be estimated inside the unit cell are defined by 3R structural parameters. When symmetry is present, the number of dependent parameters is less than 3R, sometimes considerably less; in extreme cases, some atoms can even be precisely located from symmetry considerations alone.

We shall be concentrating in this thesis on the centrosymmetric case, in which for every atom at (x,y,z)

25

there is an equivalent atom at (-x,-y,-z). In this case, equation (1.2.11) reduces to

$$FC(\underline{h}) = 2 \sum_{r=1}^{R/2} f_r \cos[2\pi(hx_r + ky_r + lz_r)] \quad (1.2.12)$$

where the summation is now over the asymmetric part of the unit cell and we have written FC = FC($\underline{h}$) to denote its dependence on the triple $\underline{h}$ = hkl. Other simplifications exist for different space groups. There is a complete list in the International Tables for X-ray Crystallography; selected examples are given in [51], while Chapter 4 of [54] expands upon the significance of the resulting simplifications. In particular, it should be noted that for a given space group, various hkl's will give a zero contribution to the geometrical part of the structure factor, thus indicating systematic absences on the X-ray diffraction photograph (to be introduced in the next Section); such absences characterise that particular space group.

Finally, we note that we may express the FC($\underline{h}$) in the form

$$FC(\underline{h}) = \left| FC(\underline{h}) \right| \exp(i\phi_{\underline{h}}) \quad (1.2.13)$$

where $|FC(h)|$ is called the structure amplitude and $\phi_{\underline{h}}$ the associated phase angle. Although we may measure the structure amplitude, the phases are unobservable quantities. In the centosymmetric case, the phase angle

is either 0 or π depending whether the structure factor is positive or negative; in practice, the observed structure amplitude is accorded a phase angle based on the corresponding fitted value.

1.3 Experimental Procedure in X-ray Crystallography

The diffraction effects produced when X-rays pass through a crystal are studied. X-rays are electromagnetic waves of very high frequency and short wavelength, produced when rapidly moving electrons (accelerated by potentials of tens of thousand of volts) collide with atoms; the energy released when the electrons are suddenly stopped is given out in the form of electromagnetic waves having a wavelength of around 1Å (10^{-8} cm), the order of magnitude of interatomic distances. The wavelength distribution in the X-ray beam depends on the material of the target (of the X-ray tube) used and on the accelerating voltage. For most crystallographic purposes, a monochromatic beam (i.e. a beam consisting of one wavelength λ) is desirable.

The intensities of the diffracted beams need to be considered, since these are observable quantities which depend upon the arrangement and position of atoms in the unit cell. There are two main methods for recording diffraction patterns of crystals and measuring the diffracted intensities, namely photographic methods and direct counting diffractometer methods: rotation

photographs will be briefly discussed here.

For a rotation photograph the crystal is set to rotate about some axis. The incident beam is normal to the rotation axis; the diffraction pattern is usually recorded on a cylindrical film co-axial with this axis (see Figure 1.3.1). The diffracted beams intersect the cylindrical film in a set of circles that appear as straight lines (layer lines) when the film is flattened out. It is possible to assign reflection indices (hkl) to the diffracted beams, which may be regarded as the reflection of X-rays by a particular set of parallel crystal planes as discussed in the previous Section. X-ray intensity measurements from the blackening of photographic film emulsion are made as follows.
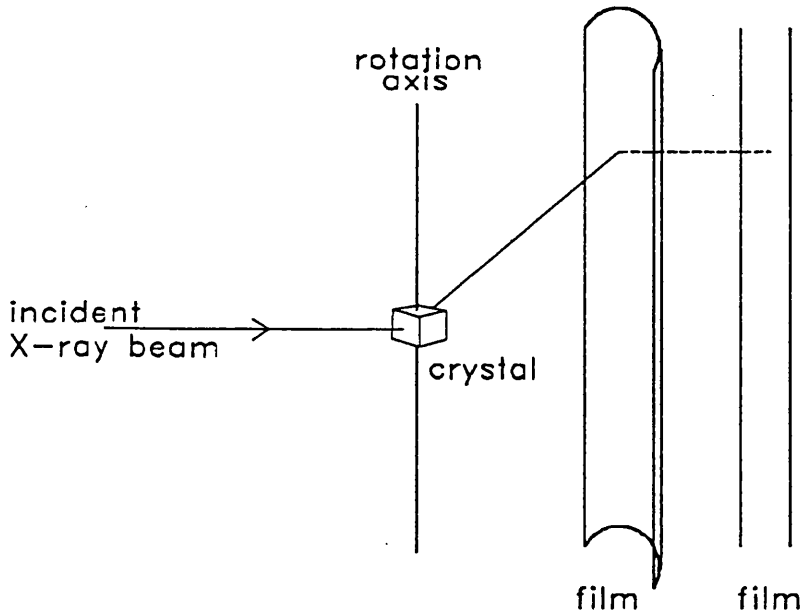
Figure 1.3.1: Simplified schematic diagram of X-ray rotation photograph procedure

Intensity scales may be prepared by allowing reflected beams to strike a film for different lengths of time and according each spot a value in proportion to this length. Intensities may then be measured by visual comparison with the scale. Alternatively, a photometric device may be used to estimate the blackening, the background intensity being measured and subtracted from the peak intensity. The general subject of accuracy in photographic measurements has been discussed exhaustively in [47], along with fuller details of the above procedure.

For a complete study of X-ray diffraction from crystals, the concept of a reciprocal lattice is needed. Such a lattice may be postulated for each direct lattice as follows. From the origin of the unit cell, draw lines normal to families of planes (hkl) in direct space; note the parentheses to denote the families of planes. Along each line, reciprocal lattice points hkl - no parentheses - are marked off such that the distance from the origin to the first point in any line is inversely proportional to the corresponding interplanar spacing d(hkl). Our spacing becomes d*(hkl) (say) given by

$$d*(hkl) = \frac{K}{d(hkl)} \qquad (1.3.1)$$

for some constant K. In practice, K is often taken to be the wavelength of the X-radiation used; reciprocal lattice units (RU) are then dimensionless.
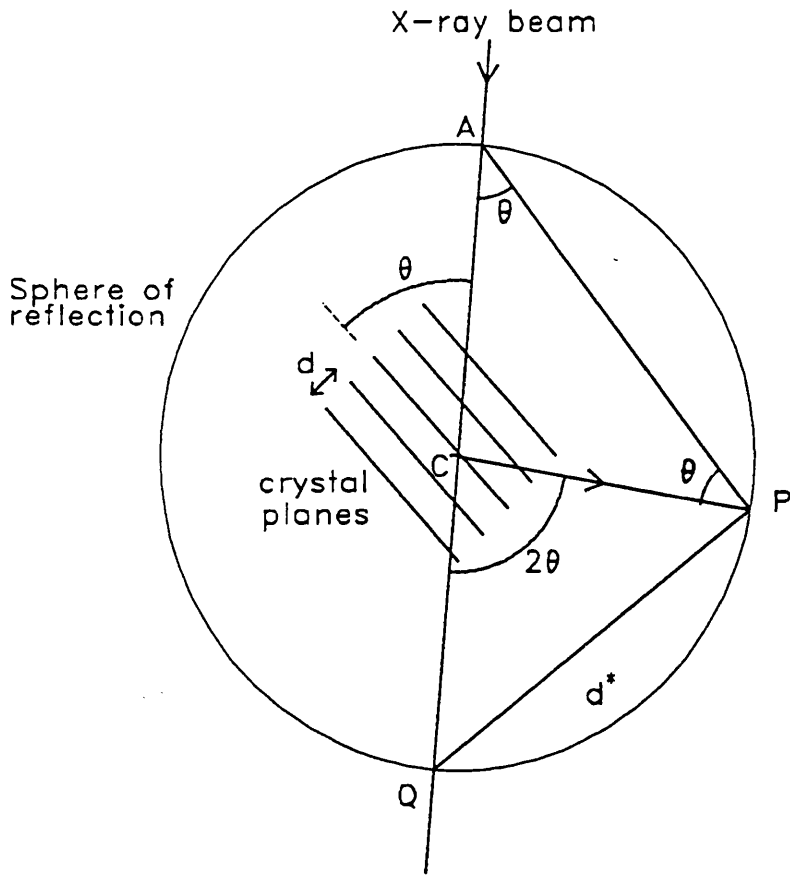
Figure 1.3.2:   Ewald's construction

We may now introduce the sphere of reflection, or Ewald's sphere, a device which facilitates the geometric interpretation of X-ray diffraction photographs. The sphere is centred on the crystal (C) and drawn with a radius of 1 RU on the X-ray beam (AQ) as diameter (see Figure 1.3.2). Let a reflected beam hkl cut the sphere in P. Then A,P and Q lie on a circular section of the sphere which passes through the centre C. It can be seen from Figure 1.3.2 that

$$QP = 2 \sin \theta (hkl) \quad .$$

From Bragg's equation (1.2.6) and from the definition of the reciprocal lattice above, we identify the point P with

the reciprocal lattice point hkl; hence

$$QP = d*(hkl)$$

(by taking K = λ in equation (1.3.1)).

The upshot of the above analysis is as follows. The condition that the crystal is in the correct orientation for a Bragg reflection hkl to take place is that the corresponding reciprocal lattice point P is on the sphere of reflection, as constructed above. As the crystal oscillates, an X-ray reflection flashes out each time a reciprocal lattice point cuts the sphere of reflection, and the direction of reflection is given by CP.

The apparently irregular variations of intensity observed are due to the effect of the relative position of the atoms in space. We recall the general structure factor equation (1.2.11):

$$FC(\underline{h}) = 2 \sum_{r=1}^{R} f_r \exp[2\pi i(hx_r + ky_r + lz_r)] \quad .$$

The modulus $|FC(\underline{h})|$ is called the structure amplitude; the intensities are proportional to $|FC(\underline{h})|^2$. In practice, observed intensities are measured from which observed structure amplitudes $|FO(\underline{h})|$ can be derived. See Appendix 1 for further comments on (integrated) intensity and related crystallographic concepts.

31

1.4 Statistical Problems Arising: Motivation For and Plan of this Thesis

The co-ordinates in a model structure need to be determined such that the set of calculated structure factors best fit (in some sense) an observed set of structure factors. The method usually used is that of (weighted) least squares. As we shall see in Chapter 2, however, although we can determine a set of structure amplitudes from our intensity measurements, direct elucidation of the crystal structure would further require a knowledge of complex numbers of the form

$$\left| FO(\underline{h}) \right| \exp(i \, \phi_h) \quad .$$

However, the "phases" ϕ_h cannot be determined from experiment: this is the so-called phase problem of crystallography. It manifests itself in the centrosymmetric case in the loss of sign information for our observed structure factors. Statistical techniques may be developed to help combat the phase problem (see Section 3.2) and provide a basis for sensible sign allocation. Similarly, the allocation of appropriate least squares weights may be considered from a statistical perspective and the allocation of appropriate weighting schemes undertaken (see Section 3.3).

Another major problem arises from the fact that our fundamental equation (1.2.12) is highly non-linear in the

32

parameters of interest. We will therefore need to be concerned with the development of procedures for handling non-linear models. Chapter 4 is consequently entirely devoted to the area of non-linear parameter estimation. Efficient techniques for dealing with "nuisance" parameters, such as the thermal parameters that enter equation (1.2.12), will also need to be developed.

Methods of handling prior information and how it may be incorporated into our analyses will be considered. In particular, statistical procedures will be employed to show how structural information available (in crystallographic data bases) may be used. The most important application will be that of sequential experimentation, to be discussed in Sections 3.4-3.6, in which the data collection process is modified in the light of results based on data currently available. The important statistical areas to be covered include the specification of relevant updating procedures, methods of selecting sampling strategies and the consideration of appropriate stopping rules.

Now that the relevant background theory has been introduced and some of the associated statistical problems outlined in their proper context, the prime motivation for this thesis may be stated as follows. We will be concerned with various statistical techniques arising in the field of X-ray crystallography, ultimately with regard

to producing fully automatic (computer-controlled) sequential experimentation procedures for efficient crystal structure determination. An overview of the necessary stages in such an analysis is presented in Figure 1.4.1. To conclude this Section a few introductory comments will be made on the various stages involved.

```
                    Objectives
                        |
                        v
                Prior Information
                        |
                        |
                        v
           Initial Sampling Scheme
                        |
                        |
                        v                       respond to
    |----------------New Estimate --------------------------->
    |                   |                       unwanted
    |                   |                       behaviour
    |                   |
    |  Re-cycle         v
    <------------ Stopping Rules  ------------> STOP
    |                   |
    |                   |
    |                   v
    <----------- More data
```
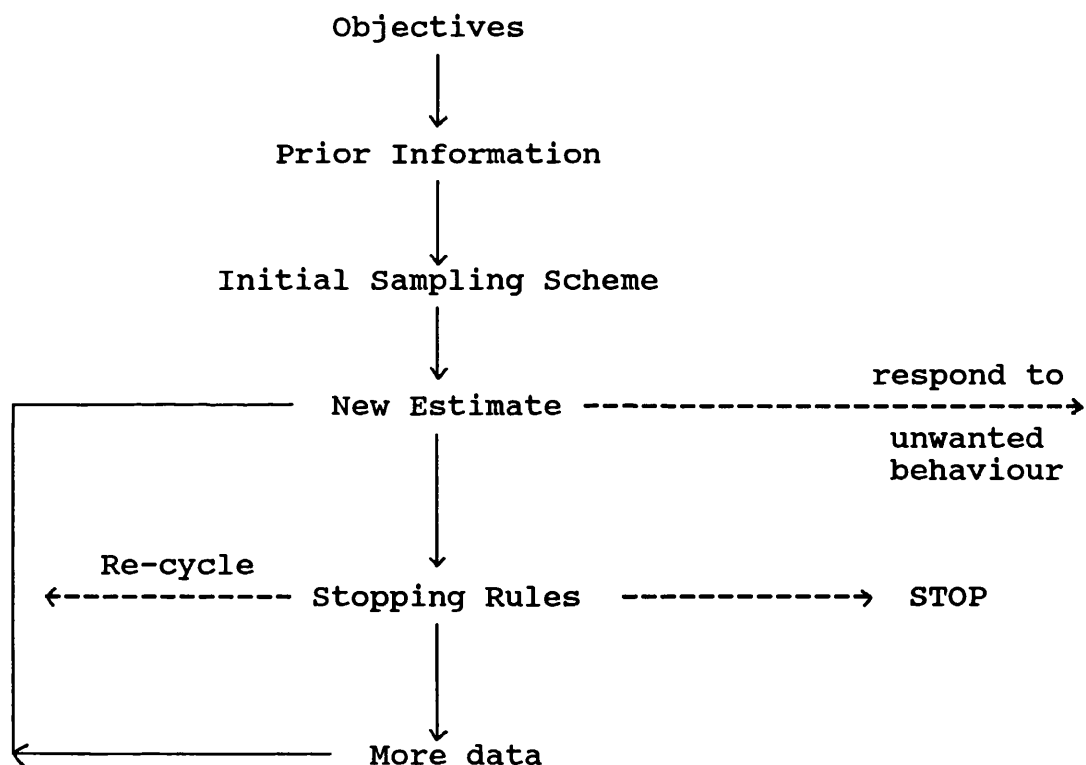
Figure 1.4.1: Overview of a sequential experimentation scheme as may be applied to X-ray crystallography

The objectives of a crystal structure analysis may influence our approach, in particular with regard to any simple matrix approximations we may wish to justify. Typical objectives include quick estimates (or packing

structures), accurate specification of selected parameters (as opposed to the whole molecule), or simply the most cost-effective estimate within a given time period. Our studies in Section 3.6 are motivated by wishing to make final refinements by measuring some but not all the intensities in the set of as yet unobserved reflections. We will seek to choose an appropriate subset that is not unduly large but still enables us to make very good final refinements. Much of our theory will be based on the minimax strategy, which seeks to reduce all associated parameter variances below specified target values. For this part of the analysis, we will be assuming we have good initial estimates and have already invoked an initial sampling scheme. Further comments will be found in Section 3.6.

An important constituent of our prior information before any experimentation is undertaken is the set of postulated atomic co-ordinate positions which represent our prior beliefs about the atomic configuration under consideration. Such a trial structure is a necessary preliminary to the standard least squares approach of crystal structure determination in view of the non-linear nature of the model equation that arises in crystallography. A suitable trial structure - together with the analogous prior beliefs about the temperature factors of Section 1.2 - may be assumed to be available to us. Were we further able to quantify our degree of belief

in our trial structure, as proposed in Section 3.2 we could proceed along the Bayesian lines outlined there and in Section 3.3. Other types of prior information that may be available and how they might be incorporated into our analysis are also given in Section 3.3.

A few brief comments on initial sampling schemes appear in Section 3.6; we merely note here that we shall generally assume appropriate initial batches of data to be readily available.

The key step of obtaining new parameter estimates raises a whole host of problems, whether we are invoking an iterative scheme based on a fixed set of data or updating existing beliefs in the light of further acquired data. Statistical considerations are dominated by the inherent non-linearity of our model and fundamental phase problem. We shall seek schemes to account for matrix inversion/ simplification wherever possible and updating procedures concerned with the recursive addition of data. Much of the associated theory is given in Section 3.5, while Chapter 4 is devoted to the general theory of non-linear parameter estimation. The standard refinement methods in general use in the crystallographic world are summarised in Chapter 2, the theory being given for the case of fixed data only. The important question of how to proceed when our new estimates are unsatisfactory (e.g. they may lead to an increase in a target function such as the residual

sum of squares) is touched upon during Chapters 3 and 4. For example, modification of a least squares step length or appeal to techniques such as steepest descent or Marquardt's compromise are cited as viable possibilities, together with our own ideas proposed in Chapter 4.

The scope for various stopping rules is wide. Standard rules exist such as cycling until resultant increments become negligible in the least squares iterative schemes for fixed data. When to stop the sequential addition of data into the analysis is not usually so clear-cut, however, though our minimax strategy presents no problems on this front. The general problem of whether further data should be acquired might best be considered in a decision theoretic framework in which we might pose the question: does the expected gain in information arising from making the extra measurement(s) justify the cost involved? Alternatively we might ask whether further acquisition of data is likely to afford us any useful additional information. Only if the answers to these sort of questions are in the negative, should we consider terminating our data collection procedure. Such an approach is not pursued here, though is recommended as an area for further study. Further general comments, including a note of standard stopping rules currently employed in crystallography, mainly of an ad hoc nature, are discussed as they arise during the formulation of the theory of Chapter 3.

In cases where we deem acquisition of further data should indeed be undertaken, we shall wish to specify how many observations to make and where to make them. The specification of appropriate such sequential sampling schemes forms a major part of the theory of Chapter 3. Emphasis is put on where to make the additional observations. One-at-a-time selection schemes are relatively easy to formulate; possible extensions to cover general batches of incoming data are given. For these ideas, including determination of the optimum size of such batches, consultation with crystallographers is recommended.

In fact, the whole science of crystallographic statistics should be based upon co-operation and interaction between the two disciplines. As we have seen, there is clearly plenty of scope for the application of various statistical techniques within the crystallographic setting; the most efficient applications are likely to be those that incorporate the relevant theory in such a way as to best utilise the expert knowledge of the crystallographers. In cases where this expert knowledge is intuitive or based on rules of thumb, sound statistical theory should be sought which parallels the appropriate results. If such a theory is not immediately apparent, it should not devalue the usefulness of these techniques that have served crystallography well over the years. The statistician

should (initially at least) be prepared to compromise his aesthetic values in search of an efficient practical procedure.

1.5 Notation

The following notation shall be used throughout this thesis. Suppose we have a known response function analogue of our structure factor equation (1.2.12)

$$\eta = FC(\underline{h}, \underline{\theta}) \qquad (\text{say}) \qquad\qquad (1.5.1)$$

where $\underline{h}$ is the hkl triple of Section 1.2 and

$$\underline{\theta} = (\theta_1 \ \cdots \ \theta_p)^T$$

is the vector of parameters to be estimated (p = number of unknown parameters). Typically, $\underline{\theta}$ consists of atomic co-ordinate positions - conventionally expressed as fractions of the lattice dimension - together with isotropic or anisotropic temperature factor parameters (up to 6 per atom). An overall scale factor may also be included in $\underline{\theta}$. Note that we shall henceforth also use the abbreviation FC($\underline{h}$) to denote the particular value of the (theoretical) structure factor evaluated at our current parameter estimate.

Suppose at any given stage we have n observations

$$\left| FO(\underline{h}_i) \right| = \left| FT(\underline{h}_i) \right| + \epsilon_i \qquad (i=1..n) \qquad (1.5.2)$$

39

where $|FO(\underline{h}_i)|$ represents the measured structure amplitude for the $\underline{h}_i$ reflection, and $FT(\underline{h}_i)$ the true (but unknown) value of the corresponding structure factor. The disturbances ϵ_i assumed to be normally distributed with mean zero and variance $\sigma_{\underline{h}_i}^2$. Thus equation (1.5.2) may be be re-expressed as

$$|FO(\underline{h}_i)| \sim N(|FT(\underline{h}_i)|, \sigma_{\underline{h}_i}^2) \qquad (i=1\ldots n) \qquad (1.5.3)$$

which represents our model. The multivariate form of our model may be obtained by grouping the observations into an (nx1) vector with the appropriate mean vector and covariance matrix $\underline{V}$ (say). The disturbances ϵ_i, which represent measurement errors, are often assumed to be independently distributed so that the latter reduces to a diagonal matrix with elements $\sigma_{\underline{h}_i}^2$.

For any iterative non-linear parameter estimation schemes that we shall be considering, let $\underline{\theta}^{(k)}$ denote the estimate of $\underline{\theta}$ after iteration k (= 0,1,2 ...). The initial estimate $\underline{\theta}^{(0)}$, however, shall often also be written $\underline{\theta}_0$. We let $\underline{\Delta\theta}^{(k)}$ likewise represent the increment for the estimate after the (k+1)th iteration, i.e.

$$\underline{\theta}^{(k+1)} = \underline{\theta}^{(k)} + \underline{\Delta\theta}^{(k)} .$$

We sometimes use the notation $\underline{\Delta\hat{\theta}}$ to emphasise that our

increment has been estimated via a least squares approach (as outlined in Section 2.2).

We shall use the notation $\Big|_{\underline{\theta}^{(k)}}$ to denote the value of a quantity evaluated at $\underline{\theta} = \underline{\theta}^{(k)}$; in particular,

$$X^{(k)}_{(i,j)} = \frac{\partial FC}{\partial \theta_j} (\underline{h}_i, \underline{\theta}) \Big|_{\underline{\theta}^{(k)}} \qquad (i=1..n; \; j=1..p) \qquad (1.5.4)$$

represents the value of the partial derivative of the structure factor equation for the $\underline{h}_i$ reflection with respect to θ_j, evaluated at $\underline{\theta}=\underline{\theta}_{(k)}$. The (nxp) design matrix $\underset{\sim}{X}^{(k)}$ at the kth iteration is defined by the matrix whose (i,j)th element is given by equation (1.5.4). In the centrosymmetric case our 'observations' at each iteration will be taken to be the (nx1) vector $\underline{Y}^{(k)}$ whose i^{th} element is given by

$$Y^{(k)}_i = FO*(\underline{h}_i) - FC(\underline{h}_i, \underline{\theta}) \Big|_{\underline{\theta}^{(k)}} \qquad (i=1..n) \qquad (1.5.5)$$

where we use the notation $FO*(\underline{h}_i)$ to represent the measured structure amplitude for the $\underline{h}_i$ reflection, given the same sign as its fitted value.

Additional notation will be introduced as and when necessary.

CHAPTER 2

CRYSTAL STRUCTURE REFINEMENT: A REVIEW OF METHODS

2.1 Introduction

In crystal structure analyses the co-ordinates of the
atoms inside a unit cell may be estimated by the technique
of least squares. However, an approach based upon the
distribution of electron density throughout the unit cell
may also be used. It can be shown that although the
former technique is parametric and the latter non-
parametric the underlying theory for the least squares
approach may be arrived at by considering a special case
of the electron density. This Chapter aims to study the
non-parametric method and to compare and contrast it with
the widely adopted parametric approach to crystal
structure determination.

The first method - that of least squares - is based upon
measured structure factors, derived from observed
intensities, which may be obtained by a study of the
diffraction effects produced when X-rays pass through a
suitably orientated crystal. The structure factors are
assumed to depend on the atomic co-ordinate positions, and
our atomic co-ordinate estimates are taken to be those
that minimise the discrepancy - in the sense of a weighted
residual sum of squares - between the consequent
calculated and observed values. The second method, which

also makes use of measured structure factors, seeks to create a three-dimensional image of the scattering matter inside the unit cell - the so-called electron density - and to identify atomic co-ordinate positions with the centres of the peaks of such a map.

Sections 2.2 and 2.3 will explain the least squares approach and electron density approach to structure determination in more detail. The method of differential synthesis, a refinement procedure based on the electron density approach, will be discussed in Section 2.4. Section 2.5 will briefly compare and contrast the two differing approaches to structure determination, based on the salient points of Sections 2.2 and 2.3. Section 2.6 will introduce a few of the more popular methods that have been developed to use in conjunction with out refinement procedures in order to help overcome the phase problem of Section 1.4. The Chapter concludes in Section 2.7 with a summary of the related notions of Patterson maps, heavy atom methods and isomorphous replacement.

2.2   The Standard Least Squares Approach

The method of weighted least squares much used in structure determination by crystallographers is to minimise the function $\psi = \psi(\underline{\theta})$ below with respect to the parameters θ_j (j=1..p):

$$\psi = \sum_{\underline{h}} w_{\underline{h}} \left(\left| FO(\underline{h}) \right| - FC(\underline{h}, \underline{\theta}) \right| \right)^2 \qquad (2.2.1)$$

where $w_{\underline{h}}$ is the 'weight' of the observation and the summation is taken over all triples $\underline{h}$ at which observations have been taken. We note that this criterion function is equivalent to

$$\psi = \sum_{\underline{h}} w_{\underline{h}} \left(FO*(\underline{h}) - FC(\underline{h}, \underline{\theta}) \right)^2 \qquad (2.2.2)$$

in the centrosymmetric cases with which we will be concerned.

The weights associated with the X-ray data may be based on a combination of theoretical assumptions and practical experience (see e.g. [26] p.217) though are usually just taken to be the inverse of the variance of the corresponding observation. In the original application of the method of least squares to structure analysis, however, only relative weights could be estimated based on the assumption that

$$w_{\underline{h}} \propto \left| FO(\underline{h}) \right|^{-2} ;$$

other feasible alternatives include weighting schemes such as [22]

$$w_{\underline{h}} \propto \left(a + \left| FO(\underline{h}) \right| + b \left| FO(\underline{h}) \right|^2 \right)^{-1}$$

44

where a and b are about 2FO(min) and 2/FO(max) respectively (where FO(min) and FO(max) are the minimum and maximum observed structure amplitudes). The appropriateness of any given weighting scheme may be tested by grouping structure amplitudes according to any factor (e.g. $|FO(h)|$ or position in reciprocal space) that is suspected to influence the standard deviation, and checking that the average value of

$$w_{\underline{h}}(FO*(\underline{h}) - FC(\underline{h}))^2$$

is approximately the same in each group. The question of an appropriate weighting scheme to use in conjunction with the theory of this Section is returned to in Section 3.3.

Returning to our criterion function ψ of (2.2.2) it should be noted that a more general formulation would introduce cross-product terms leading to

$$\psi'(\text{say}) = \sum_{\underline{h},\underline{h}'} w_{\underline{h}\ \underline{h}'}(FO*(\underline{h}) - FC(\underline{h}))(FO*(\underline{h}') - FC(\underline{h}'))$$

$$(2.2.3)$$

where the weights $w_{\underline{h}\ \underline{h}'}$ are now typically elements of the inverse of the covariance matrix $\underset{\sim}{V}$ of our observations. We may therefore think of the standard weighted least squares approach based on (2.2.2) as taking $\underset{\sim}{V}$ to be the fixed diagonal matrix with elements $w_{\underline{h}}^{-1} = \sigma_{\underline{h}}^2$.

The essence of the least squares approach is that the

$FC(\underline{h},\underline{\theta})$ are non-linear functions of $\underline{\theta}$ and so may be expanded in a Taylor series about $\underline{\theta}^{(0)}$, an initial parameter value. By ignoring second order terms, the resulting expression is linear in terms of increments of the parameters $(\underline{\Delta\theta}^{(0)})$. Using this Taylor series approximation, minimisation of the criterion function (2.2.2) - with respect to these increments - may now be carried out. This is achieved by solving the normal equations for the parameter increments obtained by differentiating and setting equal to zero. The resulting increments are then added to the initial set of parameter values $\underline{\theta}^{(0)}$, the $FC(\underline{h},\underline{\theta})$ expanded about the new point in parameter space, and the next set of normal equations solved for the new increments. This procedure (assuming it converges) is repeated until resulting increments become negligible.

The observational equations obtained at iteration k (derived in detail in Section 3.6) are

$$(\underset{\sim}{X}^{(k)T} \underset{\sim}{W} \underset{\sim}{X}^{(k)}) \ \underline{\Delta\theta}^{(k)} = \underset{\sim}{X}^{(k)T} \underset{\sim}{W} \underset{\sim}{Y}^{(k)} \qquad (2.2.4)$$

where $\underset{\sim}{X}^{(k)}$, $\underset{\sim}{Y}^{(k)}$ are defined by equations (1.5.4), (1.5.5) and W is the diagonal matrix with elements $w_{\underline{h}}$. Equation (2.2.4) must be solved for $\underline{\Delta\theta}^{(k)}$ and used in conjunction with the updating scheme

$$\underline{\theta}^{(k+1)} = \underline{\theta}^{(k)} + \underline{\Delta\theta}^{(k)} \ . \qquad (2.2.5)$$

Recall that the matrix $\underset{\sim}{X}^{(k)}$ and the vector $\underline{Y}^{(k)}$ must be re-

evaluated at each updated parameter estimate. Following

[55] we may choose to cycle until

$$\left| \theta_j^{(k+1)} - \theta_j^{(k)} \right| \leq \frac{10^{-4}}{\theta_j^{(k)} + 10^{-3}} \quad (j=1..p). \quad (2.2.6)$$

The standard least squares scheme just described may be
represented as in Figure (2.2.1).

$$\underline{\theta}^{(0)}$$

$$\downarrow$$

$$k=0$$

$$\downarrow$$

Evaluate $\underline{X}^{(k)}, \underline{Y}^{(k)}$

$$\downarrow$$

Solve (for $\underline{\Delta\theta}^{(k)}$)

$$(\underline{X}^{(k)T} \underline{W} \underline{X}^{(k)} \underline{\Delta\theta}^{(k)} = \underline{X}^{(k)T} \underline{W} \underline{Y}^{(k)}$$

$$\downarrow$$

$$\underline{\theta}^{(k+1)} = \underline{\theta}^{(k)} + \underline{\Delta\theta}^{(k)}$$

$$\downarrow$$

No $? \left| \theta_j^{(k+1)} - \theta_j^{(k)} \right| \leq \dfrac{10^{-4}}{\theta_j^{(k)} + 10^{-3}} \quad (j=1..p)?$

YES

$$\downarrow$$

STOP

(ACCEPT $\underline{\theta}^{(k+1)}$)
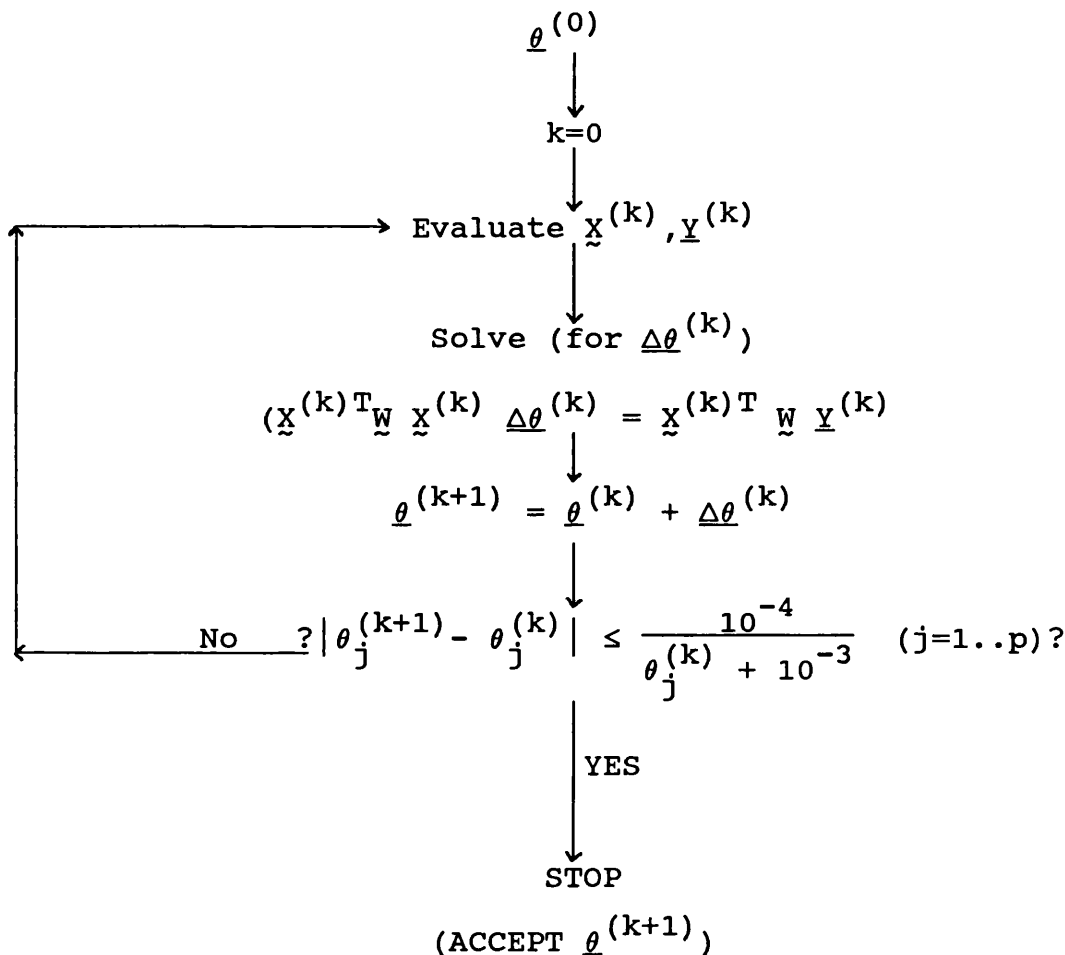
Figure 2.2.1: Standard least squares cycle

47

Provided

$$\underset{\sim}{N}^{(k)} = \underset{\sim}{X}^{(k)T} \underset{\sim}{W} \underset{\sim}{X}^{(k)} \tag{2.2.7}$$

is non-singular, equation (2.2.4) gives

$$\underline{\Delta\theta}^{(k)} = \left[\underset{\sim}{N}^{(k)} \right]^{-1} \underset{\sim}{X}^{(k)T} \underset{\sim}{W} \underset{\sim}{Y}^{(k)} \ . \tag{2.2.8}$$

This is the so-called multiple linear regression formula. It is trivial to deduce that (2.2.8) pertains when $\underset{\sim}{X}^{(k)}$ and $\underset{\sim}{Y}^{(k)}$ are defined via

$$X^{(k)}_{(i,j)} = \left. \frac{\partial \left| FC(\underline{h}_i, \underline{\theta}) \right|}{\partial \theta_j} \right|_{\underline{\theta}=\underline{\theta}^{(k)}} \tag{2.2.9}$$

and

$$Y_i^{(k)} = \left. \left| FO(\underline{h}_i) \right| - \left| FC(\underline{h}_i, \underline{\theta}) \right| \right|_{\underline{\theta}=\underline{\theta}^{(k)}} \tag{2.2.10}$$

respectively. From our (approximate) observational equations based on equation (1.5.3), namely

$$\underset{\sim}{Y}^{(k)} \sim N(\underset{\sim}{X}^{(k)}\underline{\Delta\theta}^{(k)}, \underset{\sim}{V}) \ , \tag{2.2.11}$$

we can now deduce that our estimate based on (2.2.8) is linear unbiased. The covariance matrix of the sampling distribution of our estimate is

$$[\underset{\sim}{N}^{(k)}]^{-1} \underset{\sim}{X}^{(k)T} \underset{\sim}{W} \underset{\sim}{V} \underset{\sim}{W} \underset{\sim}{X}^{(k)} [\underset{\sim}{N}^{(k)}]^{-1}$$

which reduces to

48

$$[\underset{\sim}{N}^{(k)}]^{-1}$$

in the case where the weight matrix $\underset{\sim}{W} = \underset{\sim}{V}^{-1}$. The Gauss-Markov Theorem says that amongst all linear unbiased estimates this is the one with smallest variance in the sense that it minimises the determinant of the resultant matrix. Furthermore, in the case of normal errors, the estimate is efficient.

The above analysis is based on a fixed weight matrix W which makes no allowance for any Taylor series residual of higher order (2nd, 3rd...) terms. Were we to know the correct appropriate covariance matrix we would of course use it. Appreciative of the fact that these residuals change dependent on the parameter estimate, dynamic weighting schemes could be proposed which modify the weights used in the analysis above accordingly as the parameter estimate varies. Thus in general it may be reasonable to expect to re-evaluate a weight matrix $\underset{\sim}{W}^{(k)}$ at each iteration. Further related ideas appear in Section 6.1.

The least squares theory discussed here required only an initial parameter value $\underset{\sim}{\theta}^{(0)}$ before the proposed iterative procedure could be invoked. General prior parameter distribution information available may be incorporated naturally into the analysis by appealing to the Bayesian inferential paradigm [52] as outlined in Section 3.3. It

should also be noted that the general problem of minimising equation (2.2.1) under conditions such as the atomic co-ordinate parameters having to satisfy certain equations of constraint may be incorporated into the least squares analysis along the lines proposed in [8] or may be solved by the standard method of undetermined (Lagrangian) multipliers as discussed in e.g. [26].

Once the least squares refinement procedure - which is especially suited for automatic, iterative computer operation - is invoked, however, it is essentially blind in the following sense. It will not be able to suggest the nature of any additional features whose inclusion may improve the level of agreement between the observed data and the underlying parametric model. As we shall see in the following Sections this is in marked contrast to the interactive nature of the electron density approach to crystal structure determination.

In extreme cases, there may prove to be problems of collinearity amongst the columns of the matrix $\underset{\sim}{X}^{(k)}$, which may manifest itself in terms of a high condition number for the matrix $\underset{\sim}{N}^{(k)}$, consequently yielding unreliable estimates. In cases where it is in fact singular resort may be made to the theory of pseudoinverses as discussed in e.g. [11]. Potentially more serious, however, are various other computational difficulties associated with the least squares approach: in particular, we shall be

concerned in Section 3.5 with:

(1)   (Avoidance of) matrix inversion routines

(2)   Relationship between $\underset{\sim}{X}^{(k+1)}$ and $\underset{\sim}{X}^{(k)}$

(3)   Updating procedures to account for the recursive addition of data (recursive least squares).

The least squares approach represented in Figure 2.2.1 is for a fixed set of data and as such will only form an inner cycle of any sequential schemes we shall be discussing.   Again, relevant comments appear in Section 3.5.   However, it gives the best structure for the data already available, in the sense of smallest associated residual sum of squares ψ.

2.3   Electron Density Approach

Let $\rho(\underline{r})$ denote some continuous electron density distribution inside the unit cell i.e. a small volume $\underline{dr}$ at a vector distance $\underline{r}$ from the origin contains $\rho(\underline{r})$ $\underline{dr}$ electrons (so that $\rho(\underline{r})$ is necessarily non-negative). Let the directions of incident and scattered waves be indicated by vectors $\underline{s}_0$ and $\underline{s}$ (of magnitude $1/\lambda$) as in Figure 1.2.1.   Since a decrease in path length by δ corresponds to an increase in phase angle of $2\pi\delta/\lambda$ we find that the phase of the wave scattered at $\underline{r}$ (in the direction of $\underline{s}$) is

$$2\pi \underline{r} . \underline{S}$$

with respect to the origin, where

$$\underline{S} = \frac{\underline{s} - \underline{s}_0}{\lambda}$$

is the scattering vector of Section 1.2.   The expression

for the wave scattered at $\underline{r}$ then contains a term

$$\rho(\underline{r}) \ \underline{dr} \ \exp(2\pi i \ \underline{r} . \underline{S}) ;$$

consequently, the superposition of all the wavelets

scattered by the distribution is obtained by integration,

viz

$$F(\underline{S}) \ (\text{say}) = \int \rho(\underline{r}) \ \exp[2\pi i \ \underline{r} . \underline{S}] \ \underline{dr} \ . \qquad (2.3.1)$$

Note that equation (2.3.1) is a more general equation for

the structure factor of which equation (1.2.11) is a

special example.

For the derivation of (1.2.11) we are concerned with those

$\underline{S}$ such that the path difference $\lambda \underline{r} . \underline{S}$ is an integral number

of wavelengths whenever $\underline{r}$ is at the origin of another unit

cell - so that the corresponding waves are in phase, thus

reinforcing each other - and we find that $\underline{S}$ yields an

integer triplet (which we associate with our previous $\underline{h}$ =

hkl).   The continuous electron density is now assumed to

be discrete, being concentrated at the 'points' occupied by the atoms in the unit cell (so the integration is replaced by a finite sum), with the appropriate densities now being atomic scattering factors.

Returning to equation (2.3.1) we see that the structure factor $F(\underline{S})$ is the Fourier transform of the electron density $\rho(\underline{r})$. In turn we can express the electron density as the inverse Fourier transform of the structure factor, namely

$$\rho(XYZ) = \frac{1}{v} \sum_{\underline{h}} F(\underline{h}) \exp[-i\alpha_{\underline{h}}] \qquad (2.3.2)$$

where $F(\underline{h}) = F(\underline{S})$ - $\underline{h}$ being the integer triplet hkl associated with $\underline{S}$ as discussed above -

$$\alpha_{\underline{h}} = 2\pi(hX + kY + lZ) \ ,$$

v is the volume of the unit cell and XYZ represents any point inside the unit cell ($0 \leq X,Y,Z \leq 1$). Replacing $F(\underline{h})$ by the theoretical structure factor $FC(\underline{h})$ given by equation (1.2.11), the theoretical electron density inside the unit cell may be calculated. Note that $\sum_{\underline{h}}$ is used to denote the triple summation $\sum_h \sum_k \sum_l$. In principle this summation should include an infinite number of terms. Further remarks on this are deferred to later in the Section.

Using equation (1.2.13) - with the more general structure

53

factor $F(\underline{h})$ in place of FC - equation (2.3.2) may be re-written as

$$\rho(XYZ) = \frac{1}{v} \sum_{\underline{h}} |F(\underline{h})| \cos(\alpha_{\underline{h}} - \phi_{\underline{h}}) . \qquad (2.3.3)$$

From equation (2.3.3) we see that were the phases $\phi_{\underline{h}}$ to be known for each $\underline{h}$ we could combine the structure amplitudes $|F(\underline{h})|$ in a triple Fourier series to reconstruct the electron density. Then, assuming atoms to be at the centres of peaks of the resulting three-dimensional map, we would know the entire structure.

In practice, however, there are three main problems associated with the above approach. Firstly, for the structure amplitudes $|F(\underline{h})|$ we would like to use measured values $|FO(\underline{h})|$. However, our corresponding expression for the 'observed electron density'

$$\rho_O(XYZ) = \frac{1}{v} \sum_{\underline{h}} |FO(\underline{h})| \cos(\alpha_{\underline{h}} - \phi_{\underline{h}}) \qquad (2.3.4)$$

(where the subscript 'o' denotes the dependence on the observed amplitudes $|FO(\underline{h})|$) is now necessarily a finite sum (with observations only being available for those $\underline{h}$ whose associated Bragg angle of reflection θ satisfies

$$\frac{\sin \theta}{\lambda} < S_{max}$$

where λ is the wavelength of the X-ray radiation used and S_{max} corresponds to the radius of the limiting sphere in

54

the reciprocal lattice). Furthermore, the sensible assignment of the phases $\phi_{\underline{h}}$ (which are not observable quantities) is non-trivial and forms the basis of the phase problem mentioned in Section 1.4. Finally our observed amplitudes will be subject to some experimental error, however small.

As we have observed above the summation for our observed electron density ρ_0 in equation (2.3.4) is necessarily limited by the finite number of experimentally available observations. Since our derived electron density equations (2.3.2) and (2.3.3) incorporated infinite sums a termination-of-series error is introduced; its main effect is to produce a set of approximately spherical ripples surrounding each atomic peak. The strength of the ripples is approximately proportional to the peak strength and problems of interpretation may be posed by overlapping of weak peaks with ripples from strong ones. For the purposes of differential synthesis to be discussed in the next Section the termination-of-series errors present in the observed electron density will be assumed to be present in the theoretical electron density based on FC($\underline{h}$) if both summations are taken over exactly the same terms.

Returning to the theory above, however, we still need some method to evaluate unobserved structure amplitudes and all phases. One possibility is for them to be assigned theoretical values based on any current available

parameter estimate $\underline{\theta}$ (which may have been obtained by least squares analysis based upon the limited set of observed data available). This may only be regarded as a first approximation, however, depending as it does on an underlying parametric approach. For the allocation of phases, however, it may be better to resort to the so-called 'direct methods' which have been developed to help combat the phase problem. Correct allocation of phases is vitally important as the general features of an electron density map are known to depend much more on the phases than on the structure amplitudes. Some of the more popular direct methods in current use are discussed briefly in Section 2.6.

Alternative methods for dealing with partial (finite) data on the $\underline{h}$'s include one based on maximum entropy, a method which seeks to fit the 'best' distribution of electrons throughout the unit cell consistent with the experimental data available. Solution of a centrosymmetric structure by such an approach is discussed in [36].

Despite the foregoing comments, it is common crystallographic practice to set all unobservable structure amplitudes to zero so that the infinite sum in equation (2.3.3) is reduced to the finite summation in equation (2.3.4) over the $\underline{h}$'s at which observations may experimentally be observed. The justification for this is based on the parametric approach inherent in deriving our

simplified theoretical expression for the structure factor equation (1.2.11). Fake isotropic temperature factors may be thought of as being introduced into the atomic scattering factors used in equation (1.2.11); these have the effect of artificially lowering the values of f_r for the $\underline{h}$'s where we are unable to take measurements so that equations (1.2.11) and (1.2.13) now yield negligible values for the (theoretical) structure amplitudes in this region. The effect of this analysis on the resultant electron density map is that of the series-termination-errors discussed above. The peaks become less sharply defined, though their positions - and the consequent atomic co-ordinate estimates - are unaltered. The loss of sharpness in the peaks will be reflected by the associated standard errors of these estimates. In most analyses, crystallographers claim that the introduction of fake temperature factors is unnecessary, with those postulated in a given trial structure being sufficient to ensure negligibly small theoretical structure amplitudes for the appropriate $\underline{h}$'s.

Suppose we have augmented our set of available observed structure amplitudes by 'pseudo-observations' (including all phases) so that we have an approximation to the electron density based partly on the data available and partly on theoretical assumptions. We may find that measurement of certain further structure amplitudes will lead to improved information on phases to be deduced from

certain direct methods. Consequently, the electron density approach may be considered as being conducive to the technique of sequential experimentation, with additional observations being taken at the $\underline{h}$'s that will lead to most additional phase information.

One of the major advantages of the electron density approach to crystal structure determination outlined in this Section is that it is an interactive process i.e. the derived electron density maps are open to interpretation from the crystallographer and may consequently suggest paths of action. This also pertains to the difference maps of the next Section and further comments are made there. This is in stark contrast to the refinement process based on least squares, discussed in Section 2.2, which is an automatic process requiring no intervention from the crystallographer.

Perhaps the biggest single drawback to the approaches to crystal structure determination based on the ideas of electron density presented here is the resolution of the peaks of the resultant three-dimensional map. In particular, it was found that the computational effort required, especially in cases where the number of atoms in the unit cell was relatively small, was prohibitive and alternative methods of structure determination were sought. This heralded the advent of automatic iterative schemes such as the least squares refinement procedure.

2.4 The Method of Differential Synthesis

We recall that equation (2.3.4) for the 'observed electron density' is a <u>finite</u> series. We shall now consider one method of correcting for the ensuing series-termination-errors which involves a synthesis with coefficients

$$|FO(\underline{h})| - |FC(\underline{h})|,$$

where the theoretical structure factors $FC(\underline{h})$ and phases $\phi_{\underline{h}}$ are derived from a suitable postulated structure.

We define the difference function of the electron density by

$$\rho_D(XYZ) = \rho_o(XYZ) - \rho_c(XYZ) \qquad (2.4.1)$$

where $\rho_o(XYZ)$ is given by equation (2.3.4) and $\rho_c(XYZ)$ by

$$\rho_C(XYZ) = \frac{1}{v} \sum_{\underline{h}} |FC(\underline{h})| \cos(\alpha_{\underline{h}} - \phi_{\underline{h}}) \qquad (2.4.2)$$

(where the subscript 'c' denotes the dependence on the calculated structure factors $FC(\underline{h})$ and the summation is over precisely those reflections that appear in ρ_o). As before, a parametric model is needed with the $FC(\underline{h})$ being evaluated at the current parameter estimate in accordance with equation (1.2.11). The phases $\phi_{\underline{h}}$ are taken to be the same in both ρ_o and ρ_c.

The method of differential synthesis proceeds by

59

successively refining ρ_D based on the convergence criterion that all slopes and curvatures of ρ_D at the atomic positions are to be made equal to zero. It has been shown in [15] that this method is equivalent to a least squares refinement method based on an appropriate weighting scheme, and that comparison of the two methods reduces to comparison of the associated least squares weights $w_{\underline{h}}$.

The main justification for the above method is that it is open to informative interpretation in the sense that the prominent features of the difference synthesis enable various kinds of errors in the model structure to be recognised and corrected, e.g. missing or superfluous atoms or incorrectly placed atoms. Also, such maps will have peaks where insufficient electron density is included in the trial structure and troughs where too much is. Simultaneous refinement of the electron density is therefore also provided. Individual types of error may, however, be compounded to yield complicated patterns in the residual density. The procedure followed in practice is to correct only for the prominent features. These corrections are included in a revised structure model and the difference synthesis recalculated. We proceed in this way until the residual density contains no significant systematic features. Our resultant structural model will agree as well as possible with the observed electron density.

Finally, an expression for the variance of atomic co-ordinate estimates obtained via such an approach has been derived in [21]. For example, for co-ordinate x_r we find (after n observations have been taken)

$$\sigma^2_{x_r} = (\frac{2\pi}{v})^2 \sum_{\underline{h}} h^2 \frac{(|FO(\underline{h})| - |FC(\underline{h})|)^2}{|(C_{xx})_r|} \frac{n}{n-p} \qquad (2.4.3)$$

where $(C_{xx})_r$ (together with $(C_{xy})_r..(C_{zz})_r$) represents the curvature of the calculated electron density. For comparison, the corresponding variance derived from the least squares approach of the previous Section is given for cases in which only the relative observational errors are known, namely

$$\sigma^2_{x_r} = \sum_{\underline{h}} W_{\underline{h}} \frac{(|FO(\underline{h})| - |FC(\underline{h})|)^2}{n-p} N^{-1}_{jj} \qquad (2.4.4)$$

where $x_r = \theta_j$ and N^{-1}_{jj} is the jth diagonal element of the covariance matrix of the sampling distribution of the least squares estimate.

2.5 A Comparison and Contrast of the Two Standard Methods of Structure Determination

The main difference between the techniques of structure determination via the least squares and electron density approaches is that the former is a parametric approach

61

whilst the latter is non-parametric. The electron density approach seeks to determine the continuous distribution of the electrons throughout the unit cell, associating molecular formation with the areas of high concentration. Conversely, we would expect that molecule formation is accompanied by a certain re-distribution of the electron density of the individual atoms to bring more density into some regions, e.g. between the nuclei, with a corresponding density deficit in other regions. This information is not taken into account by the least squares model in which it is assumed that the electron density distribution can be regarded as a superposition of electron density peaks corresponding to free, spherically symmetric atoms. Alternatively, the least squares model may be viewed as that corresponding to the electron density which has delta functions at the points occupied by atomic centres and is zero elsewhere, which is clearly an over-simplification.

Despite the fundamental differences between the two approaches, it is interesting to note that the use of appropriate weighting schemes in the least squares setting leads to equivalence with differential synthesis, the refinement procedure based on electron density as discussed in Section 2.4. Expressions for the variances of atomic co-ordinate estimates under the two approaches are given by equations such as (2.4.3) and (2.4.4).

The least squares theory of Section 2.2 concentrates on the centrosymmetric case. Our criterion function (2.2.1) is based on observed structure amplitudes and the analogous fitted values. The phase problem reduces to that of correct sign allocation. In the formulation of Section 2.3, which is not confined to the centrosymmetric case, we see that the electron density approach is based on all structure factors - amplitudes and general phase angles - obtained from both observed and theoretical values. The general phase problem becomes that of the sensible assignment of such phase angles. This may be tackled via the direct methods of Section 2.6. Further methods for tackling the phase problem in the least squares setting will be introduced in Chapter 3.

Prior knowledge may readily be incorporated into the least squares approach. Once invoked, however, the refinement procedure is essentially blind as discussed in Section 2.2. On the other hand the electron density approach is very much an interactive procedure, relying on the crystallographer to use prior knowledge in the interpretation of the resultant electron density maps.

Other than the question of correct sign allocation, the problems involved in the least squares approach are mainly computational, including possible collinearity and various updating procedures (as discussed in Section 3.5). The main problem in the electron density approach - apart from

the determination of all structure factors and consequent phase problem - is the resolution of peaks.

Although electron density considerations will continue to play a role in structure determination, this thesis concentrates on the widely adopted parametric approach, which is found to be particularly conducive to automated sequential experimentation techniques of the type with which we will be concerned in Chapter 3. We will continue there to concentrate on the centrosymmetric case.

2.6   A Review of Some Direct Methods

Direct methods seek to tackle the phase problem by deriving phases of a structure by consideration of relationships among the indices and among the structure factor amplitudes, particularly of stronger reflections. Perhaps the single most important direct method is that of structure invariants and it is this that is described first.

The unitary structure factor $U(\underline{h})$ is defined by [54]

$$U(\underline{h}) = \sum_r n_r \exp[2\pi i (hx_r + ky_r + lz_r)] \quad (2.6.1)$$

where n_r, the unitary scattering factor, is given by

$$n_r = f_r / \sum_r f_r \quad .$$

The normalised structure factor $E(\underline{h})$ may now be defined by

$$E(\underline{h}) = |E(\underline{h})|\exp(i\phi_{\underline{h}}) = U(\underline{h})/(\sum_r n_r^2)^{\frac{1}{2}} \quad . \qquad (2.6.2)$$

By taking the scattering factor, f_r, to be that corresponding to a Bragg angle $\theta = 0$ (viz $f_r = Z_r$, the atomic number of the rth atom), the ensuing $E(\underline{h})$ values may be thought of as eliminating the effects of thermal vibration and treating each atom as if all its electrons were concentrated at a point. We may then wish to determine those (x_r, y_r, z_r) which minimise the modified weighted residual sum of squares

$$\psi_E(\text{say}) = \sum_{\underline{h}} W_{\underline{h}} (|E(\underline{h})| - \left| \frac{1}{(\sum_r z_r^2)^{\frac{1}{2}}} \sum_r z_r \exp[2\pi i \right.$$

$$\left. \cdot (hx_r + ky_r + lz_r)] \right|)^2 \quad . \qquad (2.6.3)$$

The co-ordinates (which are unknown) may be eliminated from (2.6.3) to yield relationships among the $E(\underline{h})$'s having probabilistic validity, which in turn lead to approximate values of the unknown phases $\phi_{\underline{h}}$. These individual values are found to depend on the structure and also the choice of origin; however, certain linear combinations of the phases - the structure invariants - are determined by the structure alone and are independent of the origin. Important structure invariants are the

linear combinations of three phases (triplets) and four phases (quartets), as will presently be introduced.

One important underlying part of the theory is that the value of a given structure invariant is primarily determined, in favourable cases, by the values of one or more small sets of magnitudes $E(\underline{h})$ - the so-called 'neighbourhoods' - and is relatively insensitive to the values of the remaining $E(\underline{h})$ magnitudes (the neighbourhood principle). Before proceeding further, it should be stressed that recovery of the individual phases from the appropriate structure invariants available should pose no great problem. We first consider the case of triplets.

Let H denote the collection of all reciprocal lattice vectors $\underline{h}$ = hkl, and let R_1, R_2 and R_3, be fixed non-negative numbers. The ordered triple of reciprocal vectors $\underline{h}_1$, $\underline{h}_2$, $\underline{h}_3$ is assumed to be uniformly distributed over the subset of HxHxH defined by

$$\left| E(\underline{h}_1) \right| = R_1, \quad \left| E(\underline{h}_2) \right| = R_2, \quad \left| E(\underline{h}_3) \right| = R_3 \qquad (2.6.4)$$

and

$$\underline{h}_1 + \underline{h}_2 + \underline{h}_3 = \underline{0} \, . \qquad (2.6.5)$$

. The structure invariant (triplet)

$$\phi_3 = \phi_{\underline{h}_1} + \phi_{\underline{h}_2} + \phi_{\underline{h}_3} \qquad (2.6.6)$$

(note that we require (2.6.5)) is a function of the random

66

variables and therefore is itself a random variable. The three magnitudes (2.6.4) define the first neighbourhood of ϕ_3, and given these, the conditional probability distribution of ϕ_3 is given by [41]

$$P_{1/3}(\text{say}) \simeq \frac{1}{2\pi I_0(A)} \exp (A \cos \Phi) \qquad (2.6.7)$$

where I_0 is the modified Bessel function and

$$A = \frac{2 \sigma_3}{\sigma_2^{3/2}} R_1 R_2 R_3$$

with

$$\sigma_n = \sum_{r=1}^{R} z_r^n .$$

This distribution will always have a unique maximum at $\Phi = 0$ in the interval $(-\pi, \pi)$ so that the most probable value of ϕ_3 is zero. The larger the value of A the smaller is the variance of the distribution and the more reliable is the estimate (zero) of ϕ_3.

Unfortunately, as we have just seen, the above theory gives no reliable non-zero estimate for any structure invariant (triplet), a situation which is paralleled in the analagous theory for quartets when only the appropriate first neighbourhood is considered. However, were we to consider the second neighbourhood (7 magnitudes in all) the theory for quartets may well yield non-zero estimates as discussed next.

Let R_1, R_2, R_3, R_4, R_{12}, R_{23} and R_{31} be fixed non-negative numbers. The ordered quadruple of reciprocal vectors $\underline{h}_1$, $\underline{h}_2$, $\underline{h}_3$ and $\underline{h}_4$ is assumed to be uniformly distributed over the subset of HxHxHxH defined by

$$\left| E(\underline{h}_1) \right| = R_1, \quad \left| E(\underline{h}_2) \right| = R_2, \quad \left| E(\underline{h}_3) \right| = R_3, \quad \left| E(\underline{h}_4) \right| = R_4$$

$$(2.6.8)$$

$$\left| E(\underline{h}_1 + \underline{h}_2) \right| = R_{12}, \quad \left| E(\underline{h}_2 + \underline{h}_3) \right| = R_{23}, \quad \left| E(\underline{h}_3 + \underline{h}_1) \right| = R_{31}$$

$$(2.6.9)$$

and

$$\underline{h}_1 + \underline{h}_2 + \underline{h}_3 + \underline{h}_4 = \underline{0} \ . \qquad (2.6.10)$$

The structure invariant (quartet)

$$\phi_4 = \phi_{\underline{h}_1} + \phi_{\underline{h}_2} + \phi_{\underline{h}_3} + \phi_{\underline{h}_4} \qquad (2.6.11)$$

(note that we require (2.6.10)) is a function of the random variables $\underline{h}_1$, $\underline{h}_2$, $\underline{h}_3$, $\underline{h}_4$ and therefore it itself a random variable. The seven magnitudes (2.6.8) and (2.6.9) define the second neighbourhood of ϕ_4, and given these, the conditional probability distribution of ϕ_4, can be calculated [40]. This distribution, as previously indicated, can have a maximum at any angle between 0 and π, these extremes being reached according as the three magnitudes in (2.6.9) are all large or all small, respectively.

A few other direct methods will now briefly be outlined. There are also available numerous (in)equality and various probability relations between structure factors which will not be mentioned here. We start by considering the centrosymmetric case.

When the crystal structure under consideration contains a centre of symmetry, the unitary structure factor of equation (2.6.1) may be expressed as

$$U(\underline{h}) = \sum_r n_r \cos[2\pi(hx_r + ky_r + lz_r)] \quad . \qquad (2.6.12)$$

By using Cauchy's inequality

$$\left|\sum_r a_r b_r\right|^2 \leq \left(\sum_r |a_r|^2\right)\left(\sum_r |b_r|^2\right) \qquad (2.6.13)$$

with

$$a_r = \sqrt{n_r}$$

and

$$b_r = \sqrt{n_r} \cos[2\pi(hx_r + ky_r + lz_r)]$$

we may derive

$$(U(\underline{h}))^2 \leq \tfrac{1}{2}(1 + U(2\underline{h})) \quad . \qquad (2.6.14)$$

These Harker-Kasper inequalities (2.6.14) may be used to show for example that a structure factor whose indices are all even is necessarily positive. Likewise, by appropriate use of (2.6.13) we may deduce

69

$$(U(\underline{h}_1) + U(\underline{h}_2))^2 \le (1 + U(\underline{h}_1 + \underline{h}_2))(1 + U(\underline{h}_1 - \underline{h}_2))$$

$$(2.6.15)$$

and

$$(U(\underline{h}_1) + U(\underline{h}_2))^2 \le (1 - U(\underline{h}_1 + \underline{h}_2))(1 - U(\underline{h}_1 - \underline{h}_2))$$

$$(2.6.16)$$

which may be combined to give

$$(|U(\underline{h}_1)| + |U(\underline{h}_2)|)^2 \le (1 + s(\underline{h}_1)s(\underline{h}_2)s(\underline{h}_1+\underline{h}_2)|U(\underline{h}_1+\underline{h}_2)|)$$
$$\times (1 + s(\underline{h}_1)s(\underline{h}_2)s(\underline{h}_1-\underline{h}_2)|U(\underline{h}_1+\underline{h}_2)|)$$

$$(2.6.17)$$

where $s(\underline{h})$ is the sign of $U(\underline{h})$ etc.

As should be readily appreciated, the presence of higher symmetry will often lead to many inequality relationships which may be stronger than those above and prove to be very useful. The importance of these inequalities lies in the fact that they enable various relations between structure factors to be established. One such relation, which will often be suggested by (2.6.17), is the triple-product sign relationship below.

For three strong related reflections in a centrosymmetric structure, the signs are related by

$$s(\underline{h}_1)\ s(\underline{h}_2)\ s(\underline{h}_1 + \underline{h}_2) \approx +1 \qquad (2.6.18)$$

70

where $\approx$ means 'probably equals'. A relatively simple expression (good enough for most practical purposes) used to estimate the degree of probability associated with the sign above is [16]

$$P(\underline{h}_1, \underline{h}_2) = \tfrac{1}{2} + \tfrac{1}{2}\tan h \, [\epsilon_3/\epsilon^{3/2}) \, |U(\underline{h}_1)U(\underline{h}_2)U(\underline{h}_1 + \underline{h}_2)|]$$

(2.6.19)

where

$$\epsilon_3 = \sum_r n_r^3$$

and

$$\epsilon = \sum_r n_r^2 \ .$$

For a structure containing R equal atoms in the unit cell (2.6.19) reduces to

$$P(\underline{h}_1, \underline{h}_2) = \tfrac{1}{2} + \tfrac{1}{2}\tan h \, [R|U(\underline{h}_1)U(\underline{h}_2)U(\underline{h}_1 + \underline{h}_2)|] \ . \quad (2.6.20)$$

From relationships such as (2.6.18) it is often possible to derive phases for almost all strong reflections and so to determine the structure from the resulting electron density map.

The final two methods here apply to symmetric and non-centrosymmetric structures alike. Approximate values for phase angles may be derived from

$$\phi_{\underline{h}} \approx \langle \phi_{\underline{h}-\underline{h}'} + \phi_{\underline{h}'} \rangle_{\underline{h}'}$$

(2.6.21)

where < > denotes the mean value. In addition to (2.6.21) the so-called tangent formula

$$\tan \phi_{\underline{h}} \approx \frac{<|E(\underline{h})E(\underline{h}-\underline{h}')|\sin(\phi_{\underline{h}'} + \phi_{\underline{h}-\underline{h}'})>_{\underline{h}'}}{<|E(\underline{h})E(\underline{h}-\underline{h}')|\cos(\phi_{\underline{h}'} + \phi_{\underline{h}-\underline{h}'})>_{\underline{h}'}}$$

(2.6.22)

is used extensively to calculate and refine phases.

2.7 Patterson Maps, Heavy Atoms and Isomorphous Replacements

The Patterson function is defined by ($0 \leq U,V,W \leq 1$)

$$P(UVW) = \frac{1}{v} \sum_{\underline{h}} |F(\underline{h})|^2 \cos[2\pi(hU + kV + lW)] \quad (2.7.1)$$

(c.f. expression for electron density (2.3.3). Note that no phase information is required for this map - only the relative positions of atoms.

The peaks in this map occur at points whose distances from the origin correspond in magnitude and direction with distances between atoms in the crystal. Ideally this map can be interpreted in terms of an atomic arrangement (consequently bypassing altogether the necessity of phase determination); however, it is more likely to serve as a guide and use of it may be made via the two approaches about to be described. For further details see e.g. [38].

If one or a few atoms of high atomic number are present,

they will dominate the scattering, as shown by Figure
2.7.1. If these atom(s) can be located from a Patterson
map, the phases of the entire structure may be
approximated by the phases of the heavy atom(s). In the
resulting electron density map, portions of the remainder
of the structure will usually be revealed. A 'heavy' atom
(M) with a much higher atomic number will have a much
longer vector, F(M) say, in a diagram such as Figure
2.7.1. Since the steps or f_r values for the lighter atoms
are relatively small, there is a high probability that $\phi_\underline{h}$
will be close to ϕ_M.
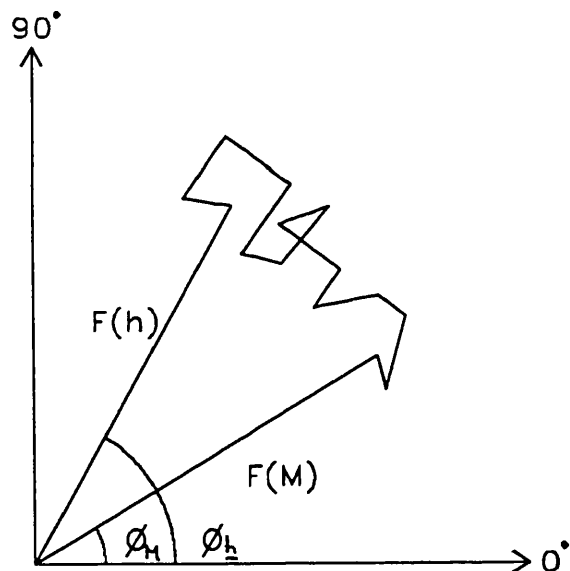
Figure 2.7.1   Vector representation of structure factors

Finally, the method of isomorphous replacement is one of
the best methods for the direct determination of phase
angles, being a practical approach particularly useful for
solving large structures such as those of proteins.   The

basic idea is as follows. Two crystals are said to be isomorphous if their space groups are the same and their unit cells and atomic arrangements are essentially identical. Suppose that atoms may be added to or replaced in a molecule to produce a new crystal isomorphous with those formed by the parent molecule. If the positions of these added or replaced atoms can be found from Patterson maps, their contributions to the phase angle of each reflection can be calculated, and if the atoms are sufficiently heavy, differences in intensities for the two isomorphs can be used to determine the phase angle for each reflection, either algebraically or graphically.

CHAPTER 3

STATISTICAL ASPECTS OF STRUCTURE DETERMINATION

3.1 Introduction

This Chapter expands upon some of the statistical problems outlined in Section 1.4, the discussion being confined throughout to the centrosymmetric case. Section 3.2 is concerned with the phase problem, which has occupied crystallographers for years. Historically, it has proved the biggest stumbling block to successful crystal structure determination [48]. Here, we outline the problem, and try and look at it from a statistical perspective. Some of the Bayesian ideas introduced in Section 3.2 carry over naturally to Section 3.3 in which we concentrate on various aspects of crystal structure analysis. In particular, we shall be considering the incorporation of prior information into our analysis and the formulation of appropriate weighting schemes. Suggestions will be made as to how the new results obtained may be used to alter existing procedures.

Section 3.4 sees us turn our attention to the important area of sequential experimentation. The general theory is outlined with particular emphasis being given to our specific crystallographic application. Some of the relevant updating procedures required for efficient sequential experimentation in our current setting are

75

given in Section 3.5.

Section 3.6 then considers various selection schemes we might use in conjunction with the ideas of the two previous Sections. General theory will be developed based on a minimax strategy and various approximations considered. We then go on to consider the similar results motivated by the criterion of D-optimality, before briefly discussing various related stratified sampling schemes. We make a few comments on the possible use of non-linearity weights before concluding the Section with brief notes on the practical implementation of our procedures.

Finally, we note that although the important problem of non-linearity is necessarily discussed throughout this Chapter, Chapter 4 is further dedicated to this key area.

3.2 A Bayesian Approach to Handling the Phase Problem of Centrosymmetric Structure Determination

The atomic arrangement in the unit cell of a crystal may be determined based upon the measured intensities of a sufficient number of X-ray diffraction maxima. From these intensities a set of numbers $|FO(\underline{h})|$ can be derived, one corresponding to each intensity. However, the elucidation of the crystal structure requires, as we have seen in Chapter 2, a knowledge of the complex numbers

$$FO(\underline{h}) = |FO(\underline{h})| \exp(i \phi_{\underline{h}}) ,$$

of which only the magnitudes $|FO(\underline{h})|$ can be determined from experiment. Thus a 'phase' $\phi_{\underline{h}}$ must be assigned to each $|FO(\underline{h})|$, and the problem of determining the phases when only the magnitudes $|FO(\underline{h})|$ are known is called the 'phase problem'. It is without doubt the biggest single problem in crystal structure determination. Standard methods to help combat the phase problem were introduced in Chapter 2. We continue here to concentrate on the centrosymmetric case, in which the phase problem reduces to that of correct sign allocation to the (necessarily positive) measured structure amplitudes $|FO(\underline{h})|$.

If the correct sign is not pre-determined (e.g. via the direct methods of Section 2.6) the measured structure amplitude is incorporated into the usual least squares analysis with the sign of the corresponding theoretical structure factor $FC(\underline{h})$ based on a specified trial atomic configuration. Thus an observed amplitude of 20 will enter the analysis with a positive sign whether our corresponding theoretical value is 1 or 21. We see that two different scenarios are being treated in the same way; the latter $|FC|$ value is sufficiently close to $|FO|$ to enable us to assign a positive sign with some degree of confidence, whereas the former value is not. Were our corresponding theoretical value to be 41, it might further be strongly argued that FO* should be taken to be positive on the grounds that 41 is much nearer to +20 than -20. We

have already commented in Section 2.3 that general features of electron density maps depend far more on correct allocation of phases than the actual amplitudes involved. In view of this we may therefore wish to incorporate into our analysis those observations, the signs of which we are most confident. For such an approach, however, we will not only need a prior estimate of the structure factor (FC), but also some measure of its accuracy. Such a measure shall now be derived.

Consider first the following general theory. Suppose that a random variable X has a normal distribution with mean μ and variance σ^2 i.e.

$$X \sim N(\mu, \sigma^2) \tag{3.2.1}$$

where μ, σ^2 are both known. Then (for small σ^2) we may approximate the mean and variance of cos X by

$$E(\cos X) = \cos \mu (1 - \frac{\sigma^2}{2}) \tag{3.2.2}$$

and

$$Var(\cos X) = \sigma^2 \sin^2 \mu \tag{3.2.3}$$

respectively. Similarly, if $X_1 \ldots X_R$ are independent normal random variables with distributions

$$X_r \sim N(\mu_r, \sigma_r^2) \qquad (r = 1 \ldots R) \tag{3.2.4}$$

then $\sum_r \cos X_r$ has approximate mean and variance

$$E(\Sigma \cos X_r) = \Sigma \cos \mu_r (1 - \frac{\sigma_r^2}{2}) \qquad (3.2.5)$$

and

$$Var(\Sigma \cos X_r) = \Sigma \sigma_r^2 \sin^2 \mu_r \qquad (3.2.6)$$

respectively. The above analysis may be extended to find the first two moments of

$$F = \Sigma_r C_r e^{-B_r} \cos X_r \qquad (3.2.7)$$

which incorporates additional independent (of each other and the X's) random variables $B_1 \ldots B_R$ with distributions

$$B_r \sim N(\alpha_r, \tau_r^2) \qquad (r=1 \ldots R) \qquad (3.2.8)$$

and constants $C_1 \ldots C_R$. The resultant approximate mean and variance are

$$E(F) = \Sigma_r C_r e^{-\alpha_r} \cos \mu_r (1 + \frac{\tau_r^2}{2} - \frac{\sigma_r^2}{2}) \qquad (3.2.9)$$

and

$$Var(F) = \Sigma_r C_r^2 e^{-2\alpha_r} (\sigma_r^2 \sin^2 \mu_r + \tau_r^2 \cos^2 \mu_r) \; .$$

$$(3.2.10)$$

The ramifications of equations (3.2.9) and (3.2.10) in the crystallographic setting are as follows. Note that F will be used to represent the theoretical structure factor

79

equation with FC being reserved for this expression evaluated at a given trial atomic configuration.

For a structure factor with indices hkl, the structure factor equation of crystallography in the centrosymmetric case may be represented in the form of equation (3.2.7), with

$$X_r = 2\pi(hx_r + ky_r + lz_r) \qquad (3.2.11)$$

where (x_r, y_r, z_r) represent the atomic co-ordinates of the rth atom; the B_r represent the corresponding linear combination of (an)isotopic temperature factor as discussed in Section 1.2, with C_r the associated atomic scattering factor. Although we do not know exactly the true co-ordinates (x_r, y_r, z_r) a Bayesian approach typically furnishes us with prior beliefs expressible in the form

$$
\begin{bmatrix} x_r \\ y_r \\ z_r \end{bmatrix} \sim N \left(\begin{bmatrix} x_{r_0} \\ y_{r_0} \\ z_{r_0} \end{bmatrix} , \begin{bmatrix} \sigma^2_{x_r} & \sigma_{x_r y_r} & \sigma_{x_r z_r} \\ & \sigma^2_{y_r} & \sigma_{y_r z_r} \\ & & \sigma^2_{z_r} \end{bmatrix} \right) .
$$

$$(3.2.12)$$

Using a diagonal approximation to the covariance matrix of equation (3.2.12), and assuming independence between atoms, equations (3.2.11) and (3.2.12) yield equation (3.2.4) with

$$\mu_r = 2\pi(hx_{r_0} + ky_{r_0} + lz_{r_0}) \qquad (3.2.13)$$

and

$$\sigma_r^2 = 4\pi^2(h^2 \sigma_{x_r}^2 + k^2 \sigma_{y_r}^2 + l^2 \sigma_{z_r}^2) \quad . \tag{3.2.14}$$

Similar expressions exist for the α_r and τ_r^2 appearing in equation (3.2.8). The equation (3.2.9) for the expected value of the (true) structure factor is <u>not</u> our usual FC value based on our current approximation, namely

$$FC = \sum_r C_r e^{-\alpha_r} \cos \mu_r \quad . \tag{3.2.15}$$

This represents instead the mode of our distribution for F based on our beliefs expressed by equations (3.2.12), though in practice - small τ_r^2, σ_r^2 - these two quantities will be very similar.

By writing equation (3.2.7) in the form

$$F = 2 \sum_r f_r \cos 2\pi(hx_r + ky_r + lz_r) \tag{3.2.16}$$

where

$$f_r = f_{or} \exp(- B_r \frac{\sin^2 \theta}{\lambda^2}) \quad ,$$

equation (3.2.10) for the variance of our distribution for F gives

$$\text{Var}(F) = 4 \sum_r f_{ro}^2 \left[4\pi^2 (h^2 \sigma_{x_r}^2 + k^2 \sigma_{y_r}^2 + l^2 \sigma_{z_r}^2) \right.$$

$$\cdot \sin^2 2\pi(hx_{r_0} + ky_{r_0} + lz_{r_0}) + \left(\frac{\sin^2\theta}{\lambda^2}\right)^2 \sigma_{B_r}^2$$

$$\left. \cdot \cos^2 2\pi(hx_{r_0} + ky_{r_0} + lz_{r_0}) \right] \tag{3.2.17}$$

where

$$f_{ro} = f_{or} \exp(-\alpha_r \frac{\sin^2\theta}{\lambda^2}) \quad .$$

Equation (3.2.17) is seen to reduce to the form

$$\text{Var}(F) = \sum_j \left[\frac{\partial F}{\partial \theta_j} \Big|_{\underline{\theta}_0} \sigma_j \right]^2 \tag{3.2.18}$$

where $\underline{\theta}$ is our vector of parameters comprising all atomic co-ordinates and isotropic temperature factors, and $\underline{\theta}_0$ represents the mode of the distributions representing our prior beliefs about these parameters, with associated variances σ_j^2 for the estimate θ_{j0} of θ_j. The form of equation (3.2.18) is very important in practice as it is a relatively simple expression based on quantities that are readily available to us: the partial derivatives involved form the appropriate row of the design matrix $\underline{X}^{(0)}$.

We are now in a position to form approximate prediction intervals for structure factors based on our prior

82

beliefs; we use the normal approximation

$$FT \sim N(FC, \sigma_c^2) \hspace{4cm} (3.2.19)$$

where σ_c^2 is given by equation (3.2.18), and we are using the notation FT to represent the true structure factor. We use the conventional value FC in place of that suggested by equation (3.2.9). In accordance with our original motivation, equation (3.2.19) may now enable us to estimate how likely we think the corresponding |FO| value is to have the same sign as FC. The extension of these ideas, in particular the interpretation of an observed |FO| value with a given standard error in the light of prior beliefs of the form of equation (3.2.19), is deferred to the next Section. Here we note that prediction intervals based on (3.2.19) of the form

$$FC \pm c_\alpha \, \sigma_c \hspace{4cm} (3.2.20)$$

- where the constant c_α is determined by the confidence level α - may be used to determine which further (as yet unobserved) reflections to measure. For example, we may wish to measure only when the e.g. 95% confidence interval does not contain zero, so that we are indeed confident of correct sign allocation. We must also note that the observations that are likely to yield the most information will be those where the associated prediction interval is big, since in these cases the theoretical structure factor is sensitive to small shifts in atomic co-ordinate positions. Consequently locating the true structure factor

accurately (amplitude <u>and</u> sign) should tie down these co-ordinate positions. As is often the case in crystallographic studies we have conflicting requirements; here we require prediction intervals that are large, but at the same time one-sided. Note, however, that sensible use of diffraction photographs may be made in the sense that if a prediction interval is predominantly positive with a slight negative tail and we visually observe a high intensity, then a positive sign may safely be assumed.

Our theory above may also be used to justify the crystallographer's intuitive notion of "shells" in reciprocal space based on

$$S = \frac{\sin \theta}{\lambda}$$

values, whereby various ad hoc rules have been developed in which they only consider reflections within specified S-limits at various stages of the refinement procedure (see e.g. [58]). For reflections with similar S-values- for which the factors f_r below will be approximately constant - our structure factor equation of equation (1.2.12)), viz

$$F = 2 \sum_r f_r \cos 2\pi (hx_r + ky_r + lz_r),$$

yields (on average)

$$F^2 \simeq 2 \sum_r f_r^2 .$$

Similarly, by using the approximation

$$h^2 \sigma_{x_r}^2 + k^2 \sigma_{y_r}^2 + l^2 \sigma_{z_r}^2 \simeq (h^2 + k^2 + l^2)\, \bar{\sigma}^2 \qquad (3.2.21)$$

where $\bar{\sigma}^2$ is the average value of $<(\sigma_{x_r}^2 ,\ \sigma_{y_r}^2 ,\ \sigma_{z_r}^2)>_{r}$,

equation (3.2.18) gives (on average)

$$\sigma_c^2 \simeq 8\pi^2 (h^2 + k^2 + l^2)\, \bar{\sigma}^2 \sum_{r} f_r^2 \ . \qquad (3.2.22)$$

The appropriateness of the above strategy is now apparent: by using equations (3.2.20) and (3.2.22) we may address the phase problem by only using those observations for which

$$c_\alpha\, \sigma_c < |F| ,$$

or equivalently,

$$c_\alpha^2\, \sigma_c^2 < F^2 ,$$

which leads us to consider those hkl values that satisfy

$$h^2 + k^2 + l^2 < \frac{1}{4\pi^2\, c_\alpha^2\, \bar{\sigma}^2} \ . \qquad (3.2.23)$$

We should note, however, that here we have used the most general approximation in equation (3.2.21); alternative expressions such as

$$h^2 \sigma^2_{x_r} + k^2 \sigma^2_{y_r} + 1^2 \bar{\sigma}^2_{z_r} \simeq h^2 \bar{\sigma}^2_x + k^2 \bar{\sigma}^2_y + 1^2 \bar{\sigma}^2_z \qquad (3.2.24)$$

where $\bar{\sigma}^2_x$ is the average value of $<(\sigma^2_{x_r})>_r$ etc may also be used, the degree to which we might approximate being dictated by the dimensions of the unit cell present.

Most information will be derived from those reflections for which the (average) relative uncertainty, $\sigma_c/|F|$, hence σ_c^2/F^2, is largest, thus suggesting we take measurements towards the extremities of our limiting 'shell' defined by such as equations (3.2.23). Obviously, a successful refinement, in which $\bar{\sigma}^2$ reduces at each stage will continuously extend our limiting shell, so that ultimately we may be able to incorporate all available data safety into our analysis.

3.3 Further Application of Bayesian Ideas to Crystal Structure Analysis

Inferential statistics is concerned with reaching conclusions extending beyond the immediate data. Statistical decision theory is further concerned with utilising the available information in order to choose among a number of alternative procedures. Bayesian statistics attempts to reduce the amount of uncertainty present in an inferential or decision-making problem by combining new information as it is obtained with any previous information to form the basis for statistical

procedures. The appropriate combination of the new and the old information is achieved via Bayes' theorem (hence the term Bayesian statistics). This shall now be stated and some applications of Bayesian ideas to crystal structure analysis considered.

We consider the one parameter case in which the prior distribution of the continuous random variable $\underline{\theta} = \theta$ is represented by a density function $f(\theta)$. The posterior distribution $f(\theta/y)$ is the conditional density of θ, given the observed value y of the sample statistic. Bayes' theorem for continuous random variables states that [73]

$$f(\theta/y) \propto f(y/\theta) \, f(\theta) \qquad\qquad (3.3.1)$$

where $f(y/\theta)$ is often termed the likelihood function and represents the conditional density of our sample statistic, given the parameter value (see Section 4.2). The general (multivariate) form of Bayes' theorem is the natural analogue of equation (3.3.1), stating that the posterior distribution of (the vector of parameters) $\underline{\theta}$ is proportional to the product of the prior distribution and the likelihood function.

Our first application of the above ideas concerns the incorporation of any relevant prior beliefs into our crystal structure analysis. For example, based on the results of previous related studies, we may have reason to believe that

$$\underset{\sim}{\theta} \sim N(\underset{\sim}{\theta}^{(0)}, \underset{\sim}{\Sigma}^{(0)}) \qquad\qquad (3.3.2)$$

where the covariance matrix $\underset{\sim}{\Sigma}^{(0)}$ is known and non-singular, so that

$$\underset{\sim}{\Delta\theta}^{(0)} = \underset{\sim}{\theta} - \underset{\sim}{\theta}^{(0)} \sim N(\underset{\sim}{0}, \underset{\sim}{\Sigma}^{(0)}) \quad . \qquad\qquad (3.3.3)$$

Following the Taylor series approach of Section 2.2 and expanding about $\underset{\sim}{\theta}^{(0)}$ at the first stage of our iteration procedure, we obtain the observational equations

$$\underset{\sim}{Y}^{(0)}/\underset{\sim}{\theta} \sim N(\underset{\sim}{X}^{(0)}\underset{\sim}{\Delta\theta}^{(0)}, \underset{\sim}{V}) \quad . \qquad\qquad (3.3.4)$$

[Note that here we are using the notation for $\underset{\sim}{X}^{(0)}$, $\underset{\sim}{Y}^{(0)}$ defined by equations (2.2.9), (2.2.10)]. Bayesian analysis [53] yields the posterior distribution for the parameter increment $\underset{\sim}{\Delta\theta}^{(0)}$, given observed data $\underset{\sim}{Y}^{(0)}$ = $\underset{\sim}{y}^{(0)}$ to be

$$\underset{\sim}{\Delta\theta}^{(0)} \sim N(\underset{\sim}{D}^{(0)}\underset{\sim}{d}^{(0)}, \underset{\sim}{D}^{(0)}) \qquad\qquad (3.3.5)$$

where

$$[\underset{\sim}{D}^{(0)}]^{-1} = \underset{\sim}{X}^{(0)T} \; \underset{\sim}{W} \; \underset{\sim}{X}^{(0)} + [\underset{\sim}{\Sigma}^{(0)}]^{-1} \qquad\qquad (3.3.6)$$

and

$$\underset{\sim}{d}^{(0)} = \underset{\sim}{X}^{(0)T} \; \underset{\sim}{W} \; \underset{\sim}{y}^{(0)} \quad . \qquad\qquad (3.3.7)$$

Note that now $\underset{\sim}{W} = \underset{\sim}{V}^{-1}$. [These formulae pertain with

$\underset{\sim}{x}^{(0)}$, $\underset{\sim}{y}^{(0)}$ having their original specification (1.5.4), (1.5.5)]. Note further that as $[\underset{\sim}{\Sigma}^{(0)}]^{-1}$ approaches the zero matrix, consideration of the mode of the posterior distribution (3.3.5) yields our previous least squares results (for the case $\underset{\sim}{W} = \underset{\sim}{V}^{-1}$); however, such cases – unbounded $\underset{\sim}{\Sigma}^{(0)}$ – are unrealistic. Merely by specifying $\underline{\theta}^{(0)}$ we are expressing a preference for the true underlying parameter to lie in this region of parameter space, and, in any case, since fractional positional co-ordinates necessarily lie between -1 and +1 we automatically have an upper bound on the corresponding parameter variances. It is not unreasonable to expect $\underset{\sim}{\Sigma}^{(0)}$ to be accurately specified.

Naturally enough, we use the posterior mean for $\underline{\Delta\theta}^{(0)}$ to obtain

$$\underline{\theta}^{(1)} = \underline{\theta}^{(0)} + \underset{\sim}{D}^{(0)} \underset{\sim}{d}^{(0)} . \tag{3.3.8}$$

The procedure may be repeated, maintaining our prior beliefs of equation (3.3.2), but with our least squares approximate linearisation now being about $\underline{\theta}^{(1)}$. Our observational equations become

$$\underset{\sim}{y}^{(1)} \sim N (\underset{\sim}{X}^{(1)} \underline{\Delta\theta}^{(1)}, \underset{\sim}{V}); \tag{3.3.9}$$

equation (3.3.9) may be combined with the prior beliefs

$$\underline{\Delta\theta}^{(1)} = \underline{\theta} - \underline{\theta}^{(1)} \sim N(\underline{\theta}^{(0)} - \underline{\theta}^{(1)}, \underset{\sim}{\Sigma}^{(0)}) \tag{3.3.10}$$

89

to obtain a new posterior mean and covariance matrix for $\underline{\Delta\theta}^{(1)}$, given observed data $\underline{Y}^{(1)} = \underline{y}^{(1)}$. We obtain the result [53]

$$\underline{\Delta\theta}^{(1)} \sim N(\underline{D}^{(1)} \ \underline{d}^{(1)}, \ \underline{D}^{(1)}) \qquad\qquad (3.3.11)$$

where

$$[\underline{D}^{(1)}]^{-1} = \underline{X}^{(1)T} \ \underline{W} \ \underline{X}^{(1)} + [\underline{\Sigma}^{(0)}]^{-1} \qquad\qquad (3.3.12)$$

and

$$\underline{d}^{(1)} = \underline{X}^{(1)T} \ \underline{W} \ \underline{y}^{(1)} + [\underline{\Sigma}^{(0)}]^{-1}(\underline{\theta}^{(0)} - \underline{\theta}^{(1)}) \ . \qquad (3.3.13)$$

The extra term appearing in equation (3.3.13) arises due to the non-zero prior mean appearing in equation (3.3.10). The process of substitution and re-estimation may be continued until resulting increments become negligible in some specified sense. The resultant procedure may be represented as in Figure 3.3.1.

At this stage it seems sensible to comment on the prior information that may be available and how it can be incorporated into the Bayesian framework. The three most important types of prior information in crystallographic studies are that

1)   Certain bonded atoms are at pre-assigned distances from each other

$$\underline{\theta}^{(0)}, \ \underline{\Delta\theta}^{(0)} \ \sim \ \mathbf{N}(\underline{o}, \ \underline{\Sigma}^{(0)})$$

$$k = 0$$

Evaluate $\underline{X}^{(k)}, \ \underline{Y}^{(k)}$

$$[\underline{D}^{(k)}]^{-1} = \underline{X}^{(k)T} \ \underline{W} \ \underline{X}^{(k)} + [\underline{\Sigma}^{(0)}]^{-1}$$

$$\underline{d}^{(k)} = \underline{X}^{(k)T} \underline{W} \ \underline{Y}^{(k)} + [\underline{\Sigma}^{(0)}]^{-1}(\underline{\theta}^{(0)} - \underline{\theta}^{(k)})$$

k=k+1

$$\underline{\Delta\theta}^{(k)} = \underline{D}^{(k)} \ \underline{d}^{(k)}$$

$$\underline{\theta}^{(k+1)} = \underline{\theta}^{(k)} + \underline{\Delta\theta}^{(k)}$$

NO

? Accept $\underline{\theta}^{(k+1)}$?

YES

STOP
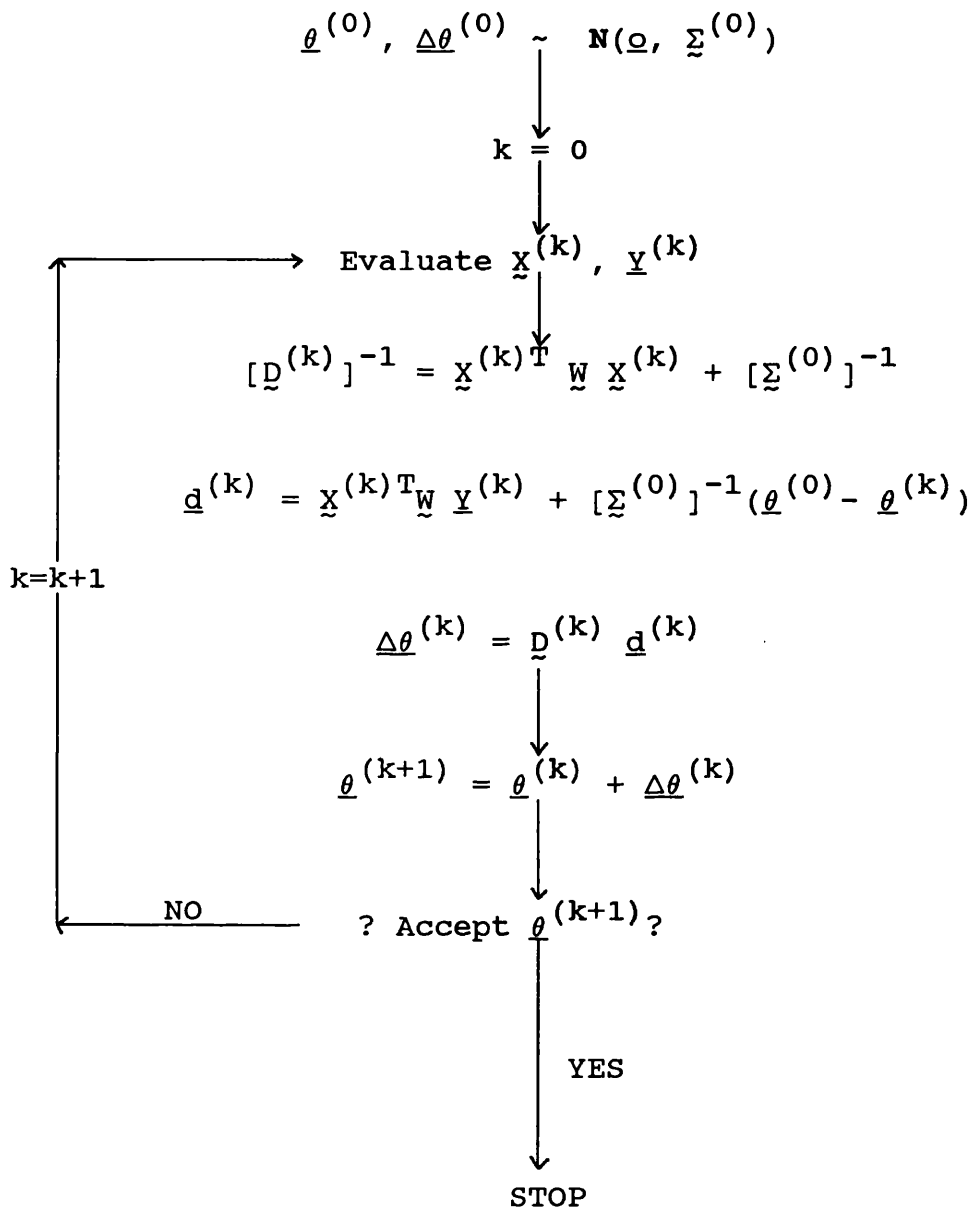
Figure 3.3.1: Incorporation of Bayesian beliefs of the form of equation (3.3.2) into a standard least squares cycle

2)   Certain atom triplets define pre-assigned bonding
      angles

3)   Certain atoms are co-planar.

These conditions are often imposed on the atomic
parameters to ensure that they represent a given molecule.
One way of utilising such information is to express 2) and
3) as distance requirements also - in terms of the three
distances involved and specification of bonded/non-bonded
atomic distances in the given plane respectively - and
proceed along the lines proposed in [67]. We shall very
briefly outline this approach before reverting to our
Bayesian ideas.

The approach proposed in [67] is to treat the subsidiary
equations obtained like observational equations and modify
the least squares refinement method accordingly. The new
least squares equations obtained are

$$\frac{\partial R_1}{\partial \theta_j} + \frac{\partial R_2}{\partial \theta_j} = 0 \qquad (j=1\ldots p) \qquad\qquad (3.3.14)$$

where

$$R_1 = \sum_{r,s} w_{rs} ([(x_r - x_s)^2 + (y_r - y_s)^2$$

$$+ (z_r - z_s)^2]^{\frac{1}{2}} - d_{rs})^{\frac{1}{2}} \qquad\qquad (3.3.15)$$

with d_{rs} being the appropriate prescribed distances for
select pairs of atoms r and s, and w_{rs} the associated

weights; R_2 corresponds to the usual residual sum of squares ψ of equation (2.2.1). Iterative least squares schemes may now proceed along the same lines as before.

The Bayesian assumes that conditions 1), 2) and 3) may typically be expressed in the form

$$\underline{P} \sim \mathbf{N} \ (\underline{\mu}_p \ , \ \underset{\sim}{\Sigma}_p) \tag{3.3.16}$$

where $\underline{P}$ represents our distances (non-linear functions of $\underline{\theta}$) and the covariance matrix $\underset{\sim}{\Sigma}_p$ is known and non-singular. As with our observational equations, the non-linear functions of $\underline{\theta}$ appearing in $\underline{P}$ may be expanded in a 1st order Taylor series so that equation (3.3.16) yields the approximation

$$X\underset{\sim}{P}^{(0)} \ \underline{\Delta\theta}^{(0)} \sim \mathbf{N} \ (\underline{\mu}_p - \underline{P}^{(0)}, \ \underset{\sim}{\Sigma}_p) \tag{3.3.17}$$

where $\underline{P}^{(k)}$ represents $\underline{P}$ evaluated at the parameter value $\underline{\theta} = \underline{\theta}^{(k)}$ and likewise $X\underset{\sim}{P}^{(k)}$ the associated design matrix. Then by defining (assuming it exists)

$$\underset{\sim}{H}^{(k)} = [X\underset{\sim}{P}^{(k)T} \ X\underset{\sim}{P}^{(k)}]^{-1} \ X\underset{\sim}{P}^{(k)T} \tag{3.3.18}$$

equation (3.3.17) suggests

$$\underline{\Delta\theta}^{(0)} \sim \mathbf{N} \ (\underset{\sim}{H}^{(0)} (\underline{\mu}_p - \underline{P}^{(0)}), \ \underset{\sim}{H}^{(0)T} \ \underset{\sim}{\Sigma}_p \ \underset{\sim}{H}^{(0)}) \ . \tag{3.3.19}$$

Equation (3.3.19) may be combined with the observational equation (3.3.4) to obtain the following posterior distribution for our parameter increment, given observed

93

data $\underline{y}^{(0)} = y^{(0)}$:

$$\Delta \underline{\theta}^{(0)} \sim N \left(\underline{D}_p^{(0)} \underline{d}_p^{(0)}, \underline{D}_p^{(0)} \right) \qquad (3.3.20)$$

where

$$[\underline{D}_p^{(0)}]^{-1} = \underline{X}^{(0)T} \underline{W} \underline{X}^{(0)} + [\underline{H}^{(0)T} \Sigma_p \underline{H}^{(0)}]^{-1} \qquad (3.3.21)$$

and

$$\underline{d}_p^{(0)} = \underline{X}^{(0)T} \underline{W} \underline{y}^{(0)} + [\underline{H}^{(0)T} \Sigma_p \underline{H}^{(0)}]^{-1} \underline{H}^{(0)} (\underline{\mu}_p - \underline{P}^{(0)}) \ .$$

$$(3.3.22)$$

The iterative cycle of substitution and re-estimation may now proceed similar to before.

In practice, however, the above approach is likely to be infeasible since the number of distance requirements specified will typically be less than the number of parameters to be estimated; this renders the matrix

$$\underline{XP}^{(k)T} \underline{XP}^{(k)}$$

singular and equation (3.3.18) consequently becomes meaningless. To rectify this, we may seek to incorporate prior beliefs of the form of equation (3.3.2) into equation (3.3.16) and proceed as above. However, the chemical prior information available as expressed by equation (3.3.16) will generally differ from the distribution for $\underline{P}$ derived from equation (3.3.2) alone. Bearing this in mind, the following alternative approach is proposed in such cases which seeks first to amalgamate

94

the distributions of equations (3.3.2) and (3.3.16) into a single reconciled distribution.

Suppose as before we have a prior distribution for $\underset{\sim}{\theta}$ given by equation (3.3.2). Then

$$X \underset{\sim}{P}^{(0)} \, \underset{\sim}{\Delta\theta}^{(0)} \sim N \, (\underset{\sim}{0}, \, X\underset{\sim}{P}^{(0)} \, \underset{\sim}{\Sigma}^{(0)} \, X\underset{\sim}{P}^{(0)T}) \qquad (3.3.23)$$

where $\underset{\sim}{\Delta\theta}^{(0)} = \underset{\sim}{\theta} - \underset{\sim}{\theta}^{(0)}$. Providing

$$\underset{\sim}{M}^{(0)} = [X\underset{\sim}{P}^{(0)} \, \underset{\sim}{\Sigma}^{(0)} \, X\underset{\sim}{P}^{(0)T}]^{-1} \qquad (3.3.24)$$

exists, we obtain the conditional distribution

$$\Delta\theta / XP \, \Delta\theta \sim N(\Sigma \, XP^T \, M \, XP \, \Delta\theta, \, \Sigma - \Sigma \, XP^T \, M \, XP \, \Sigma) \qquad (3.3.25)$$

(where all $^{(0)}$s and $\sim/$-s have been dropped for notational convenience). By using the approximation (3.3.17) to equation (3.3.16), namely

$$XP \, \Delta\theta \sim N \, (\mu_p - P_p, \, \Sigma_p)$$

(with $P_p = P^{(0)}$), we find that

$$\Sigma \, XP^T \, M \, XP \, \Delta\theta \sim N(\Sigma \, XP^T \, M(\mu_p - P_p), \, \Sigma \, XP^T \, M \, \Sigma_p \, M \, XP \, \Sigma) \quad .$$

$$(3.3.26)$$

Retaining the conditional distribution (3.3.25), i.e.

$$\Delta\theta - \Sigma \, XP^T \, M \, XP \, \Delta\theta \sim N(o, \, \Sigma - \Sigma \, XP^T \, M \, XP \, \Sigma) \qquad (3.3.27)$$

independently of equation (3.3.26) we get

$$\Delta\theta \sim \mathbf{N}(\Sigma \ XP^T \ M(\mu_p - P_p), \ \Sigma + \Sigma \ XP^T \ M(\Sigma_p - M^{-1})M \ XP \ \Sigma) \quad .$$

$$(3.3.28)$$

We may proceed to cycle similar to before; the analogue of equation (3.3.23) for $XP^{(k)} \underline{\Delta\theta}^{(k)}$ will have non-zero mean for $k > 0$ and the ensuing results must be modified accordingly. The resultant scheme is represented in Figure 3.3.2. Having accepted $\underline{\theta}^{(k)}$ at some stage of our iteration procedure, we may update the covariance matrix in accordance with equation (3.3.28) to obtain the beliefs

$$\underline{\theta} \sim \mathbf{N}(\underline{\theta}^{(0)} + \underset{\sim}{\Sigma}^{(0)} \ X\underset{\sim}{P}^{(k)T} \ \underset{\sim}{M}^{(k)} [(\mu_p - \underline{P}^{(k)})$$

$$- X\underset{\sim}{P}^{(k)}(\underline{\theta}^{(0)} - \underline{\theta}^{(k)})],$$

$$\underset{\sim}{\Sigma}^{(0)} + \underset{\sim}{\Sigma}^{(0)} \ X\underset{\sim}{P}^{(k)T} \ \underset{\sim}{M}^{(k)} [\Sigma_p - (\underset{\sim}{M}^{(k)})^{-1}]\underset{\sim}{M}^{(k)} \ X\underset{\sim}{P}^{(k)} \ \underset{\sim}{\Sigma}^{(0)}) \quad .$$

$$(3.3.29)$$

We may then use these beliefs (3.3.29) in conjunction with the observational data as discussed previously.

We shall now turn our attention towards the following alteration to existing methods of structure analysis. Rather than incorporating an (appropriately signed) observed structure factor into our least squares analysis with a weight based on the associated experimental error alone, we shall use a Bayesian procedure to obtain a more realistic weighting scheme (based on an improved estimate of the variance of the signed structure factor). Such a

$$\underline{\theta}^{(0)}, \; \underline{\Delta\theta}^{(0)} \sim N(\underline{o}, \; \underset{\sim}{\Sigma}^{(0)}), \; \underline{P} \sim N(\underline{\mu}_p, \; \underset{\sim}{\Sigma}_p)$$

$$k = 0$$

Evaluate $X\underset{\sim}{P}^{(k)}, \; \underline{P}^{(k)}$

$$\underset{\sim}{M}^{(k)} = [X\underset{\sim}{P}^{(k)} \; \underset{\sim}{\Sigma}^{(0)} \; X\underset{\sim}{P}^{(k)T}]^{-1}$$

$$\underline{\Delta\theta}^{(k)} = (\underline{\theta}^{(0)} - \underline{\theta}^{(k)}) + \underset{\sim}{\Sigma}^{(0)}X\underset{\sim}{P}^{(k)}\underset{\sim}{M}(k)[(\underline{\mu}_p - \underline{P}^{(k)}) - X\underset{\sim}{P}^{(k)}(\underline{\theta}^{(0)} - \underline{\theta}^{(k)})]$$

k=k+1

$$\underline{\theta}^{(k+1)} = \underline{\theta}^{(k)} + \underline{\Delta\theta}^{(k)}$$

NO

? Accept $\underline{\theta}^{(k+1)}$?

YES

Accept (3.3.28)

STOP
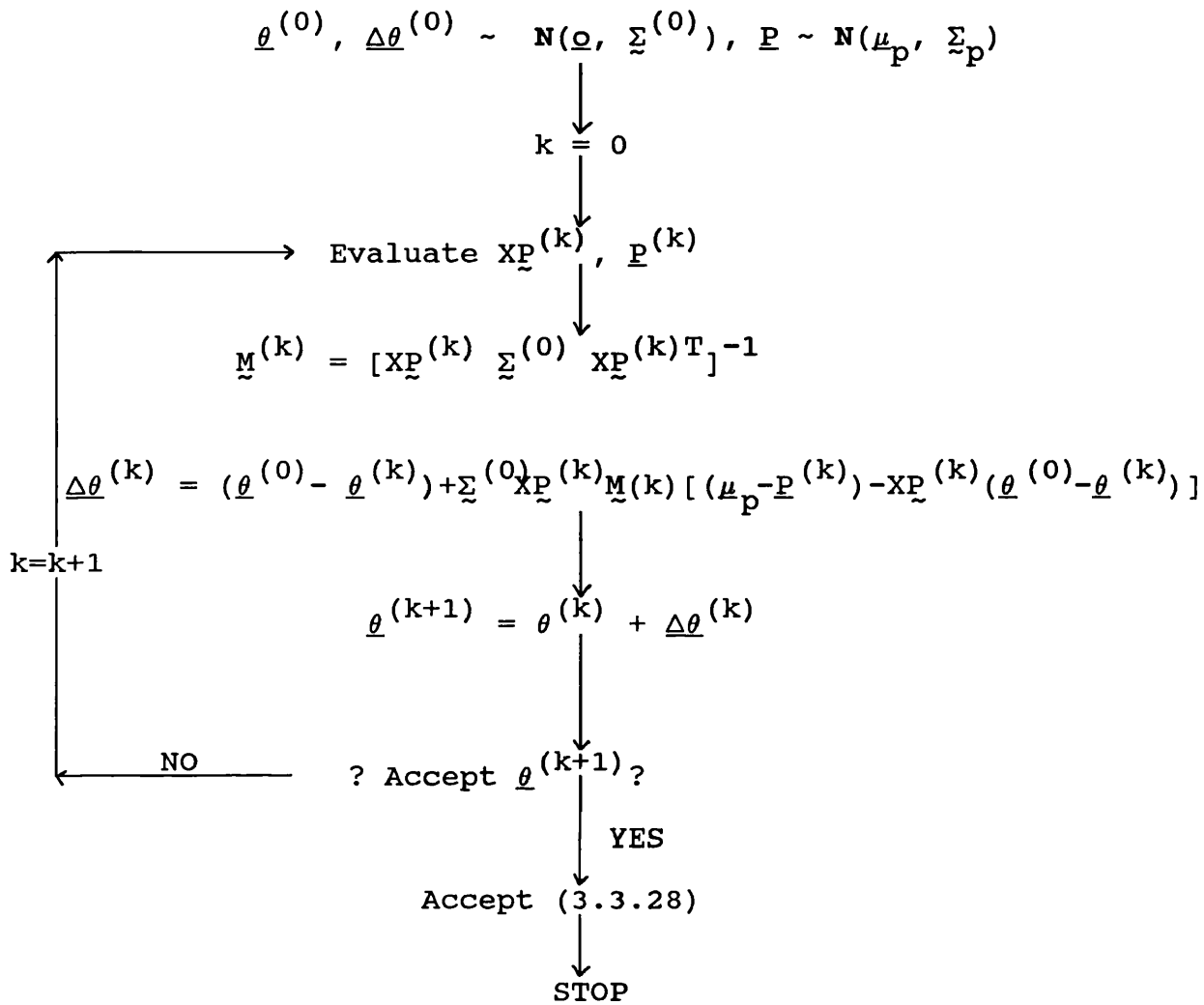
Figure 3.3.2:   Amalgamation of Bayesian beliefs of the form of equation (3.3.2) and prior chemical information of the form of equation (3.3.16)

scheme may influence the data entering the least squares procedure at any stage, with preference going to those observed values with the larger associated weights.

Let us suppose that we are sampling from a univariate normal population with known variance σ_T^2, but unknown mean μ_T. Suppose further that we have a prior distribution for the unknown quantity μ_T that is normal with parameters $\mu^{(0)}$ and $\sigma^{(0)2}$. If we observe a sample mean m from a sample of size n, the posterior distribution for μ_T given the sample results is again normal. Formulae for the new parameters may be found in e.g. [73]. In the crystallographic setting, for a given hkl reflection, let μ_T represent the true structure factor FT (say), with the parameters corresponding to $\mu^{(0)}$ and $\sigma^{(0)2}$ being given by the FC and σ_c^2 of equation (3.2.19). The observation y(= m) corresponds to our observed $|FO(\underline{h})|$ value (n = 1), and we assume our error variance $\sigma_o^2 = \sigma_T^2$ is known prior to measurement. The above theory will <u>not</u> pertain directly, however, because of the phase problem, the effect of which is to generate the model (c.f. (1.5.3))

$$|FO(\underline{h})|/FT = ft \sim N(|ft|, \sigma_o^2) \quad . \qquad (3.3.30)$$

Note the modulus now appearing in the mean term. It does not seem unreasonable to expect our observed structure amplitude to be unbiased on the modulus of the true structure factor; furthermore, symmetricity of errors seems a natural condition to impose in cases where they

are small relative to the mean term so that possible skewness due to the necessary non-negativity does not have a significant effect. Under such circumstances, the above normal approximation may be justified and consequently is considered to represent a realistic model (further related comments appear in Section 6.1).

Using the prior distribution (3.2.19) for FT, viz

$$FT \sim N(FC, \sigma_c^2)$$

the above approach will need to be modified accordingly. This is achieved as follows. By replacing FT by X for notational convenience, equations (3.3.30) and (3.2.19) may be combined via Bayes' theorem to give the following posterior distribution for X:

$$f_{X/|FO(\underline{h})|} = |FO|^{(x)} \alpha \exp\left[-\frac{1}{2\sigma_c^2}(x-FC)^2\right.$$

$$\left. -\frac{1}{2\sigma_o^2}(|FO| - |x|)^2\right] . \qquad (3.3.31)$$

The posterior mean and variance (μ_1, and σ_1^2, (say)) may be evaluated via

$$\mu_1 = \frac{I_1}{I_0} \qquad (3.3.32)$$

and

$$\sigma_1^2 = \frac{I_2}{I_0} - \mu_1^2 \qquad (3.3.33)$$

99

where

$$I_r = \int_{-\infty}^{\infty} x^r \exp\left[\frac{-1}{2\sigma_c^2}(x-FC)^2 - \frac{1}{2\sigma_o^2}(|FO| - |x|)^2\right]dx$$

$$(r=0,1,2). \tag{3.3.34}$$

Because of the $|x|$ term appearing in (3.3.34), I_r must be decomposed into the sum of the two integrals

$$\int_{-\infty}^{0} x^r \exp\left[\frac{-1}{2\sigma_c^2}(x-FC)^2 - \frac{1}{2\sigma_o^2}(|FO| + x)^2\right]dx \tag{3.3.35}$$

and

$$\int_{0}^{\infty} x^r \exp\left[\frac{-1}{2\sigma_c^2}(x-FC)^2 - \frac{1}{2\sigma_o^2}(|FO| - x)^2\right]dx . \tag{3.3.36}$$

The results of the above calculations yield

$$\mu_1 = \frac{\sigma_o^2}{\sigma_o^2 + \sigma_c^2} FC + \frac{\sigma_c^2}{\sigma_o^2 + \sigma_c^2} |FO|(\tan h \ \xi)$$

$$+ \frac{\phi\left[\frac{\mu^+}{\sigma}\right] e^a - \phi\left[\frac{\mu^-}{\sigma}\right] e^{-a}}{\Phi\left[\frac{\mu^+}{\sigma}\right] e^a + \Phi\left[\frac{\mu^-}{\sigma}\right] e^{-a}} \ \sigma \tag{3.3.37}$$

and

$$\sigma_1^2 = \sigma^2 + \left[\frac{\sigma_o^2 \text{ FC}}{\sigma_o^2 + \sigma_c^2}\right]^2 + \left[\frac{\sigma_c^2 |\text{FO}|}{\sigma_o^2 + \sigma_c^2}\right] + \frac{2\sigma_o^2 \sigma_c^2 |\text{FO}| \text{ FC}}{(\sigma_o^2 + \sigma_c^2)^2}$$

$$\cdot (\tan h \ \xi) + \frac{\mu^+ \phi(\frac{\mu^+}{\sigma}) \ e^a + \mu^- \phi(\frac{\mu^-}{\sigma}) \ e^{-a}}{\Phi(\frac{\mu^+}{\sigma}) \ e^a + \Phi(\frac{\mu^-}{\sigma}) \ e^{-a}} \ \sigma - \mu_1^2$$

$$(3.3.38)$$

where

$$\xi = a + \tfrac{1}{2} \log \left[\frac{\Phi(\frac{\mu^+}{\sigma})}{\Phi(\frac{\mu^-}{\sigma})}\right] \ , \qquad (3.3.39)$$

$$\mu^{\pm} = \frac{\sigma_c^2 |\text{FO}| \pm \sigma_o^2 \text{ FC}}{\sigma_o^2 + \sigma_c^2} \ , \qquad (3.3.40)$$

$$\sigma = \frac{\sigma_c \ \sigma_o}{(\sigma_o^2 + \sigma_c^2)^{\frac{1}{2}}} \ , \qquad (3.3.41)$$

and

$$a = \frac{|\text{FO}| \ \text{FC}}{\sigma_o^2 + \sigma_c^2} \ ; \qquad (3.3.42)$$

furthermore we use the standard notation

$$\phi(z) = \frac{1}{\sqrt{2\pi}} \ e^{-z^2/2}$$

and

$$\Phi(z) = \int_{-\infty}^{z} \phi(t) \ dt$$

101

for the density and distribution functions of $N(o,1)$ random variables.

Were we to replace the integrals appearing in (3.3.35) and (3.3.36) by integrals over the complete range $-\infty$ to ∞, we would obtain the approximations

$$\mu_1 = \frac{\sigma_o^2 \, FC}{\sigma_o^2 + \sigma_c^2} + \frac{\sigma_c^2}{\sigma_o^2 + \sigma_c^2} \, |FO| \; (\tan h \; a) \qquad (3.3.43)$$

and

$$\sigma_1^2 = \sigma^2 + \left[\frac{\sigma_c^2 |FO|}{\sigma_o^2 + \sigma_c^2} \right]^2 (\sec h^2 \; a) \quad . \qquad (3.3.44)$$

Such approximations may loosely be justified on the grounds that the terms $(|FO| + x)^2$ and $(|FO|-x)^2$ appearing in the respective exponential exponents will become large enough to render the integrand negligible over the extension of the range. Comparison of exact and approximate results for typical $|FO|$, σ_o^2, FC, σ_c^2 values reveals great similarity (see Table 3.3.1), the main differences occurring when $|FO| \sim \sigma_o^2$ and $|FC| \sim \sigma_c^2$. We shall henceforth use the approximate results (3.3.43) and (3.3.44).

The first – Bayesian – stage of our present analysis is now over. The second stage consists of considering how to utilise the information afforded us by knowledge of μ_1 and σ_1^2. In Chapter 2 we used

102

				Exact			Approximate		
|FO|	σ_0	FC	σ_c	μ_1	σ_1	$100 \times [(FO*-\mu_1)^2 + \sigma_1^2]^{-1}$	μ_1	σ_1	$100 \times [(FO*-\mu_1)^2 + \sigma_1^2]$
.1	.1	2	2	.0114	.1507	3572	.0096	.1411	3570
2	.1	-2	2	-1.522	1.299	52.19	-1.522	1.299	52.19
2	2	2	2	1.677	1.734	32.09	1.462	1.669	32.51
2	.1	-2	5	-0.318	1.976	14.85	-0.318	1.976	14.85
5	1	10	5	4.991	1.690	35.03	4.991	1.690	35.03
20	4	-1	10	-3.081	17.390	0.17	-3.081	17.390	0.17

Table 3.3.1:   Comparison of some exact and approximate posterior moments for the true structure factor based on equations (3.3.37), (3.3.38) and (3.3.43), (3.3.44) respectively, together with weights of the form of equation (3.3.45)

$$FO* = |FO| \times \text{sign } FC$$

in the least squares analysis with weights $w \propto 1/\sigma_o^2$ determined by observational errors. These weights are imposed to quantify how reliable FO* is as an estimate for FT, assuming the sign is correct and in the light of observational error. In view of our posterior moments for FT, it is now possible to quantify this reliability in terms of variance about FO*, suggesting the modified weighting scheme

$$w \propto \frac{1}{\sigma_1^2 + (FO* - \mu_1)^2} \quad . \qquad (3.3.45)$$

Standard iterative procedures may be invoked as before; the implementation of such schemes forms the third stage of our procedure.

Care must be taken whenever (3.3.45) is to be used as a dynamic weighting scheme, in which the weights are continually updated as our FC, σ_c^2 values of equation (3.2.19) change; in keeping with our original Bayesian ideals, such updated 'prior' information should be combined with new data (not the original $|FO|, \sigma_o^2$ values) in the evaluation of the appropriate weights (3.3.45). Once the final batch of measurements have been obtained a static weighting scheme should be employed and our problem reduced to that of minimising a criterion function of the form of (2.2.2). Crystallographers maintain the original

data, however, when using similar dynamic weighting schemes of their own - see Section 5.2 - which although being a more practicable approximation to the approach above, is theoretically not as sound.

Note also that our theory is only an approximate one. Although, by considering individual structure factors we may not be using all the information available to us in an optimal manner (see Section 6.1), we are nevertheless making sensible use of relevant information that previously lay dormant. This fact, coupled with the relative tractability of our computations and consequent practicability of our approach, makes it a worthwhile one to adopt in practice.

The approach above may be further tuned by paying attention to the following points:

1)   We will not in practice know the population variance σ_T^2

2)   Our $|FO|$ value and associated σ_o^2 will in fact represent the mean and variance of (typically) n = 4 observations.

These points may be addressed as follows. Replacing σ_T^2 by Y for notational convenience, we may replace equation (3.2.19) - recall X = FT - by the prior joint density

$$f_{X,Y}(x,y) = \frac{1}{(2\pi\, y/n_0)^{\frac{1}{2}}} \exp\left[\frac{-(x-m_0)^2}{2y}\, n_0\right]$$

$$\cdot \frac{\exp\left[\frac{-vd}{2y}\right] \left(\frac{vd}{2y}\right)^{d/2-1} \left(\frac{vd}{2}\right)}{\Gamma\left(\frac{d}{2}\right)} \qquad (3.3.46)$$

where $\Gamma(x) = \int_0^\infty t^{x-1} e^{-t}\, dt$

represents the gamma function. The two components of the right hand side of (3.3.46) have the following respective interpretations:

1) The conditional distribution of X given Y = y is normally distributed with parameters m_0 and y/n_0

2) The gamma distribution with parameters $d/2$ and $vd/2$ is the marginal distribution of the reciprocal of the process variance, $1/Y$.

Prior beliefs for the first two moments of the process variance being e and var respectively, would suggest

$$v = e, \qquad d = 2 + \frac{2e^2}{var}. \qquad (3.3.47)$$

Our original prior beliefs (3.2.19) would now suggest

$$m_0 = FC, \qquad n_0 = \frac{e}{\sigma_c^2}. \qquad (3.3.48)$$

Now suppose that given X=x and Y=y, $Z_1 \ldots Z_n$ are independent and identically distributed $N(|x|, y)$ random variables. Then we have [35] the distributional results

$$\bar{Z} \sim N(|x|, \frac{y}{n}) \qquad\qquad (3.3.49)$$

and

$$Z_{var} \sim \frac{y}{n-1} \chi^2(n-1) \qquad\qquad (3.3.50)$$

where

$$\bar{Z} = \sum_{i=1}^{n} \frac{Z_i}{n} \quad , \qquad Z_{var} = \sum_{i=1}^{n} \frac{(Z_i - \bar{Z})^2}{n-1} \quad . \qquad (3.3.51)$$

The quantities $\bar{Z}$, Z_{var} are independent, and consequently have joint density function (conditional on X=x and Y=y)

$$f_{\bar{Z}, Z_{var}/X=x, Y=y}(\bar{z}, z_{var}) = \frac{1}{(2\pi y/n)^{\frac{1}{2}}} \exp[-\frac{(\bar{z} - |x|)^2 n}{2y}]$$

$$\cdot \frac{(\frac{1}{2})^{n-1} \exp[-\frac{(n-1)}{2y} z_{var}] [\frac{(n-1)}{y} z_{var}]^{(n-3)/2}}{\Gamma(\frac{n-1}{2})} \frac{n-1}{y} \quad .$$

$$(3.3.52)$$

(3.3.52) is the analogue of our original model equation (3.3.30); z and z_{var} represent our measured quantities $|FO|$ and σ_o^2 respectively. Equations (3.3.46) and (3.3.52) may now be combined via Bayes' theorem to derive the joint posterior distribution for X and Y given observed values $|FO|$ and σ_o^2. The calculations are similar to those

previously derived and yield the following approximate results.

Conditional on $\sigma_T^2 = \sigma_t^2$, FT has posterior mean and variance, μ_2 and σ_2^2 (say), given by

$$\mu_2 = \frac{e\,FC}{n\sigma_c^2 + e} + \frac{n\sigma_c^2 |FO|}{n\sigma_c^2 + e} \tan h\left[\frac{|FO|\,FC}{\sigma_t^2} \frac{ne}{n\sigma_c^2 + e}\right] \qquad (3.3.53)$$

and

$$\sigma_2^2 = \frac{\sigma_c^2 \sigma_t^2}{n\sigma_c^2 + e} + \left[\frac{n\sigma_c^2 |FO|}{n\sigma_c^2 + e}\right]^2 \sec h^2\left[\frac{|FO|\,FC}{\sigma_t^2} \frac{ne}{n\sigma_c^2 + e}\right].$$

$$(3.3.54)$$

Note that these reduce to equations (3.3.43) and (3.3.44) when $n=1$ and $\sigma_t^2 = e = \sigma_o^2$. The posterior marginal distribution of $1/\sigma_T^2$ is gamma with parameters

$$\tfrac{1}{2}(n + 3 + \frac{2e^2}{var}) \qquad \text{and} \qquad \frac{(n-1)\sigma_o^2}{2} + e(1 + \frac{e^2}{var}) \qquad (3.3.55)$$

which suggests taking σ_t^2 to be the posterior mean for σ_T^2, viz

$$\sigma_t^2 = \frac{(n-1)\sigma_o^2\, var + 2e(var + e^2)}{(n+3)\, var + 2e^2} \qquad (3.3.56)$$

in the above equations. Some typical results are shown in Table 3.3.2.

We have already seen how the results of this Section may be used to modify our least squares weights.

n	|FO|	σ_o	FC	σ_c	e	var	μ_2	σ_2	$100 \times [FO* - \mu_2)^2 + \sigma_2^2]^{-1}$
1	2	2	2	2	4	0	1.462	1.669	32.51
4	2	2	-2	2	4	0	-1.462	1.494	39.68
4	5	.1	5	.1	.01	.002	5	0.038	69610
4	5	1	-10	5	1	.01	-4.871	1.406	50.13
8	10	2	20	20	9	4	5.144	8.634	1.02
4	20	4	-1	10	9	4	-3.774	19.254	0.16

Table 3.3.2: Some posterior moments for the true structure factor based on equations (3.3.53), (3.3.54), together with weights of the form of equation (3.3.45)

Alternatively, in the spirit of the previous Section, we may wish to determine the likelihood of correct sign allocation by considering such quantities as $P(FT > 0)$ based on the posterior distribution, and only use the corresponding FO* value if this probability is sufficiently big/small. The following probabilistic argument may be employed.

In cases where ft is positive, (3.3.30) gives the conditional distribution

$$\left| FO(\underline{h}) \right| / FT = ft \sim N(ft, \sigma_o^2) \ .$$

This may be combined with the prior distribution (3.2.19) to give the posterior distribution for the true structure factor

$$FT / \left| FO(\underline{h}) \right| = \left| FO \right| \sim N(\sigma^2 (\frac{FC}{\sigma_c^2} + \frac{\left| FO \right|}{\sigma_o^2}), \ \sigma^2) \qquad (3.3.57)$$

where σ is given by (3.3.41). Similarly, when ft is negative, so that (3.3.30) gives

$$\left| FO(\underline{h}) \right| / FT = ft \sim N(-ft, \ \sigma_o^2),$$

we obtain

$$FT / \left| FO(\underline{h}) \right| = \left| FO \right| \sim N(\sigma^2 (\frac{FC}{\sigma_c^2} - \frac{\left| FO \right|}{\sigma_o^2}), \ \sigma^2) \ . \qquad (3.3.58)$$

If we denote the probability of the true structure factor being positive by p, we would expect the posterior

distributions (3.3.57) and (3.3.58) to pertain with probabilities p and (1-p) respectively. This gives rise to the expected posterior mean

$$\frac{\sigma_o^2 \ FC}{\sigma_o^2 + \sigma_c^2} + (2p-1) \ \frac{\sigma_c^2 \ |FO|}{\sigma_o^2 + \sigma_c^2} \ . \qquad (3.3.59)$$

By equating this expression with the posterior mean μ_1 of equation (3.3.43), we obtain

$$p = \tfrac{1}{2} + \tfrac{1}{2} \tan h \ (\frac{|FO| \ FC}{\sigma_o^2 + \sigma_c^2}) \qquad (3.3.60)$$

which we may use as our assessed value of the probability of the true structure factor being positive.

We have seen in this Section some of the ways in which Bayesian ideas may be applied to crystal structure analysis. In particular, they have enabled us to incorporate naturally into the analysis any prior beliefs that we may hold. The Bayesian approach has furthermore, in accordance with our initial ideals, provided a sensible framework on which to make inferences and base decisions, strong arguments indeed for the inclusion of such ideas in the field of structure determination.

3.4 An Introduction to Sequential Experimentation in Crystallography

Sequential experimentation is concerned with the modification of a data collection process as a result of analysing data already collected. Such analysis of the data during the actual collection process may enable us to determine whether further acquisition of data is likely to be cost-effective at any given stage; if so, we would also like to be able to use statistical techniques to select what the appropriate new data should be in the light of some specified criteria. We shall therefore be concerned with procedures consisting of the three components:

1)   An initial sampling scheme

2)   A sequential algorithm for addition of new data

3)   A stopping rule.

We shall first, however, give a brief historical review of the development of sequential experimentation.

The role of experimental design in statistics was firmly established by Fisher's pioneering work at Rothamsted Experimental Station in the 1920's and 1930's, which relied on the key insight [63] that:

> 'statistical analysis of data could be informative only if the data themselves were informative, and that informative data could best be assured by applying statistical ideas to the way in which the data were collected in the first place.'

His work in this field culminated in 1935 with the publication of what is widely regarded as the definitive book on experimental design [28].

The area of sequential experimentation lay surprisingly dormant, however, until the classic works [65], [66] by Wald some 10-15 year later, from which much of today's thinking originates. His most important discovery was the sequential probability ratio test (see [65]), which was designed to choose between two simple hypotheses. This still remains an important area of sequential analysis, and is one in which the theory is now highly developed.

Applications of sequential experimentation have been dominated by sequential medical trials: when comparing two treatments, ethical considerations often demand that a trial be stopped whenever there is clear evidence of the superiority of one of the treatments, thus making such trials particularly conducive to sequential analysis. Alternative theory and methods to those originally proposed by Wald were developed in this context by Armitage and Bross, and may be found in [1], together with further details of the methodology of sequential medical trials.

Other important areas in the development of sequential analysis are discussed in [70]; these include sampling inspection and recent work on sequential estimation and

sequential resource allocation. A review of sequential statistical procedures (pre-1975) is also provided in [34], concentrating on sequential estimation and sequential hypothesis testing.

Our specific crystallographic application of sequential experimentation procedures will be concerned with the acquisition and analysis of X-ray diffraction data. We shall wish to make the process efficient in relation to specified criteria such as available resources and the information actually required from structural studies. We will be concerned with such aspects as cost, accuracy, speed and resolution - which will presently be introduced - and their relative importance, which will vary as the ultimate objectives of our structure determination vary. Brief comments will now be made on each of these four areas.

The financial cost involved in any given structure determination may be split into 6 components, namely

(1) The cost of mounting the specimen under consideration

(2) The cost of turning on the appropriate data-acquisition machinery

(3) The cost involved with the 'energy' source

(4) The cost of recording intensities

(5) The cost of measuring intensities

(6)   The cost of computing.

All 6 components will vary with the hardware and software
used, though for any given combination (1) and (2) will be
constant as will the turning on cost involved in (3). The
other component of (3) will be a running cost. Although
we   are   concentrating   here   on   X-ray   diffraction,
alternative energy sources may be used.   For example,
neutron beams - much used in crystallographic studies for
the analysis of proteins - are much more expensive and in
short supply.   General sequential procedures reducing
numbers of measurements taken will assume magnified
importance in the neutron beam case in view of the
consequent increase in (3).   The running cost will
obviously depend upon the time taken for experimentation;
this in turn will depend upon the (random) number of
intensities to be recorded and measured. The time taken
for individual intensities may not be constant across
reflections, however, nor may the costs (4) or (5).   In
the case of direct counting diffractometer methods, for
example, two basic procedures are in common usage: we may
either count for a specified time or until a specified
number of counts has been attained.   Finally, the cost
involved in (6) will depend on the complexity of our
calculations, in particular the nature of our sequential
algorithm and relevant updating procedures involved, which
may   be   governed   by   our   original   objectives.   Some

appropriate such updating procedures are discussed in the next Section.

In any structure determination we will be concerned with providing atomic co-ordinate estimates. There will also be an associated covariance matrix which we may take as our measure of the accuracy involved. When we speak of maximising the accuracy we need to minimise this matrix in some sense. One measure of the size of a matrix is its determinant and we may try to minimise this; alternatively, for example, we may try to minimise its trace, which is computationally much easier to evaluate and represents the sum of all the individual variances. The context should make it clear in what sense we are trying to maximise the accuracy in any given case.

The speed of any given procedure will be the time taken to reach the appropriate stopping rule. We shall generally assume that the overall time taken for any given sequential procedure is unlikely to be restrictively long, even in cases in which we are primarily interested in obtaining a quick estimate.

The concept of resolution shall now be introduced. We recall from Chapter 2 the essential premise of crystal structure analysis that atoms are located at the peaks of electron density maps, which may be constructed from observed structure amplitudes and calculated phases. Such maps may be calculated after eliminating all observed

amplitudes measured beyond a given 2θ value: the 'resolution' obtained is usually expressed in terms of the interplanar spacings

$$d = \frac{\lambda}{2\sin\theta}$$

corresponding to the maximum observed 2θ values. This definition emphasises the crystallographer's pre-occupation with collecting data in the order of ascending $(\sin\theta)/\lambda$. We see that the better the resolution, the more detail will be portrayed by the corresponding electron density maps (thus giving rise to better atomic co-ordinate estimates). It would seem reasonable to try to extend the definition of resolution to be a measure of the detail portrayed by successive density maps, in the general case where data has been acquired sequentially.

A major practical consideration will be the testing of any theoretical procedures that we may propose. In particular relevant sequential experimentation procedures may be tried for various combinations of the available data-acquisition hardware and software (though certain combinations may have been deemed unsuitable by theoretical considerations). It may well prove necessary to modify the theoretical results to develop procedures which perform well in practice. The need for co-operative interaction between statisticians and crystallographers is apparent. Where we find it expedient to resort to rules

117

of thumb or ad hoc methods, we shall still wish to be able to obtain some form of theoretical justification.

We consider now how we may hope to improve on existing procedures. The structural information available in crystallographic data bases has traditionally only been used to assist constrained refinements after the data has been collected; little attention has previously been paid to the possibility of analysing the data during the collection process in line with the ideals of sequential experimentation. Crystallographers have continued to take excessive measurements as though completely unaware that statisticians not only have extremely powerful techniques for estimating the value of information extracted from incomplete data sets, but also have appropriate stopping rules associated with their sampling procedures. Such an approach has been particularly remiss in the field of computer-controlled diffractometry, in which the data are acquired serially (in order of ascending $(\sin \theta)/\lambda$). The operational speed of modern computers means that the time taken to re-position the crystal and measure a new reflection - of the order of a minute - is long enough for a substantial amount of calculation to be undertaken before the crystal has to be reset and another measurement taken.

The possibility of allowing the data collection procedure to be modified according to the data already collected is

particularly attractive in the case of crystal structure solution and refinement for the following reason. The resolution attainable is linearly dependent on

$$S = \frac{\sin \theta}{\lambda} ,$$

whereas the number of available reflections is proportional to S^3; consequently, for even a modest increase in resolution, many extra reflections may be needed. To gain an acceptable degree of accuracy or resolution many more observations (typically 3-10 times) than variable parameters to be estimated are often needed. It is hoped that the sequential procedures envisaged in Section 3.6 will diminish this degree of over-determination.

We conclude this Section by noting the following three components of crystal structure analysis using diffractometer data:

1)   Models which are themselves derived by computer techniques involving ab initio calculations and/or data retrieved from computerised data bases

2)   Computer control of the data collection process, which is serial, and in which each additional observation can be used to update the prior information on a realistic real-time basis

3) Structure refinement using iterative computer procedures.

These components make the crystallographic setting ideal for the general exploration of techniques for computer-controlled sequential experimentation.

3.5 <u>Some Relevant Updating Procedures for Use in Sequential Structure Determination</u>

A sequential crystal structure analysis may be concerned with 3 updating procedures:

(1) Recursive addition of data

(2) Iterations to avoid matrix inversion

(3) Iterations to account for non-linearity.

We shall address (1) first. Following the least squares approach of Section 2.2 we shall assume we have available a least squares estimate after n_1 observations have been made, and formulate our problem as follows. We wish to find a sequential method for computing the least squares estimate after a further n_2 observations have been made, without having to re-do all our calculations from scratch. This will be particularly desirable in the sense of computer efficiency and real time operation. The recursive least squares (RLS) algorithm helps us achieve this and is given below. It provides exact results for

120

the true linear model, but can be incorporated into our non-linear theory as suggested.

Suppose we have a linear model of the form

$$\underline{Y} = \underline{X} \, \Delta\underline{\theta} + \underline{\epsilon} \qquad\qquad (3.5.1)$$

where

$$\underline{\epsilon} \sim N \, (\underline{O}, \underline{V}).$$

Suppose also we have a target function $\underline{\epsilon}^T \, \underline{W} \, \underline{\epsilon}$ (where $\underline{W} = \underline{V}^{-1}$) to be minimised after each batch of measurements has been received, giving updated estimates (for $(\Delta\underline{\theta})$) at each stage. Summarise our first n_1 observations by (3.5.1) with

$$\underline{X} = \underline{X}_1, \ \underline{Y} = \underline{Y}_1, \ \underline{V} = \underline{V}_1 \qquad\qquad (3.5.2)$$

where $\underline{X}_1$ is the design matrix for the first batch (n_1 observations) of data etc. Similarly, our first (n_1+n_2) observations may be summarised by (3.5.1) with

$$\underline{X} = \begin{bmatrix} \underline{X}_1 \\ \hline \underline{X}_2 \end{bmatrix}, \ \underline{Y} = \begin{bmatrix} \underline{Y}_1 \\ \underline{Y}_2 \end{bmatrix}, \ \underline{V} = \begin{bmatrix} \underline{V}_1 & \underline{V}_{12} \\ \hline \underline{V}_{12}^T & \underline{V}_2 \end{bmatrix} \qquad (3.5.3)$$

where $\underline{X}_2$ is the design matrix for the second batch (n_2 observations) of data etc. We shall here assume independence between batches, so that the covariance term $\underline{V}_{12} = 0$.

Let

$$\Delta\hat{\underline{\theta}}_1 = [\underline{X}_1^T \ \underline{W}_1 \ \underline{X}_1]^{-1} \ \underline{X}_1^T \ \underline{W}_1 \ \underline{Y}_1 \tag{3.5.4}$$

(where $\underline{W}_1 = \underline{V}_1^{-1}$) be the least squares estimator of $\Delta\underline{\theta}$ based on the first n_1 observations, and $\Delta\hat{\underline{\theta}}_2$ the least squares estimator to be calculated from the first (n_1+n_2) observations; similarly,

$$\underline{A}_1 = [\underline{X}_1^T \ \underline{W}_1 \ \underline{X}_1]^{-1} \tag{3.5.5}$$

and $\underline{A}_2$ the respective covariance matrices. Then

$$\Delta\hat{\underline{\theta}}_2 = \Delta\hat{\underline{\theta}}_1 + \underline{A}_1 \ \underline{X}_2^T \ [\underline{X}_2 \ \underline{A}_1 \ \underline{X}_2^T + \underline{V}_2]^{-1} \ (\underline{Y}_2 - \underline{X}_2 \ \Delta\hat{\underline{\theta}}_1)$$

$$\tag{3.5.6}$$

and

$$\underline{A}_2 = \underline{A}_1 - \underline{A}_1\underline{X}_2^T \ [\underline{X}_2\underline{A}_1 \ \underline{X}_2^T + \underline{V}_2]^{-1} \ \underline{X}_2 \ \underline{A}_1 \ . \tag{3.5.7}$$

Equations (3.5.6) and (3.5.7) form the basis of the recursive least squares algorithm; clearly, if the sizes of incoming batches of data (n_2) are small compared to the amount of data already held (n_1), then much less computational effort (in terms of matrix inversion etc) is required using the RLS algorithm than in reprocessing the whole data set at each stage.

The above analysis assumed a linear model. For problems

that are non-linear in the parameters, linearisation via a Taylor series expansion is used in order to reach the form required by the algorithm. In order to produce a sequence of least squares estimates a Gauss-Newton iteration should be undertaken at each stage. Figure (2.2.1) may be correspondingly modified to give rise to the sequential procedure for our non-linear model shown in Figure (3.5.1).

The procedure may be summarised as follows. For each new data set, the whole model is linearised about the current parameter estimate. The RLS formulae are then applied to get a parameter increment, and consequently updated parameter estimate. A Gauss-Newton iteration scheme may now proceed on the whole model linearised about this new parameter estimate and continue as before until convergence. A new data set may be received and the whole process repeated.

Clearly, since the matrices for the whole set of model equations must be inverted at each cycle of the Gauss-Newton iterations, the computational advantages of using the RLS algorithm will be lost, unless we can dispense with any such cycles (and cycle along the dotted line of Figure (3.5.1)) This may indeed be possible in the latter stages of a crystal structure refinement in which we might reasonably expect linearisation to be good enough that one iteration at each stage will give essentially correct

```
INITIAL DATA SET
       |
       v
   NEW DATA SET  <- - - - - - - - - - - - - - - - -
       |                                          ^
       |                                          |
       v                                          |
   RLS ALGORITHM                                  |
       |                                          |
       |              No                     No   |
       v                                          |
?fixed cycle iteration(s)? - - -> ?stopping rule? - - ->
       |                                     |
       | YES                     YES         | YES
       v                                     v
   ?stopping rule?  ----------------->     STOP
       |
       | NO
       v
```
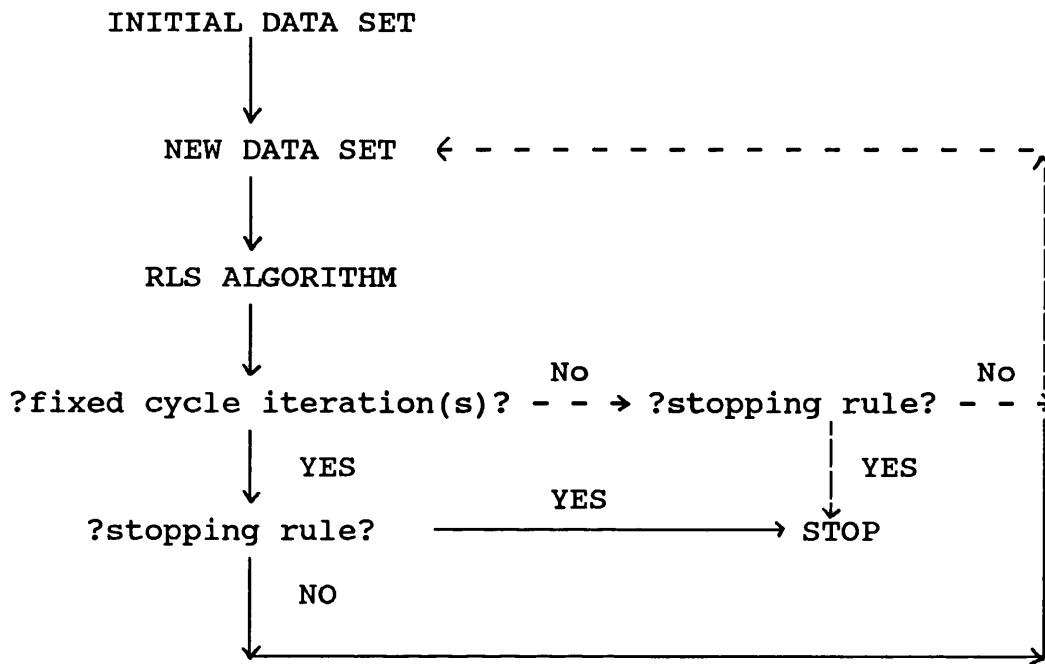
Figure 3.5.1: Non-linear sequential procedure based on
              RLS algorithm

estimates. Then we would hope that estimates obtained by
ignoring Gauss-Newton cycles would tend towards those
gained by following the complete procedure in the
following sense. The initial data set would yield a least
squares estimate sufficiently close to the true value that
further adjustments will be very small and each cycle of
our RLS algorithm would be roughly equivalent to one
Gauss-Newton cycle, especially as the amount of new
information received in each batch becomes (relatively)
increasingly small.

In terms of computer efficiency, avoidance/simplification
of matrix inversion routines (2) is the single most
important point.   A typical result is the matrix

approximation for small $\underset{\sim}{\varepsilon}$ relative to $\underset{\sim}{A}$ –

$$[\underset{\sim}{A} + \underset{\sim}{\varepsilon}]^{-1} \simeq \underset{\sim}{A}^{-1} - \underset{\sim}{A}^{-1} \underset{\sim}{\varepsilon} \underset{\sim}{A}^{-1} \tag{3.5.8}$$

– which may be used to update the covariance matrix for our parameter estimates when a (relatively small) new set of data becomes available. A similar result may also be used in conjunction with point (3) above, in which we find an approximation for the covariance matrix ($\underset{\sim}{A}_2^{-1}$) evaluated at an updated parameter estimate in terms of the original covariance matrix ($\underset{\sim}{A}_1^{-1}$). We assume $\underset{\sim}{A}_2$ is readily available and estimate its inverse via $\underset{\sim}{B}_2$ where

$$\underset{\sim}{B}_2 = 2\underset{\sim}{A}_1^{-1} \underset{\sim}{A}_1^{-1} \underset{\sim}{A}_2 \underset{\sim}{A}_1^{-1} . \tag{3.5.9}$$

Generally, given an approximation $\underset{\sim}{B}$ to $\underset{\sim}{A}^{-1}$ we could use (3.5.9) in the form

$$\underset{\sim}{B} \rightarrow 2\underset{\sim}{B} - \underset{\sim}{B} \underset{\sim}{A} \underset{\sim}{B}$$

to give us successive approximations to $\underset{\sim}{A}^{-1}$. Convergence of such methods obviously depends on the accuracy of our initial approximation; this may be quantified as in the related discussion below.

Returning to our observational equations (2.2.4) and putting

$$\underset{\sim}{A}^{(k)} = \underset{\sim}{X}^{(k)^T} \underset{\sim}{W} \underset{\sim}{X}^{(k)} \tag{3.5.10}$$

and

$$\underline{b}^{(k)} = \underline{X}^{(k)^T} \underline{W} \underline{Y}^{(k)} \qquad (3.5.11)$$

we see that

$$\underline{A}^{(k)} \Delta\underline{\theta}^{(k)} = \underline{b}^{(k)} . \qquad (3.5.12)$$

Letting $\underline{B}^{(k)}$ be an approximation to $[\underline{A}^{(k)}]^{-1}$, (3.5.12) becomes, approximately,

$$\Delta\underline{\theta}^{(k)} = \underline{B}^{(k)} \underline{b}^{(k)} . \qquad (3.5.13)$$

The (k+1)th iterate of $\underline{b}$ derived by using fixed data and an exactly linear model in which $\underline{A}^{(k)} = \underline{A}$ (say) for all k is [61]

$$\underline{b}^{(k+1)} = \underline{b}^{(k)} - \underline{A} \Delta\underline{\theta}^{(k)} . \qquad (3.5.14)$$

With (for all k) $\underline{B}^{(k)} = \underline{B} \simeq \underline{A}^{-1}$, equations (3.5.13) and (3.5.14) give

$$\Delta\underline{\theta}^{(k+1)} = [\underline{I} - \underline{B} \underline{A}]^k \underline{B} \underline{b}^{(1)} \qquad (3.5.15)$$

where $\underline{b}^{(1)}$ is the initial vector $\underline{b}$. For convergence, we require

$$\sum_{k=1}^{\infty} \Delta\theta_i^{(k)} < \infty \qquad (i=1...p) . \qquad (3.5.16)$$

Let the vector $\underset{\sim}{B}\ \underline{b}^{(1)}$ be expanded in terms of the eigenvectors $\underset{\sim}{\psi}_j$ (say) of $\underset{\sim}{B}\ \underset{\sim}{A}$; thus if

$$\underset{\sim}{B}\ \underline{b}^{(1)} = \sum_{j=1}^{p} c_j\ \underset{\sim}{\psi}_j \qquad\qquad (3.5.17)$$

equation (3.5.15) yields

$$\underline{\Delta\theta}^{(k+1)} = \sum_{j=1}^{p} (1-\lambda_j)^k\ c_j\ \underset{\sim}{\psi}_j \qquad\qquad (3.5.18)$$

where λ_j is the eigenvalue corresponding to $\underset{\sim}{\psi}_j$. Hence condition (3.5.16) becomes

$$\sum_{k=1}^{\infty}\sum_{j=1}^{p} (1-\lambda_j)^{k-1}\ c_j\ \psi_{ji} = \sum_{j=1}^{p} [\sum_{k=1}^{\infty} (1-\lambda_j)^{k-1}]\ c_j\psi_{ji} < \infty$$

$$(i=1\ldots p) \qquad\qquad (3.5.19)$$

which is satisfied if $|1-\lambda_j| < 1$ ($j=1\ldots p$). This condition is therefore sufficient to ensure convergence.

Ideally, combination of some of the procedures above into one iterative least squares cycle would be desirable e.g. something along the lines of the following simplified possibility (assuming initial estimate $\underline{\theta}^{(1)}$ available):

(1)   Set k=1; $\underline{\Delta\theta}^{(0)} = \underline{0}$

127

(2) Obtain initial data set

(3) Evaluate deign matrix ($\underset{\sim}{X}$), vector of observed values
     - current fitted values ($\underline{Y}$), weight matrix ($\underset{\sim}{W}$), based
     on <u>all</u> available data and evaluated at $\underline{\theta}^{(k)}$

(4) Set $\underset{\sim}{A} = \underset{\sim}{X}^T \underset{\sim}{W} \underset{\sim}{X}$

(5) Set $\underset{\sim}{B} = 2\underset{\sim}{B} - \underset{\sim}{B} \underset{\sim}{A} \underset{\sim}{B}$ (on first cycle take $\underset{\sim}{B} \simeq \underset{\sim}{A}^{-1}$)

(6) Set $\underline{b} = \underset{\sim}{X}^T \underset{\sim}{W} \underline{Y}$

(7) Evaluate $\underline{\Delta\theta}^{(k)} = \underset{\sim}{B} \underline{b}$ (EXIT when sufficiently
     small)

(8) Evaluate $\underline{\theta}^{(k+1)} = \underline{\theta}^{(k)} + \underline{\Delta\theta}^{(k)}$

(9) Set $k = k+1$

(10) (a) Return to (3) (for fixed cycle iteration)
     (b) Obtain new data set (small compared to current
          available data)

(11) Evaluate design matrix ($\underset{\sim}{X}$), weight matrix ($\underset{\sim}{W}$), based
     on new data set only and evaluated at $\underline{\theta}^{(k)}$

(12) Set $\underset{\sim}{A}* = \underset{\sim}{X}^T \underset{\sim}{W} \underset{\sim}{X}$

(13) Set $\underset{\sim}{B} = \underset{\sim}{B} - \underset{\sim}{B} \underset{\sim}{A}* \underset{\sim}{B}$

(14) Return to (3).

In practice, however, many crystallographic investigations
resort to approximating the matrices $\underset{\sim}{A}^{(k)}$ of equation

128

(3.5.10), thus dispensing with the need for full-matrix calculations. Approximations are sought for which the inverse is readily available; two of the classical candidates are given below.

The diagonal approximation sets all off-diagonal elements of $\underset{\sim}{A}^{(k)}$ equal to zero. A measure of acceptability of such an approximation could be obtained by e.g. evaluating a test statistic of the form

$$\sum_{j<i} A_{ji}^{(k)2}$$

though no such check is currently undertaken. This approach implies independence between all the individual parameters of our model; a slightly more realistic approximation is therefore the block-diagonal matrix in which all off-diagonal elements of $\underset{\sim}{A}^{(k)}$ are set equal to zero with the following exceptions:

(1)   intra-atomic terms between position parameters

(2)   intra-atomic terms between temperature parameters.

Such approximations generally led to overestimates for the parameter increments in their original applications (e.g. [20]); divergence was prevented by the introduction of empirical partial shift rules of the form

$$\underset{\sim}{\theta}^{(k+1)} = \underset{\sim}{\theta}^{(k)} + \eta \, \underset{\sim}{\Delta\theta}^{(k)} \qquad\qquad (3.5.20)$$

for some constant η less than one. The applicability of the above approximations will often be governed by the underlying objectives of our crystal structure analysis; in particular, such matrix simplification is most likely to be justified when we are primarily interested in obtaining quick estimates/packing structures.

Finally, with regard to iterative schemes to account for non-linearity, we may resort to cycling procedures such as that suggested by equation (3.5.9). Alternatively, appeal may possibly be made to some form of Taylor series analogue for matrices to help us evaluate/estimate $\underset{\sim}{x}^{(k+1)}$ in terms of $\underset{\sim}{x}^{(k)}$, though this is very much an open area.

3.6 Various Selection Schemes for use in Sequential Crystal Structure Determination

In accordance with our previous ideas, we shall be concerned with determining the structure of a centrally symmetric crystal. We shall assume throughout that our parameters, the atomic co-ordinate positions (x_1, y_1, z_1) .... (x_R, y_R, z_R), are of equal interest to us: the minimax strategy we shall be using will therefore be that of reducing all the corresponding variances below a specified target value. The theory may readily be adapted in the case of unequal interest. It may be further modified were we to wish to work in real space - rather than with co-ordinates in the unit cell - by noting that the co-

130

ordinates in real space may be obtained from those in the unit cell by pre-multiplication by a fixed matrix (determined by the unit cell).

Our theory concentrates on the final stages of a crystal structure refinement. We assume that we have already measured the intensities of reflections with indices $\underline{h} =$ hkl lying in a set S(1) (say) which has enabled us to obtain very good approximations $(x^{(r)}, y^{(r)}, z^{(r)})$ (r=1...R) to the exact atomic co-ordinates. As usual, we suppose that our determination of $(x^{(r)}, y^{(r)}, z^{(r)})$ (r=1...R) was carried out by choosing those parameters to minimise

$$\sum_{\underline{h} \in S(1)} W_{\underline{h}} \; (FC(\underline{h}) - FO(\underline{h}))^2$$

where

(1) FO($\underline{h}$) is an appropriately signed square root of the observed intensity $|FO(\underline{h})|^2$

(2) $$FC(\underline{h}) = 2 \sum_{r=1}^{R} f_r(\underline{h}) \cos 2\pi(hx_r + ky_r + lz_r)$$

where the $f_r(\underline{h})$ are temperature factors which may incorporate unknown thermal parameters

(3) $W_{\underline{h}}$ (weighting factor) is typically $1/\sigma_{\underline{h}}^2$ where $\sigma_{\underline{h}}^2$ is the best available estimate for the variance of $|FO(\underline{h})|$.

131

Our studies were then motivated by wishing to make final refinements by measuring some but not all of the intensities for $\underline{h}$ in the set T(say) of remaining indices for which intensity measurements are available to us. We seek a subset S(2) of T that is not unduly large but still enables us to make a very good final refinement.

Once S(2) has been determined, our final atomic co-ordinate estimates $(\xi^{(r)}, \eta^{(r)}, \varsigma^{(r)})$ (say) are obtained by minimising

$$\sum_{\underline{h} \in S(1) \cup S(2)} W_{\underline{h}} \, (FC(\underline{h}) - FO(\underline{h}))^2 \quad .$$

The estimates are of course unknown until we have carried through the minimisation process. This minimisation process shall now be given. The least squares equations of Section 2.2. are derived in detail, though we shall here be using a notation specific to this Section.

We shall use t to stand for any of x,y,z; likewise r for any of $\xi, \eta, \varsigma,$ and m for any one of h,k,l. It will be clear from the context that when e.g. t stands for y, then r stands for η and m stands for k. The co-ordinates $(\xi^{(r)}, \eta^{(r)}, \varsigma^{(r)}$ are determined as the solutions of the least squares equations

$$0 = \frac{\partial}{\partial r^{(r)}} \sum_{\underline{h} \in S(1) \cup S(2)} W_{\underline{h}} \, (FC(r) - FO(\underline{h}))^2$$

$$(r = \xi, \eta, \varsigma; \; r = 1 \ldots R) \qquad (3.6.1)$$

where FC(r) indicates that the calculations are based on the $(\xi^{(r)}, \eta^{(r)}, \varsigma^{(r)})$ co-ordinates. The exact (to within errors in the numerical methods) solutions of (3.6.1) are found by iteration of linearised versions of (3.6.1). By writing

$$\xi^{(r)} = x^{(r)} + \Delta x^{(r)}, \quad \eta^{(r)} = y^{(r)} + \Delta y^{(r)},$$

$$\varsigma^{(r)} = z^{(r)} + \Delta z^{(r)} \quad (r=1...R) \tag{3.6.2}$$

we obtain

$$FC(r) = 2 \sum_{r=1}^{R} f^{(r)}(\underline{h}) \cos 2\pi(h\xi^{(r)} + k\eta^{(r)} + l\varsigma^{(r)})$$

$$\approx 2 \sum_{r=1}^{R} f^{(r)}(\underline{h}) \cos 2\pi(hx^{(r)} + ky^{(r)} + lz^{(r)})$$

$$- 4\pi \sum_{r=1}^{R} f^{(r)}(\underline{h})[\sin 2\pi(hx^{(r)} + ky^{(r)} + lz^{(r)})]$$

$$\cdot \; (h\Delta x^{(r)} + k\Delta y^{(r)} + l\Delta z^{(r)}) \; , \tag{3.6.3}$$

where $f^{(r)}(\underline{h})$ represents theoretical temperature factors, which we shall assume fixed for our present purposes. Equation (3.6.3) may be re-written in the form

$$FC(r) \approx FC(t) - 2\pi \sum_{r=1}^{R} G_{(t)}^{(r)} (h\Delta x^{(r)} + k\Delta y^{(r)} + l\Delta z^{(r)})$$

$$\tag{3.6.4}$$

where FC(t) indicates that the calculations are based on

133

the $(x^{(r)}, y^{(r)}, z^{(r)})$ co-ordinates, and

$$G_{(t)}^{(r)} = G_{(t)}^{(r)}(\underline{h}) = 2f^{(r)}(\underline{h}) \sin 2\pi(hx^{(r)} + ky^{(r)} + lz^{(r)}),$$

$$(r=1\ldots R) \quad . \tag{3.6.5}$$

The least squares equations obtained by setting the differentials of

$$\sum_{\underline{h}\epsilon S(1)\cup S(2)} W_{\underline{h}} \; (FC(t) - FO(\underline{h}) - 2\pi \sum_{r=1}^{R} G_{(t)}^{(r)}$$

$$. \; (h\Delta x^{(r)} + k\Delta y^{(r)} + l\Delta z^{(r)}))^2$$

with respect to $\Delta t^{(r)}$ $(t=x,y,z;\; r=1\ldots R)$ equal to zero reduce to

$$2\pi \sum_{s=1}^{R} [\sum_{\underline{h}\epsilon S(1)\cup S(2)} W_{\underline{h}} \; m \; h \; G_{(t)}^{(s)} \; G_{(t)}^{(r)}] \; \Delta x^{(s)}$$

$$+ 2\pi \sum_{s=1}^{R} [\sum_{\underline{h}\epsilon S(1)\cup S(2)} W_{\underline{h}} \; m \; k \; G_{(t)}^{(s)} \; G_{(t)}^{(r)}] \; \Delta y^{(s)}$$

$$+ 2\pi \sum_{s=1}^{R} [\sum_{\underline{h}\epsilon S(1)\cup S(2)} W_{\underline{h}} \; m \; l \; G_{(t)}^{(s)} \; G_{(t)}^{(r)}] \; \Delta z^{(s)} \tag{3.6.6}$$

$$= \sum_{\underline{h}\epsilon S(1)\cup S(2)} W_{\underline{h}} \; (FC(t) - FO(\underline{h})) \; m \; G_{(t)}^{(r)}$$

$$(m=h,k,l;\; r=1\ldots R) \quad .$$

In matrix form the least squares equations above may be written (dropping tildes for notational convenience)

$$[A_1 + A_2] \; \Delta\theta = b_1 \; \Delta F_1 + b_2 \; \Delta F_2 \tag{3.6.7}$$

134

where

(1) A_i (i=1,2) contains elements of the form

$$2\pi \sum_{\underline{h}\epsilon S(i)} W_{\underline{h}} \; m \; m' \; G^{(s)}_{(t)} \; G^{(r)}_{(t)}$$

$$(m,m' =h,k,l; \; r,s=1...R)$$

(2) $\Delta\theta$ contains elements of the form $\Delta t^{(s)}$ (t=x,y,z; s=1...R)

(3) b_i (i=1,2) contains elements of the form

$$W_{\underline{h}} \; m \; G^{(r)}_{(t)} \qquad (\underline{h}\epsilon S(i); \; m=h,k,l; \; r=1...R)$$

(4) ΔF_i (i=1,2) contains elements of the form

$$FC(t) - FO(\underline{h}) \qquad (\underline{h}\epsilon S(i)) \quad .$$

Suppose now that the contributions from S(2) are small in comparison to those from S(1). This will indeed be the case in the latter stages of a sequential crystal structure refinement in which batches of incoming data will be small in comparison to the data already acquired. We may then use the matrix approximation (3.5.8) –

$$[A_1 + A_2]^{-1} \approx A_1^{-1} - A_1^{-1} A_2 A_1^{-1}$$

– to obtain the following result:

$$\Delta\theta \approx (P - Q) \Delta F_1 + R \Delta F_2 \qquad\qquad (3.6.8)$$

135

where

$$P = A_1^{-1} b_1, \quad Q = A_1^{-1} A_2 A_1^{-1} b_1, \quad R = A_1^{-1} b_2 . \qquad (3.6.9)$$

From equation (3.6.8), we can estimate the variance of $\tau^{(r)}$ by

$$\sum_{\underline{h} \in S(1)} (p_{\tau(r)}(\underline{h}) - q_{\tau(r)}(\underline{h}))^2 \sigma_{\underline{h}}^2 + \sum_{\underline{h} \in S(2)} (r_{\tau(r)}(\underline{h}))^2 \sigma_{\underline{h}}^2$$

$$(3.6.10)$$

where $p_{\tau(r)}(\underline{h})$ represents elements of the $\tau^{(r)}$ - row of P

etc. Further reduction of (3.6.10) based on the assumption of small $S(2)$ terms in comparison to $S(1)$ terms produces the variance estimate

$$\sigma_{\tau(r)}^2 = \sum_{\underline{h} \in S(1)} (p_{\tau(r)}(\underline{h}))^2 \sigma_{\underline{h}}^2 - 2 \sum_{\underline{h} \in S(1)} p_{\tau(r)}(\underline{h})$$

$$\cdot q_{\tau(r)}(\underline{h}) \sigma_{\underline{h}}^2 . \qquad (3.6.11)$$

Note that the first term on the right hand side of equation (3.6.11) is in fact the variance estimate based on the $S(1)$ measurements alone. Our minimax strategy is to introduce indices $\underline{h} = hkl$ into $S(2)$ in order to maximise the reduction of the single biggest $\sigma_{\tau(r)}^2$ i.e. to maximise (for a particular $\tau^{(r)}$)

$$\sum_{\underline{h} \in S(1)} p_{\tau(r)}(\underline{h}) \, q_{\tau(r)}(\underline{h}) \sigma_{\underline{h}}^2 ,$$

the S(2) terms entering the above formula via the $q_{\tau}(r)(\underline{h})$

terms. Selection of the appropriate indices is greatly

simplified in practice by the following matrix

manipulations. Note that we henceforth assume that our

refinement procedure is sufficiently far advanced for us

to use a weighting scheme proportional to observational

errors alone, in line with the original least squares

theory of Section 2.2.

The matrix A_2 may be written as

$$A_2 = \sum_{\underline{h} \in S(2)} g(\underline{h}) \, g(\underline{h})^T \qquad (3.6.12)$$

where the column vector $g(\underline{h})$ contains elements of the form

$$(2\pi \, W_{\underline{h}})^{\frac{1}{2}} m \, G^{(r)}_{(t)} \qquad (m = h,k,l; \ r=1\ldots R) \; . \qquad (3.6.13)$$

Hence $Q \ (=A_1^{-1} A_2 A_1^{-1} \, b_1)$ may be written

$$Q = \sum_{\underline{h} \in S(2)} A_1^{-1} g(\underline{h}) \, g(\underline{h})^T A_1^{-1} b_1 \; . \qquad (3.6.14)$$

We further define $e_{\tau}(r)$ to be the column vector consisting

of zeros except for a 1 in the $\tau^{(r)}$ - row, and $D(\sigma^2)$ to be

the diagonal matrix with elements $\sigma_{\underline{h}}^2$ ($\underline{h} \in S(1)$).

Then

$$\sum_{\underline{h} \in S(1)} p_{\tau}(r)(\underline{h}) \, q_{\tau}(r)(\underline{h}) \, \sigma_{\underline{h}}^2 = (e_{\tau}(r)^T P) \, D(\sigma^2) \, (e_{\tau}(r)^T Q)^T$$

$$= \sum_{\underline{h} \in S(2)} \left((e_{\tau(r)}{}^T A_1{}^{-1}) b_1 \right) D(\sigma^2) (e_{\tau(r)}{}^T A_1{}^{-1} g(\underline{h}) g(\underline{h})^T A_1{}^{-1} b_1)^T$$

$$= \sum_{\underline{h} \in S(2)} e_{\tau(r)}{}^T A_1{}^{-1} (b_1 D(\sigma^2) b_1{}^T) (e_{\tau(r)}{}^T A_1{}^{-1} g(\underline{h}) g(\underline{h})^T A_1{}^{-1})^T.$$

<div align="right">(3.6.15)</div>

By using the result (putting $w_{\underline{h}} = 1/\sigma_{\underline{h}}{}^2$)

$$b_1 D(\sigma^2) b_1{}^T = \sum_{\underline{h} \in S(1)} w_{\underline{h}} \, m \, G_{(t)}^{(r)} \, m' \, G_{(t)}^{(s)} = \frac{1}{2\pi} A_1$$

<div align="right">(3.6.16)</div>

and the fact that

$$e_{\tau(r)}{}^T A_1{}^{-1} g(\underline{h})$$

is a scalar, equation (3.6.15) yields

$$\sum_{\underline{h} \in S(1)} p_{\tau(r)}(\underline{h}) \, q_{\tau(r)}(\underline{h}) \, \sigma_{\underline{h}}{}^2 = \frac{1}{2\pi} \sum_{\underline{h} \in S(2)} (e_{\tau(r)}{}^T A_1{}^{-1} g(\underline{h}))^2.$$

<div align="right">(3.6.17)</div>

Our minimax strategy, having isolated the worst variance $\sigma_{\tau(r)}{}^2$ after calculations based on our $S(1)$ measurements, is therefore to choose $S(2)$ to consist of those $\underline{h}$ that maximise

$$(e_{\tau(r)}{}^T A_1{}^{-1} g(\underline{h}))^2 .$$

<div align="right">(3.6.18)</div>

A simplified version of the selection scheme advocated above may be obtained by replacing the least squares equations (3.6.6) by the diagonalised equations

$$2\pi \left(\sum_{\underline{h} \in S(1) \cup S(2)} W_{\underline{h}} \ (m \ G^{(r)}_{(t)})^2 \right) \Delta t^{(r)} = \sum_{\underline{h} \in S(1) \cup S(2)} W_{\underline{h}} (FC(t)$$

$$- FO(\underline{h})) \ m \ G^{(r)}_{(t)} \qquad (m=h,k,l; \ r=1...R). \qquad (3.6.19)$$

To this approximation, the variance of $r^{(r)}$ is proportional to

$$\sum_{\underline{h} \in S(1) \cup S(2)} \frac{1}{W_{\underline{h}}(m \ G^{(r)}_{(t)})^2} \ , \qquad (3.6.20)$$

whence we would be led to choose S(2) to consist of those $\underline{h}$ for which

$$W_{\underline{h}} \ (m \ G^{(r)}_{(t)})^2 \qquad (3.6.21)$$

is as large as possible. For the selection rules based on (3.6.18) and (3.6.21) we must use an estimate of the weight of the as yet unobserved reflections; such an estimate will be readily available based for example on either the current FC(t) values, or on a cursory inspection of a diffraction photograph. It may also be suggested by the very nature of certain direct counting diffractometer methods - see Section 3.4 - which may be employed.

139

We see that (as above) choosing new observations to help reduce the variances of individual parameters is easy enough. Suitable extensions of these ideas to deal with alternative strategies to our minimax procedure may also be derived in certain cases. For example, we may want to minimise the average $\tau^{(r)}$ - variance or, equivalently, minimise

$$\sum_{r=1}^{R} \quad \sum_{\tau=\xi,\eta,\varsigma} \quad \sigma^2_{\tau}(r) \quad . \tag{3.6.22}$$

We may approximate the individual variances of the form (3.6.20) using the first order binomial expansion

$$\frac{1}{c_{\tau}(r)} \left(1 - \frac{1}{c_{\tau}(r)} \sum_{\underline{h}\epsilon S(2)} W_{\underline{h}} \ (m \ G^{(r)}_{(t)})^2 \right) \tag{3.6.23}$$

where

$$c_{\tau}(r) = \sum_{\underline{h}\epsilon S(1)} W_{\underline{h}} \ (m \ G^{(r)}_{(t)})^2 \tag{3.6.24}$$

is a known constant. Then our selection rule based on minimising (3.6.22) becomes that of choosing $\underline{h}$ in order to maximise

$$W_{\underline{h}} \sum_{r=1}^{R} (G^{(r)}_{(t)})^2 \left[(\frac{h}{c_{\xi}(r)})^2 + (\frac{k}{c_{\eta}(r)})^2 + (\frac{1}{c_{\varsigma}(r)})^2 \right], \tag{3.6.25}$$

a slightly less tractable, though nevertheless still manageable, result. This latter strategy is analogous to seeking to maximise the accuracy of our refinement in the sense of minimising the trace of our covariance matrix as discussed in Section 3.4. Were we instead to wish to minimise the determinant, the natural way to proceed would be via the theory of D-optimality discussed below. As we shall see, the form of the selection rule derived generalises the minimax result based on (3.6.18).

We recall that in our current setting we are supposing the known response function analogue of our structure factor equation to be of the form

$$\eta(\underline{h}) \;=\; g(\underline{h})\, \underline{\Delta\theta} \quad,$$

with our observations $\underline{Y}$ satisfying

$$\underline{Y} \;=\; \underline{\eta} + \underline{\epsilon}$$

where the random errors $\underline{\epsilon}$ have the distribution

$$\underline{\epsilon} \sim N\,(\underline{0}\,,\,\underline{\underline{V}})\,.$$

For observations at $\underline{h} \in S(1)$, estimation of $\underline{\Delta\theta}$ by least squares gives the result — c.f. (3.6.7) —

$$\Delta\hat{\theta} \;=\; A_1^{-1}\, b_1\, \Delta F_1 \qquad\qquad (3.6.26)$$

141

with corresponding covariance matrix A_1^{-1}. It is the desire to minimise this covariance matrix A_1^{-1} – equivalently, maximise A_1 – that motivates the definition of D-optimality. Our initial sampling design is said to be D-optimal if it specifies a subset $S(1)$ of indices for our initial data set, for which the associated information matrix A_1 has maximum determinant over all possible such subsets of the same size. Thus, given our initial set $S(1)$, this criterion leads us to seek the (fixed-size) subset $S(2)$ of indices for which the determinant of $(A_1 + A_2)$ is maximised.

For sequential schemes bringing in one extra observation at each stage, the appropriate selection rule may be readily defined. We merely seek that $\underline{h}$ (ϵT) for which

$$\det (A_1 + g(\underline{h}) \ g(\underline{h})^T)$$

is maximised. By using the identity

$$\det(A_1 + g(\underline{h})g(\underline{h})^T) = (1 + g(\underline{h})^T A_1^{-1} g(\underline{h})) \ \det A_1$$

$$(3.6.27)$$

we see that our selection rule is therefore to choose that $\underline{h}$ that maximises

$$g(\underline{h})^T A_1^{-1} g(\underline{h}) \qquad\qquad (3.6.28)$$

(c.f. our previous minimax result based on (3.6.18)). For general fixed batch size (>1), maximising $\det(A_1 + A_2)$ over

all possible candidates A_2 is likely to prove computationally prohibitive.

We may however, use the appropriate number of indices based on the largest values of (3.6.28), or use a one-at-a-time selection scheme as follows. Having chosen the best single $\underline{h}$ based on (3.6.28) we may obtain

$$A_2 = A_1 + g(\underline{h}) \; g(\underline{h})^T.$$

We may then choose the best single $\underline{h}$ based on (3.6.28) with A_1^{-1} replaced by A_2^{-1} (or an approximation to it); this process may be continued until the requisite number of $\underline{h}$'s have been determined.

At this juncture it seems appropriate to make a few comments on stopping rules and average sample numbers i.e. we have derived rules to determine where to make extra observations, but little has so far been said about how many to take at any stage or, indeed, when to stop sampling altogether. Termination criteria are usually formulated in terms of the accuracy of the estimates e.g. we may demand that all individual parameter variances fall below specified levels or that the determinant of the associated covariance matrix falls below a specified value. The number n_0 (say) of additional observations required at any stage, given that we have already made n observations, may be estimated [3] using the fact that the elements of the present covariance matrix - c.f. equation

(2.4.4) - are roughly proportional to 1/n-p (where p = number of parameters to be estimated). For example, if the present determinant = a, and we wish to reduce this past b (<a), then we must solve the equation

$$(n + n_o - p)^p \, b = (n - p)^p \, a \qquad (3.6.29)$$

for n_0.

Returning to our previous ideas, we now show that the general theory of optimal design effectively enables us to handle nuisance parameters in the sense that it allows us to concentrate on certain subsets of parameters e.g. the co-ordinates of any individual atom that may be of particular interest in a given structure determination. To achieve this we partition our information matrix A based on the appropriate subset of parameters (this may first necessitate a re-ordering of our parameters) so that

$$A = \begin{bmatrix} A_{11} & A_{12} \\ A_{12}^T & A_{22} \end{bmatrix}$$

where A_{11} is based on derivatives only with respect to our parameters of interest. The inverse of the upper left hand component of the similarly partitioned matrix A^{-1} is given by

$$A_{11} - A_{12} \, A_{22}^{-1} \, A_{12}^T \qquad (3.6.30)$$

and our design criterion now becomes that of maximising this determinant. Similar techniques allow us to 'home in' on any specified linear combination of parameter values. Further details, together with much of the general underlying theory of optimal design may be found in e.g. [60].

The question of randomised sampling schemes will now be considered, including specification of appropriate initial sampling schemes. If a trial structure is available, selection rules may be based on a specific criterion such as D-optimality. Alternatively, since expected values may be calculated for all structure factors, we may wish to measure the (expected) strongest reflections: observed and calculated phases are most likely to be correct for these reflections, which also take less time to measure with a given precision than do weak reflections. If, however, there is no information as to the relative value of the data points in the specific experiment under consideration, then they can be selected at random [58], subject to general provisos such as adequate 3-d coverage. Although we shall usually expect to have a suitable trial structure available to us, success in such cases motivates us to consider selection schemes based on the ideas of stratified random sampling as discussed below.

Stratified random sampling consists of splitting a population into various strata, and then independently drawing simple random samples within each stratum. One common reason for stratified sampling is that if the units within the strata are more homogeneous than the population as a whole, then we might obtain better estimates. For example in the crystallographic setting we might wish to allocate various strata - i.e. indices at which observations are available - according to the assessed weights of the observations. Such an approach should be feasible in practice since for any given structure determination we are likely to have some insight into the appropriate weights prior to experimentation. In cases where equal weights are assumed, stratification may be carried out according to e.g. convenience and/or cost.

A standard result in the theory of stratified random sampling will now be given, and a parallel result postulated for the crystallographic analogue. Considering the situation in which we have L strata and we are interested in obtaining an estimate of the population mean. Let the variance within the ith stratum be σ_i^2; suppose also that stratum i has N_i units and the sampling cost per unit is c_i. Suppose further that we have an overhead cost c_0 so that the total cost associated with the allocation $n_1 \ldots n_l$ is

$$c_0 + \sum_{i=1}^{L} c_i\, n_i \quad .$$

Then the optimum allocation - i.e. the one that minimises the variance of the resultant estimate - associated with a fixed cost c is given by [17]

$$\sum_{i=1}^{L} n_i = n = (c - c_0)\; \frac{\sum\limits_{i=1}^{L} \left(\dfrac{N_i\, \sigma_i}{\sqrt{c_i}}\right)}{\sum\limits_{i=1}^{L} N_i\, \sigma_i\, \sqrt{c_i}} \qquad (3.6.31)$$

and

$$\frac{n_i}{n} = \frac{\left(\dfrac{N_i\, \sigma_i}{\sqrt{c_i}}\right)}{\sum\limits_{j=1}^{L}\left(\dfrac{N_j\, \sigma_j}{\sqrt{c_j}}\right)} \qquad (i=1...L) \quad . \qquad (3.6.32)$$

As a corollary, we see that in the special case where cost does not depend on stratum (i.e. $c_1 = \ldots = c_L$) the optimum (or 'Neyman') allocation for a fixed size n is given by

$$n_i = \frac{n\, N_i\, \sigma_i}{\sum\limits_{j=1}^{L} N_j\, \sigma_j} \qquad (i = 1 \ldots L). \qquad (3.6.33)$$

Finally, we see that proportional allocation, in which n_i/N_i does not depend on stratum i, is optimal if the sample size is fixed and the variances within the strata are equal.

147

In our crystallographic application, adaptations to the above theory have to be made. In particular, we note that we will be concerned with estimating a vector of parameters. We will now therefore wish to obtain a representative value from each stratum to indicate the relative importance of the stratum for structure determination. A natural candidate is a measure of the accuracy of any observation in the stratum; in such cases where we stratify via the (assessed) weight of observations, such a value is readily available. Furthermore, such a scheme is likely to be similar to stratification via $(\sin \theta)/\lambda$ values, and the assumption of constant cost across each stratum will not be unreasonable. In such cases, for a fixed sample size n, we will arrive at sampling schemes of the form of (3.6.32), where σ_i^2 is the reciprocal of the representative stratum weight.

In accordance with the above ideas, initial sampling schemes in structure determinations have usually been based on a low $(\sin \theta)/\lambda$ values: these reflections have the largest assessed weights and can consequently be measured most accurately, are also most easily attainable in terms of speed and/or convenience sampling, and cost least to measure. We continue to safely assume that appropriate initial sampling schemes are available and concentrate instead on sequential sampling schemes to use.

A final refined structure will need good agreement between theoretical and observed structure factors for all reflections, not just those with low $(\sin \theta)/\lambda$ values; in fact, agreement for high order reflections - i.e. those with high $(\sin \theta)/\lambda$ values - is of paramount importance as we shall discuss later in this Section. Such reflections will therefore need to be brought into our schemes. We may achieve this via stratification based on the concepts of trace and determinant of our information matrix in line with our previous minimax and D-optimality ideas, which are motivated by seeking to minimise our associated covariance matrix. Note that the relevant selection rules we derived based on (3.6.18) and (3.6.28) were biased towards choosing higher order reflections.

Suppose at any given stage of our structure refinement we have an information matrix A_1. Bringing in reflections at indices $\underline{h} \in S(2)$, our information matrix will become $(A_1 + A_2)$ where A_2 is of the form of equation (3.6.12). Since

$$\text{tr}(A_1 + A_2) = \text{tr } (A_1) + \sum_{\underline{h} \in S(2)} \text{tr } (g(\underline{h}) \, g(\underline{h})^T)$$

we see that we may stratify according to

$$\text{tr}(g(\underline{h}) \, g(\underline{h})^T). \qquad\qquad (3.6.34)$$

Likewise, in view of the identity (3.6.27), the

corresponding argument based on determinants leads us to stratify according to (3.6.28) i.e.

$$g(\underline{h}) \ A_1^{-1} \ g(\underline{h})^T.$$

It is now unlikely that the assumption of equal cost across strata will hold. However, for a given batch size n, we may still stratify as suggested above and use Neyman allocation (3.6.33) with σ_i replaced by an analogous representative stratum value based on (3.6.34) or (3.6.28). This will tell us how many observation to take in each stratum. In cases where equal costs do hold, or where cost is a minor consideration, we may sample randomly across the strata. If, however, costs vary randomly across strata, the sample may be chosen non-randomly corresponding to the lowest costs within the given stratum. Alternatively, considerations such as likelihood of correct sign allocation or ease of measurement may determine where we measure. For any such schemes, it is important to note that the strata will vary as our parameter estimate varies at each successive stage of our iteration procedure.

However we choose our sequential sampling schemes, the success of our statistical techniques, depending as they invariably do on linear approximations, will ultimately be influenced by the 'degree of non-linearity' present and the consequent goodness of fit of linear approximations

150

such as equation (3.6.3) over the appropriate ranges. For the individual geometrical (cosine) components of our structure factor equation, the required measure of goodness of linear approximation will depend on two things:

(1)   The value of the fitted cosine term

(2)   The range of approximation.

We merely comment here that, by the nature of the cosine curve, we would expect better approximations in cases where the individual geometrical components have small values of (1) (in modulus) and (2).   This should consequently yield faster convergence of our iteration procedures and may prove useful when seeking quick estimates/packing structures.   The requirement of small values of (2) justifies omitting most high order data in such cases.   We should be aware, however, that vital information may be missed by not including certain reflections; in particular, phase information will suffer since the (theoretical) structure factors included will now be smaller than usual, being the sums of terms which will (on average) be more concentrated round small |cosine| values. The question of overcoming this apparent conflict between good phase information and acceptable linear approximation may be addressed by seeking to modify our weighting schemes according to some non-linearity criterion (e.g. based on minimum discrepancy or mean

squared error over ranges of approximation). This is not pursued further here, but we note that help in this area might come from the pioneering paper by Beale [7], which sought to make more precise the notion of 'degree of non-linearity' present and presented various non-linearity measures to help quantify this, originally for use in the formulation of conservative confidence intervals for the parameters of interest. Bates and Watts [4] developed non-linearity measures of their own based on the geometric concept of curvature, which they showed to be related to Beale's original measures. A review of their work in the field of non-linear parameter estimation is undertaken in Section 4.3.

We consider next the specification of an appropriate procedure to adopt in practice. As mentioned in Section 1.4 this will be influenced by the underlying objectives of each individual structure determination; two general strategies emerge naturally, however, which are likely to cover most possibilities. The first is the minimax type of strategy as discussed earlier in this Section, in which we may seek to reduce all variances of our atomic co-ordinate estimates below a specified level. Associated variations such as our D-optimality based strategy are included in this category. The second approach is based on the premise that the true test is whether calculated structure factors fit those observed for high-order reflections. As we saw at the end of Section 3.2, shifts

in our atomic co-ordinate estimates had relatively larger effect for fitted high-order values. Thus we can only expect a good correspondence between such observed and calculated values when our atomic estimates are indeed very close to the true underlying values. Because of their importance, such high-order reflections (which are necessarily small due to the reduced atomic scattering factors involved - see Section 1.2) should be measured as accurately as possible. As a consequence of the different viewpoints, the main difference between the above schemes turns out to be the respective stopping rules. The former strategy has a well-defined stopping rule, whilst for the latter the question of quantifying what constitutes a good or bad fit is traditionally answered in terms of R-factors (see Section 4.2). We re-iterate that we are primarily concerned with the final stages of a crystal structure analysis in which we assume our data already have essentially correct phases, and in which minimax-type selection schemes have a bias towards choosing high-order data as we have seen; the two approaches will thus be similar in practice.

Suppose we were to carry out a sequential structure determination using our minimax criterion. We shall now consider what visual displays we may wish to use in conjunction with such a procedure. The most obvious visual aid is a graph showing how the maximum variance at each stage, σ^2_{max} (say), progresses towards the target

value σ^2_T (say); or equivalently, how $-\sigma^2_{max}$ progresses towards $-\sigma^2_T$ as in Figure 3.6.1. The horizontal scale should be '% accessible data' (the left hand point being determined by the size of the initial batch of data) and the vertical scale '$-\sigma^2_{max}$' as shown. Such graphs may be useful in indicating whether our target value is too stringent or not i.e. whether we are likely to reach our target value even with all accessible data. If thought not, we might wish to curtail our experimentation procedure or set a more realistic target value.
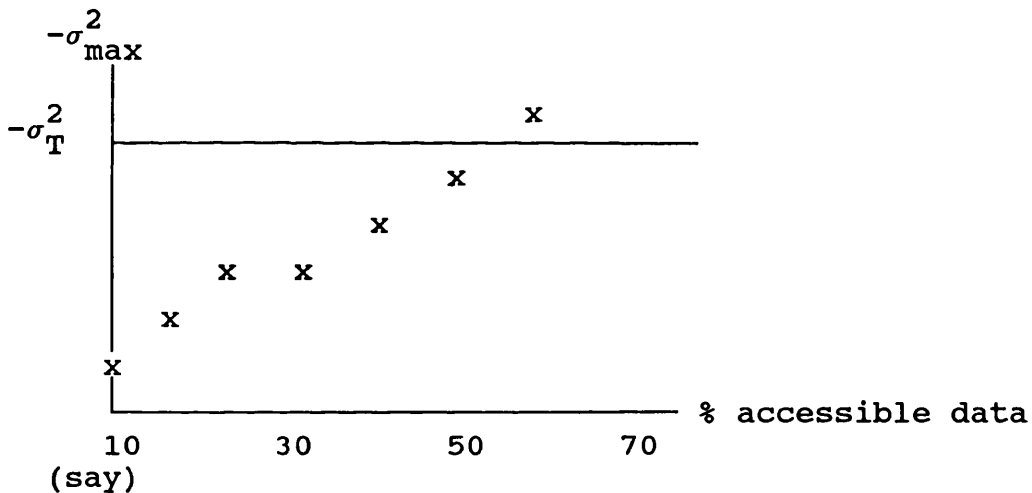
Figure 3.6.1: Chart to show progress of a sequential structure determination based on a minimax strategy

Similarly the R-values of Section 4.2 may be displayed at the various stages of our sequential procedure, either for all the data measured to date or for just the high-order reflections (or both). It is to be hoped that the initial

batches of data ensure sufficient phase discrimination that the refinement procedure will not start to diverge when higher-order reflections are included. This can be monitored by checking that successive R-values do not increase. A similar chart to Figure 3.6.1 can be produced for R-values and may be treated like a control chart in the theory of quality control [69] with associated stopping rules of its own. The two charts may of course be displayed simultaneously.

CHAPTER 4

NON-LINEAR PARAMETER ESTIMATION

4.1 Introduction

One of the basic difficulties involved in the refinement of crystal structures is the fact that the model is non-linear. Several proposals on how to measure the non-linearity of the consequent linearised models used have appeared in the literature (see e.g. [4], [7]). Hougaard [44] has used measures of this type to suggest ways of parameterising models to obtain various properties such as normal likelihood, stability of variance, zero asymptotic skewness and asymptotic unbiasedness. The adaptation and extension of related work into the area of crystal structure determination may lead to improved parameterisation of our existing model, and to alternatives to the current practice of adopting the first-order Taylor series approximation to any non-linear function. We shall investigate this further during this Chapter, relying heavily on the work [4], [5] of Bates and Watts.

First, however, a review of some standard existing methods used to tackle the problem of non-linear parameter estimation will be given in Section 4.2. The work of Bates and Watts will then be reviewed in Section 4.3; this introduces the important concepts of intrinsic curvature

and parameter effects curvature in an attempt to quantify the appropriateness of standard linear approximations and various parameterisations. These notions will be used throughout the remainder of the Chapter.

Section 4.4 considers possible re-parameterisations motivated by considering the relation between points on the approximating tangent plane and their analogues on the solution surface. In particular, mappings are sought that are orthogonal to the tangent plane in order to reduce the parameter effects curvature of Bates and Watts. Section 4.5 considers non-linear parameter estimation from a geometrical point of view and proposes various alternative estimation procedures based on curvature measures derived from the theory of Section 4.3. Section 4.6 seeks to extend the general theory and suggest how use of it may be made in the crystallographic setting. Section 4.7 briefly discusses the results of applications of some of our ideas.

4.2 A Review of Methods

This Section reviews alternative standard methods of non-linear parameter estimation. We consider first an approach based on maximum likelihood. Such an approach underlines the fundamental differences between Bayesian and classical statistics. Most notably, classical inference prescribes inferential techniques that are based

on sample information alone. The classical statistician further argues that our parameters are fixed, not random variables, and any probability statements about possible parameter values are therefore meaningless. The interpretation of results obtained under the two approaches highlights the different underlying rationales; the mode of the posterior distribution favoured by Bayesians gives the most likely parameter value whereas the maximum-likelihood estimator is that parameter value that makes the observed sample results appear most likely. Whereas Bayesians make probability statements about the parameters, classical statisticians make probability statements concerning the sample results.

The likelihood function $L(\underline{\theta})$ of a sample is the joint probability density function of the observations viewed as a function of the unknown parameters $\underline{\theta}$. The maximum likelihood estimate of $\underline{\theta}$ is that value which maximises $L(\underline{\theta})$, or equivalently log $L(\underline{\theta})$; this value is of course subject to any constraints that may be imposed. In the case where we assume our observational errors $\underline{\epsilon} = \underline{\epsilon}(\underline{\theta})$ satisfy

$$\underline{\epsilon} \sim N(\underline{0}, \underline{V})$$

the log-likelihood function for a sample of size n takes the form

$$\log L(\underset{\sim}{\theta}) = -\frac{n}{2} \log 2\pi - \tfrac{1}{2}\log \det \underset{\sim}{V} - \tfrac{1}{2} \sum_{i,i'=1}^{n} W_{ii'}\epsilon_i\epsilon_{i'}$$

$$(4.2.1)$$

where $W_{ii'}$ is the $(ii')^{th}$ entry of the matrix $\underset{\sim}{W} = \underset{\sim}{V}^{-1}$. When $\underset{\sim}{V}$ is known the maximum likelihood method requires us to minimise

$$\sum_{i,i'=1}^{n} W_{ii'}\ \epsilon_i\epsilon_{i'} \ . \qquad\qquad (4.2.2)$$

With $\underset{\sim}{\epsilon} = \underset{\sim}{Y}$ (vector of observed values - fitted values), (4.2.2) reduces to ψ' of (2.2.3) and our criterion is precisely that of weighted least squares. If, for some reason, our dispersion matrix V was unknown, however, (c.f weighted least squares with unknown weights), maximum likelihood based methods may be employed as follows. Only the results will be quoted here; for further details of their derivation see e.g. [3].

Define the cross-product matrix $\underset{\sim}{M}(\underset{\sim}{\theta})$ of the residuals by

$$\underset{\sim}{M}(\underset{\sim}{\theta}) = \underset{\sim}{\epsilon}(\underset{\sim}{\theta})\ \underset{\sim}{\epsilon}(\underset{\sim}{\theta})^T \ . \qquad\qquad (4.2.3)$$

In the case where $\underset{\sim}{V}$ is known to be a diagonal matrix, we seek those values of $\underset{\sim}{\theta}$ and V_{ii} (i=1..n) that maximise

$$\log L(\underline{\theta}) = -\frac{n}{2}\log 2\pi - \tfrac{1}{2}\sum_{i=1}^{n}\log V_{ii} - \tfrac{1}{2}\sum_{i=1}^{n}V_{ii}^{-1}M_{ii}(\underline{\theta}) \ .$$

$$(4.2.4)$$

We proceed by finding $\underline{\theta}*$ (say) which minimises

$$\sum_{i=1}^{n}\log M_{ii}(\underline{\theta}) \qquad\qquad (4.2.5)$$

and by estimating V_{ii} (i=1...n) via $\epsilon_i^2(\underline{\theta}*)$.

When $\underline{V}$ is a general unknown matrix our problem can be formulated in terms of seeking those values of $\underline{\theta}$ and $\underline{V}$ that maximise

$$\log L(\underline{\theta}) = -\frac{n}{2}\log 2\pi - \tfrac{1}{2}\log \det \underline{V} - \tfrac{1}{2}\mathrm{tr}[\underline{V}^{-1}\underline{M}(\underline{\theta})]. \quad (4.2.6)$$

We now proceed by finding $\underline{\theta}*$ which minimises

$$\log \det \underline{M}(\underline{\theta}) \qquad\qquad (4.2.7)$$

and by estimating V via $\underline{M}(\underline{\theta}*)$.

Although both estimates given for $\underline{V}$ are biased, methods exist [3] for possible bias removal. The above results are included merely for completeness and to provide a useful reference in cases where $\underline{V}$ is unknown; in crystallographic studies, however, we can generally expect to know $\underline{V}$ in which case the method of maximum likelihood is equivalent to that of weighted least squares as we have seen.

Suppose we wish to minimise the residual sum of squares function ψ of equation (2.2.1) with respect to $\underline{\theta}$ as before. Let ψ_k denote $\psi(\underline{\theta}^{(k)})$; then we call an iteration acceptable if $\psi_{k+1} < \psi_k$. Most iterative schemes encountered in non-linear parameter estimation problems such as we have here now fall into the following general framework:

(1) Provide an initial parameter estimate $\underline{\theta}^{(0)}$ (set k = 0)

(2) Determine a vector $\underline{v}^{(k)}$ in the direction of the proposed step

(3) Determine a scalar $\rho^{(k)}$ such that $\underline{\theta}^{(k+1)} = \underline{\theta}^{(k)} + \rho^{(k)} \underline{v}^{(k)}$ produces an acceptable iteration

(4) Set k=k+1 and return to step 2 unless a specified termination criterion has been met, in which case accept $\underline{\theta}^{(k)}$ as our estimate of $\underline{\theta}$.

Our standard least squares approach of Section 2.2 took $\underline{v}^{(k)} = \Delta\underline{\theta}^{(k)}$ as defined by equation (2.2.8), tacitly assuming that an associated $\rho^{(k)} = 1$ produced an acceptable iteration; the stopping criterion was based on equation (2.2.6). In view of the comments at the end of Section 3.5 we may sometimes need to introduce partial shifts $\rho^{(k)} < 1$.

Generally, an iteration is acceptable if and only if there exists a positive definite matrix $\underline{R}^{(k)}$ such that

$$\underline{v}^{(k)} = -\underline{R}^{(k)} \underline{q}^{(k)} \qquad\qquad (4.2.8)$$

where

$$\underline{q}^{(k)} = -\underline{X}^{(k)^T} \underline{W}\underline{Y}^{(k)} \left[= \left(\frac{\partial \psi_k}{\partial \underline{\theta}}\right) \right] \qquad\qquad (4.2.9)$$

so that we obtain

$$\underline{\theta}^{(k+1)} = \underline{\theta}^{(k)} - \rho^{(k)} \underline{R}^{(k)} \underline{q}^{(k)} . \qquad\qquad (4.2.10)$$

Iterative schemes based on (4.2.10) are called gradient methods and differ in the manner of choosing $\underline{R}^{(k)}$ and $\rho^{(k)}$. One or two possibilities are given below; for further details and proof of the above result see e.g [3].

The simplest gradient method is that of steepest descent, which sets $\underline{R}^{(k)} = \underline{I}_p$ (the (pxp) identity matrix); as its name implies the resulting direction $-\underline{q}^{(k)}$ is the one in which the objective function $\psi(\underline{\theta})$ decreases most rapidly, at least initially. The problem of determining an appropriate step length $\rho^{(k)}$ remains. Although not generally recommended for practical applications, the steepest descent method may prove useful in cases where our current parameter estimates are insufficiently accurate to provide convergence of the standard non-linear least squares methods.

Newton's method sets

$$\underline{R}^{(k)} = [\underline{H}^{(k)}]^{-1}$$

162

where the (pxp) Hessian matrix $\underset{\sim}{H}^{(k)}$ has (j,l)th element

$$H_{jl}^{(k)} = \frac{\partial^2 \psi_k}{\partial \theta_j \partial \theta_l} \Bigg|_{\underline{\theta}^{(k)}} \quad (j,l=1\ldots p) \qquad (4.2.11)$$

and takes $\rho^{(k)} = 1$. The underlying motivation for Newton's method is that if $\psi(\underline{\theta})$ is a quadratic function then (provided $\underset{\sim}{H}^{(k)}$ is positive definite) the method converges in a single iteration

$$\underline{\theta}^{(k+1)} = \underline{\theta}^{(k)} - [\underset{\sim}{H}^{(k)}]^{-1} \underline{g}^{(k)} \qquad (4.2.12)$$

to the minimum. In any case, provided we are within some neighbourhood of the minimum in which $\underset{\sim}{H}^{(k)}$ is positive, convergence will be quadratic.

Putting

$$\underset{\sim}{N}^{(k)} = \underset{\sim}{X}^{(k)^T} \underset{\sim}{W} \underset{\sim}{X}^{(k)} \quad ,$$

note that $\underset{\sim}{H}^{(k)}$ and $\underset{\sim}{N}^{(k)}$ differ by the correction matrix $\underset{\sim}{C}^{(k)}$ (say) whose (j,l)th element is given by

$$C_{jl}^{(k)} = \sum_{\underline{h}} W_{\underline{h}} \; (FO^*(\underline{h}) - FC(\underline{h})) \; \frac{\partial^2 FC(\underline{h})}{\partial \theta_j \partial \theta_l} \Bigg|_{\underline{\theta}^{(k)}} \quad (j,l=1\ldots p).$$

$$(4.2.13)$$

Replacing $\underset{\sim}{H}^{(k)}$ in (4.2.12) by the approximation $\underset{\sim}{N}^{(k)}$, we are led to the iterative scheme

$$\underline{\theta}^{(k+1)} = \underline{\theta}^{(k)} - [\underline{N}^{(k)}]^{-1} \underline{g}^{(k)} \qquad (4.2.14)$$

which defines the Gauss-Newton method; it is precisely our non-linear least squares method based on equations (2.2.5) and (2.2.8). The goodness of this approximation obviously depends on the matrix $\underline{C}^{(k)}$, which is likely to be small in view of the $(FO*(\underline{h}) - FC(\underline{h}))$ terms appearing in (4.2.13). Whether the Newton method is sufficiently more efficient to make the evaluation of second derivatives worthwhile is an open question, though the following tentative conclusions ([3]p.91) justify the Gauss(-Newton) approach:

'(1) If the model fits the data well, the Gauss method often requires no more iterations than the Newton method [2].

(2) If the model does not fit well, the Newton method may require fewer iterations than the Gauss method, but the computing times for the two methods are roughly the same [29].'

We next consider the Marquardt method [55] which basically combines the steepest descent and Gauss-Newton methods. It is based on

$$\underline{R}^{(k)} = [\underline{N}^{(k)} + \lambda^{(k)} \underline{I}_p]^{-1} \qquad (4.2.15)$$

where $\lambda^{(k)}$ is chosen to ensure that $\underline{R}^{(k)}$ is positive definite. We continue to take $\rho^{(k)} = 1$. Note that as $\lambda^{(k)}$ increases we approach the steepest descent direction, whilst for small $\lambda^{(k)}$ we effectively have the Gauss-Newton method. The Marquardt method typically proceeds with

initial iterations moving in the steepest descent direction before gradually changing to an approximate Gauss-Newton approach. It has proved to be very reliable in practice; it is recommended that it be tried when other standard methods fail to converge.

A few concluding remarks will now be made about non-linear parameter estimation schemes; specifically, brief comments will be made on the areas of initial parameter estimates, step lengths, stopping rules and convergence. Initial estimates are a necessary preliminary requirement before our iterative schemes may be invoked; in crystallographic studies such accurate estimates will typically be available (together with a specified covariance matrix) based on prior chemical constraints, symmetry considerations, results from similar previous analyses, etc. For crystal structure refinement methods with which this thesis is concerned, we generally assume that they are sufficiently accurate to ensure (in the centrosymmetric case) that most observed reflections within a non-trivial $(\sin \theta)/\lambda$ limit will be given the correct sign based on the corresponding calculated value, as discussed in Section 3.2.

For the standard Gauss-Newton method the assumed step lengths $(\rho^{(k)} = 1)$ will generally provide acceptable iterations, especially in the neighbourhood of the true underlying parameter values. If, however, our residual

sum of squares $\psi_{k+1} > \psi_k$ at any stage, we may resort to various correction schemes. One such scheme is given by the program of Booth and Peterson [9], which basically uses a quadratic interpolation to locate a local minimum of $\psi(\underline{\theta})$, before beginning the iterative cycle again; this method always converges [39]. Specifying appropriate step lengths is most difficult in cases where we must resort to steepest descent type methods. In such cases, however, use can probably be made of the Marquardt compromise in which the problem reduced to choice of the appropriate $\lambda^{(k)}$ in equation (4.2.15) to use in conjunction with $\rho^{(k)} = 1$. Guidelines are given in [55].

As we have seen, for our iterative estimation procedures we shall typically wish to cycle until the difference between successive estimates of $\underline{\theta}$ becomes sufficiently small in an arbitrary but specified sense e.g. as defined by (2.2.6). For general comments on appropriate such stopping rules see [14]. Crystallographers traditionally consider the R - factor defined by

$$R = \frac{\sum_{\underline{h}} |(FO*(\underline{h}) - FC(\underline{h}))|}{\sum_{\underline{h}} |FO(\underline{h})|} \qquad (4.2.16)$$

when making decisions on whether to accept the current parameter estimate; the number of observations appearing in R and their nature (e.g. position of the corresponding $\underline{h}$ in reciprocal space) influence its value. There are

166

many pitfalls associated with putting too much weight on this statistic, however, (see e.g. [33]); its value as a measure of how well our current model fits the observed data still provides a useful check at any stage of our analysis though and models producing high R-values (which crystallographers should be able to quantify for any given crystal structure analysis) are unlikely to be acceptable. More acceptable perhaps as a method of accessing the correctness of a structure is an analysis of a difference map (see Section 2.4) based on the refined structure. This should reveal no fluctuations in electron density greater than those expected on the basis of the estimated precision of the electron density. Alternative stopping rules and further comments are given in Chapter 3 in which they form an integral part of the underlying theory.

Finally, a brief note on convergence: convergence proofs are difficult for non-linear parameter estimation schemes. Also, even when a method may be known to converge in theory, convergence may be slow in practice and/or computational requirements infeasibly high. For our purposes we content ourselves with the knowledge that our schemes are likely to provide at least locally optimal values for $\underline{\theta}$ (e.g. local minima for $\psi(\underline{\theta})$) in the standard least squares framework of Section 2.2.

4.3   Review of Some of the Work of Bates and Watts

We shall be considering the non-linear model in which the relationship between the values of a response y_i (i=1..n) collected at corresponding experimental settings $\underline{x}_i$ can be written

$$Y_i = f(\underline{x}_i, \underline{\theta}) = \epsilon_i \tag{4.3.1}$$

where $\underline{\theta} = (\theta_1 \,..\, \theta_p)^T$ is our set of unknown parameters and the ϵ_i are independent and identically distributed $N(0, \sigma^2)$ errors. The solution locus is described by

$$\underline{n}(\underline{\theta}) = (\eta_1(\underline{\theta}) \,\cdots\, \eta_n(\underline{\theta})^T$$

where

$$\eta_i(\underline{\theta}) = f(\underline{x}_i, \underline{\theta});$$

$f(\underline{x}, \underline{\theta})$ is termed the model function.

Most algorithms for computing the least squares estimate $\hat{\underline{\theta}}$ and most inference methods for non-linear models are based on a local linear approximation to the model [14]. This effectively replaces the solution locus by its tangent plane at $\underline{n}(\underline{\theta}_0)$ for some fixed trial parameter value $\underline{\theta}_0$. A uniform co-ordinate system is simultaneously imposed on that tangent plane. The adequacy of the planar assumption component of the linear approximation used will be fixed i.e. invariant under re-parameterisation. This may be quantified by the intrinsic curvature derived in

the original paper [4] of Bates and Watts. We note that the planar assumption must be satisfied to a sufficient degree before it is sensible to discuss whether the uniform co-ordinate assumption is acceptable. The validity of the uniform co-ordinate assumption may then be measured by the parameter effects curvature derived in [4]; this is not fixed, and suitable re-parameterisations may reduce this curvature as discussed in their follow-up paper [5]. The main results of these two papers will now be summarised. Note also that the two papers reviewed here have since formed the basis of the book of non-linear regression analysis [6] by the same authors. Further details on the general areas of fitting non-linear models to data and improved presentation of inferential results of non-linear analyses may also be found there, together with comments on the geometrical properties of non-linear least squares.

Let an arbitrary straight line in parameter space through $\underline{\theta}_0$ be represented by

$$\underline{\theta}(b) = \underline{\theta}_0 + b\underline{h}$$

where $\underline{h}$ is any non-zero (px1) vector. The corresponding curve on the solution locus may be written

$$\underline{\eta}_{\underline{h}}(b) = \underline{\eta}(\underline{\theta}_0 + b\underline{h}) \quad .$$

The tangent to the curve $\underline{\eta}_{\underline{h}}(b)$ at b=0 may be written

$$\dot{n}_h = \underset{\sim}{V} \cdot \underset{-}{h} \qquad\qquad (4.3.2)$$

where $\underset{\sim}{V}.$ is the (nxp) matrix whose j^{th} column $\underset{-}{v}_j$ is defined by

$$\underset{-}{v}_j = \left[\frac{\partial \eta_1}{\partial \theta_j} \Big|_{\underset{-}{\theta}_0} \cdots\cdots \frac{\partial \eta_n}{\partial \theta_j} \Big|_{\underset{-}{\theta}_0} \right]^{T} \quad (j=1\ldots p) \quad . \qquad (4.3.3)$$

The tangent plane π at $\underset{-}{n}(\underset{-}{\theta}_0)$ may be defined as the set of all such linear combinations of the vectors $\{\underset{-}{v}_j\}$. Similarly, by defining second partial derivative vectors $\{\underset{-}{v}_{jk}\}$ by

$$\underset{-}{v}_{jk} = \frac{\partial^2 \underset{-}{n}}{\partial \theta_j \partial \theta_k} \Big|_{\underset{-}{\theta}_0} \quad (j,k=1..p) \qquad\qquad (4.3.4)$$

and collecting them into the (pxp) matrix of n vectors $\underset{\sim}{V}..$, the second derivative of $\underset{-}{n}_h$ (b) at b=0 can be written

$$\ddot{n}_h = \underset{-}{h}^{T} \underset{\sim}{V}.. \underset{-}{h} \qquad\qquad (4.3.5)$$

(i.e. each element of $\ddot{n}_h$ is a term of the form $\underset{-}{h}^{T} \underset{\sim}{V}_i \underset{-}{h}$ where $\underset{\sim}{V}_i$ is the ith face of the array $\underset{\sim}{V}..$). $\ddot{n}_h$ can be written as three components: $\ddot{n}_h^{N}$ normal to the tangent plane, $\ddot{n}_h^{P}$ parallel to $\dot{n}_h$, and $\ddot{n}_h^{G}$ parallel to the tangent plane normal to $\dot{n}_h$, so

$$\ddot{n}_h = \ddot{n}_h^{N} + \ddot{n}_h^{P} + \ddot{n}_h^{G} \quad .$$

If we think of $\ddot{\underline{n}}_{\underline{h}}$ as an acceleration vector, the tangential components, which are caused by the model parameterisation, may be combined into a tangential acceleration

$$ \ddot{\underline{n}}_{\underline{h}}^{T} = \ddot{\underline{n}}_{\underline{h}}^{P} + \ddot{\underline{n}}_{\underline{h}}^{G} \; . $$

This can be converted into a _parameter effects curvature_ in the direction $\underline{h}$ as

$$ K_{\underline{h}}^{T} = \| \ddot{\underline{n}}_{\underline{h}}^{T} \| \; / \; \| \dot{\underline{n}}_{\underline{h}} \|^{2} \; . \qquad (4.3.6) $$

Similarly, the _intrinsic curvature_ may be defined by

$$ K_{\underline{h}}^{N} = \| \ddot{\underline{n}}_{\underline{h}}^{N} \| \; / \; \| \dot{\underline{n}}_{\underline{h}} \|^{2} \; . \qquad (4.3.7) $$

Unlike $K_{\underline{h}}^{T}$ this is independent of the parameterisation and has the important geometrical interpretation of being the inverse of the radius of the circle that best approximates the solution locus in the direction of $\dot{\underline{n}}_{\underline{h}}$ at $\underline{\theta}_{0}$.

Defining the standard radius

$$ \rho = s \surd p \qquad (4.3.8) $$

where s^2 is an estimate of σ^2 based on ν degrees of freedom, we arrive at the final definitions for the parameter effects curvature $(\gamma_{\underline{h}}^T)$ and the intrinsic curvature $(\gamma_{\underline{h}}^N)$ to be given by

$$\gamma_{\underline{h}}^T = K_{\underline{h}}^T \rho \qquad\qquad (4.3.9)$$

and

$$\gamma_{\underline{h}}^N = K_{\underline{h}}^N \rho \quad . \qquad\qquad (4.3.10)$$

We may take as measures of the severity of non-linearity the maximum curvatures

$$\Gamma^N = \max_{\underline{h}} \gamma_{\underline{h}}^N$$

and

$$\Gamma^T = \max_{\underline{h}} \gamma_{\underline{h}}^T \ ;$$

an algorithm for the evaluation of Γ^N and Γ^T is given in [4]. We shall concentrate here on the formulation of an appropriate curvature array (and its interpretation), which will enable us to calculate the quantities (4.3.9) and (4.3.10) and will play a central role in the ensuing theory.

To determine the curvature array we first form an orthogonal-triangular decomposition of $\underline{\underline{V}}$. [13],

$$\underset{\sim}{V}\cdot \; = \; \underset{\sim}{Q}\;\underset{\sim}{R} \qquad\qquad (4.3.11)$$

where $\underset{\sim}{Q}$ is an (nxn) orthogonal matrix and $\underset{\sim}{R}$ is an (nxp) matrix with zeros below the main diagonal. The upper (pxp) sub-matrix of $\underset{\sim}{R}$ is denoted $\underset{\sim}{R}_1$ and we define

$$\underset{\sim}{L} \; = \; \underset{\sim}{R}_1^{-1} \; . \qquad\qquad (4.3.12)$$

By letting multiplications using square brackets indicate summations over the appropriate index consistent with the dimensions, while multiplications without square brackets indicate summations over the second or third index for premultiplication and postmultiplication respectively, a curvature array $\underset{\sim}{A}\cdot\cdot$ may be written

$$\underset{\sim}{A}\cdot\cdot \; = \; [\underset{\sim}{Q}^T]\;[\underset{\sim}{L}^T\;\underset{\sim}{V}\cdot\cdot\;\underset{\sim}{L}] \; . \qquad\qquad (4.3.13)$$

This can be regarded as the second derivative array for the parameters

$$\underset{\sim}{\phi} \; = \; \underset{\sim}{L}^T\;(\underset{\sim}{\theta} - \underset{\sim}{\hat{\theta}}) \qquad\qquad (4.3.14)$$

in a rotated set of sample space co-ordinates.

The first p faces of $\underset{\sim}{A}\cdot\cdot$ constitute the parameter-effects curvature array $\underset{\sim}{A}\cdot\cdot^T$. By denoting the first p columns of $\underset{\sim}{Q}$ by $\underset{\sim}{U}\cdot$, we may also write

$$\underset{\sim}{A}\cdot\cdot^T \; = \; [\underset{\sim}{U}\cdot^T]\;[\underset{\sim}{L}^T\;\underset{\sim}{V}\cdot\cdot\;\underset{\sim}{L}] \; . \qquad\qquad (4.3.15)$$

173

Similarly, by denoting the remaining (n-p) columns of $\underset{\sim}{Q}$ by $\underset{\sim}{N}$, we obtain the intrinsic curvature array

$$\underset{\sim}{A}..^{N} = [\underset{\sim}{N}^{T}] \; [\underset{\sim}{L}^{T} \; \underset{\sim}{V}.. \; \underset{\sim}{L}] \quad . \tag{4.3.16}$$

The value of the arrays $\underset{\sim}{A}..^{T}$ and $\underset{\sim}{A}..^{N}$ lies in the fact that for any unit vector $\underline{d}$, we have

$$\gamma_{\underline{L}\underline{d}}^{T} = \| \; \underline{d}^{T} \; \underset{\sim}{A}..^{T} \; \underline{d} \; \| \tag{4.3.17}$$

and

$$\gamma_{\underline{L}\underline{d}}^{N} = \| \; \underline{d}^{T} \; \underset{\sim}{A}..^{N} \; \underline{d} \; \| \tag{4.3.18}$$

i.e. relatively simple expressions for the evaluation of parameter effects and intrinsic curvatures respectively in various directions. In particular, equation (4.3.18) may be used to determine (the reciprocals of) the radii of curvature needed in Section 4.5.

Bates and Watts [5] consider further the adequacy of a (1-α) linear approximation confidence region consisting of those values of $\underline{\theta}$ for which

$$(\underline{\theta} - \hat{\underline{\theta}})^{T} \; \underset{\sim}{V}.^{T} \; \underset{\sim}{V}. \; (\underline{\theta} - \hat{\underline{\theta}}) \leq F(\rho, \nu; \; \alpha) \tag{4.3.19}$$

where $\underset{\sim}{V}.$ is calculated for the scaled responses (i.e. divided by the standard radius ρ) and $F(\rho, \nu; \; \alpha)$ is the upper α probability point of the F distribution with p and ν degrees of freedom. This may be determined by comparing the curvatures

$$a_{jkl} = (\underset{\sim}{A}..^{T})_{jkl} \qquad\qquad (4.3.20)$$

and Γ^{T} to the curvature $1/R_{\alpha}$, where R_{α} is the radius of the $(1-\alpha)$ conservative confidence region disc for $\underset{\sim}{r}$ -

$$\| \underset{\sim}{r} \| \leq R_{\alpha}$$

- where $\underset{\sim}{r}$ is defined by

$$\underset{\sim}{r} = \underset{\sim}{U}.^{T} (\underset{\sim}{n}(\underset{\sim}{\ell}) - \underset{\sim}{n}(\hat{\underset{\sim}{\ell}})) \quad . \qquad\qquad (4.3.21)$$

An explicit expression for R_{α} is derived in their paper.

We shall now give interpretations of the individual curvatures a_{jkl}. The a_{jjj} are compansion terms: they cause compression or expansion of scale along a ϕ_{j} parameter line. The a_{jkk} are arcing terms: they cause changes in the $\underset{\sim}{u}_{j}$ direction (tangent vector $\underset{\sim}{u}_{j} = \left.\dfrac{\partial \underset{\sim}{r}}{\partial \phi_{j}}\right|_{\underset{\sim}{\phi}_{0}}$) of the ϕ_{k} parameter lines as we move along them. The a_{jkj} are fanning terms: they cause changes in the $\underset{\sim}{u}_{j}$ direction of the ϕ_{k} parameter curves as we move across the ϕ_{j} parameters curves. The a_{jkl} are torsion terms: they cause a change in the $\underset{\sim}{u}_{j}$ direction of the $\underset{\sim}{u}_{k}$ tangent vector due to a unit change on ϕ_{l}. A perfect parameterisation to eliminate parameter effects would be one which reduced all individual curvatures in the array $\underset{\sim}{A}..^{T}$ (henceforth

175

abbreviated to A) to zero.

The geometric significance of the parameter effects curvature is as follows. The uniform co-ordinate assumption replaces the curved ϕ parameter lines on the approximating tangent plane by a grid of straight, parallel, equispaced lines. By interpreting r of equation (4.3.21) as the tangent plane co-ordinate vector, we see that this assumption will be justified - and zero parameter effects curvature ensue - if, for all parameter values $\underline{\theta}*$, we may recover $\underline{\theta}*$ from the projection of $\underline{n}(\underline{\theta}*) - \underline{n}(\hat{\underline{\theta}})$ onto the tangent plane. This will be the motivation behind the attempts of the next Section to find re-parameterisations to eliminate parameter effects.

Returning to our previous theory, by writing $r = H(\phi)$, appropriate confidence regions in the $\underline{\theta}$ parameters may be formed, bounded by

$$\{\hat{\underline{\theta}} + \underline{L}\ H^{-1}(\underline{r})\ :\ \|\ \underline{r}\ \| = R_\alpha\} \quad . \qquad (4.3.22)$$

Because of the difficulties in the evaluation of H^{-1} - the second order Taylor Series approximation

$$H^{-1}(\underline{r}) = \underline{r} - (\underline{r}^T\ \underline{A}\ \underline{r})/2$$

is often inadequate - it is generally found that re-parameterisation (and subsequent linear approximation) is a better approach to the problem of dealing with parameter

effects. Some relevant resultant theory [5] is summarised below.

Consider a re-parameterisation in which the new parameters $\underline{\phi}$ are non-linear transformations of $\underline{\theta}$,

$$\phi_j = G_j(\underline{\theta}) \qquad (j=1..p),$$

with inverse transformation

$$\theta_j = K_j(\underline{\beta}) \qquad (j=1...p).$$

Dropping tildes for notational convenience, form the (pxp) Jacobian matrix G. (with elements $\partial G_j/\partial \theta_k$) and the (pxpxp) second derative array G.. (with elements $\partial^2 G_j/\partial \theta_k \partial \theta_l$). Then the new parameter effects array $\widetilde{A}$ (say) may be written in terms of the original via

$$\widetilde{A} = A-[L^{-1}] \ [L^{-1}[G.^{-1}] \ [G..] \ L]. \qquad (4.3.23)$$

This equation forms the starting point for seeking re-parameterisations to give zero parameter effects.

Setting $\widetilde{A}$ equal to zero gives

$$[G.^{-1}][G..] = [(V.^T V.)^{-1} V.^T] \ [V..] = T* \ (say) \qquad (4.3.24)$$

when re-written in terms of the original parameters and derivative matrices. The target transformation $\phi* = G*(\theta)$ should therefore satisfy

$$[G.*^{-1}] \; [G..*] = T*$$

or, equivalently,

$$G..* = [G.*] \; [T*] \qquad\qquad (4.3.25)$$

where T* is the resulting target transformation curvature array given above. We could seek to solve these equations completely generally in order to provide a global solution for G* or, alternatively, we could seek to obtain a particular solution to equation (4.3.25) in which T* is evaluated at $\hat{\theta}$ i.e. to solve

$$G..* = [G.*] \; [T*(\hat{\theta})] \; . \qquad\qquad (4.3.26)$$

Neither of those methods appear to be of much practical use, however, not least because they will involve transformations which do not permit ready interpretation of the parameters. More promising may be the use of a particular class of transformation, such as Ross's expected-value transformation [59]. Problems of invertibility still exist, though it may be possible to restrict the transformation to overcome this; although this may not eliminate parameter-effects completely, it can substantially reduce them. In the next Section we shall present ideas of our own aimed at eliminating the parameter-effects curvature.

The theory of this Section has laid the foundations for

the remainder of the Chapter, in which we seek improved methods of non-linear parameter estimation.

4.4 Possible Re-parameterisations Based on Seeking Orthogonal Mappings from Tangent Plane to Solution Locus

Using the notation of the previous Section, we return to the non-linear model of equation (4.3.1). We note that observed data leads to parameter estimates (via least squares) based on our tangent plane approximation, which we shall here assume to be acceptable. It is the relation between such points on the tangent plane and their analogues on the solution that will now be discussed. This is not fixed and alternative parameter estimates may be suggested by various mappings from tangent plane to solution locus.

We recall that the least squares estimate $\hat{\theta}$ of θ is the value which minimises the sum of squares of residuals

$$S(\theta) = \sum_{i=1}^{n} (Y_i - \eta_i(\theta))^2 \quad .$$

In practice, $\eta(\theta)$ is typically approximated by

$$\eta(\theta) \simeq \eta(\theta_0) + \sum_{j=1}^{p} (\theta_j - \theta_{j0}) \left. \frac{\partial \eta}{\partial \theta_j} \right|_{\theta_0} \tag{4.4.1}$$

i.e. an approximation about initial value $\underline{\theta}_0$ which is linear in the individual parameter increments. Geometrically, we drop a perpendicular from the observed (n dimensional) data point D (say) onto the (p dimensional) approximating tangent plane π at P = $\underline{\eta}(\underline{\theta}_0)$. The corresponding point on the (p dimensional) solution locus, Q_0 = $\eta(\underline{\theta})$, will not in general lie on this perpendicular as we see below. Our discussion is related to that on the geometry of non-linear least squares found in [24].

Figure 4.4.1 shows the solution locus (with units of θ shown on it) for a non-linear model with n = 2 observations and p = 1 parameter (θ). Assume without loss of generality that our initial approximation θ_0 = 0 and
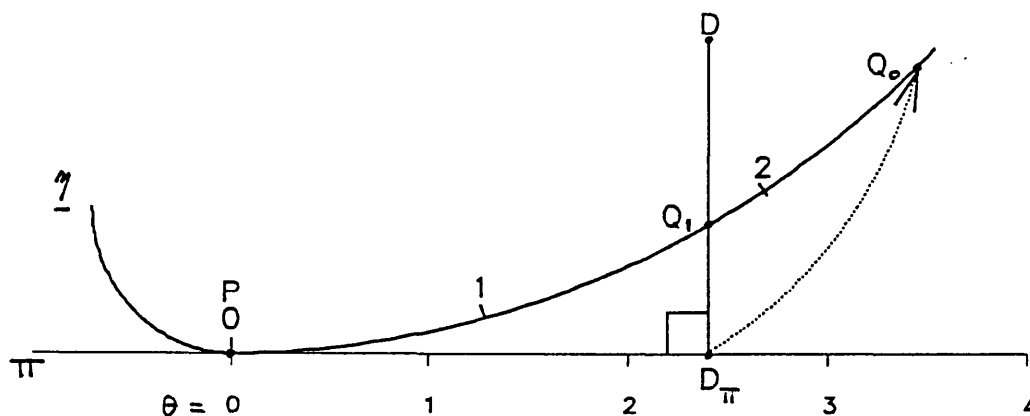
Figure 4.4.1:   Solution locus for a non-linear model
                with 2 observations and 1 parameter

that the point marked $\theta = 1$ is the point on the solution locus $\underline{\eta}$ obtained when $\theta = 1$ etc. The (line) tangent π to the solution locus at θ_o is also shown graduated with units of θ which are obtained from the rate of change found at θ_0. These units will be equally spaced; the fact that those on the solution locus are not results from the non-linearity and the non-uniformity of the co-ordinate system. It is this non-uniformity that gives rise to the parameter effects curvature of Section 4.3.

The least squares estimate of θ based on the linear assumption is the corresponding value of θ at the point, D_{π}, on the tangent line π such that DD_{π} is perpendicular to π (where D is the 2-dimensional data point). In the next iteration of the linearisation procedure we will use the tangent line at the point Q_0 (the point on $\underline{\eta}$ corresponding to our least squares estimate just found). It is proposed here to try to seek the alternative new approximation point Q_1 (the point on $\underline{\eta}$ lying on DD_{π}). In cases of high parameter effects curvature - where the rate of change at θ_0, and hence the linearised units of θ, are small, but the actual units increase sharply as in Figure 4.4.2 - such an alternative will accelerate our iteration procedure. Equivalently, we may seek a re-parameterisation such that the mapping from tangent plane to solution locus maps D_{π} to Q_1 as opposed to Q_0, in order to simultaneously speed up the convergence procedure and reduce the parameter effects curvature as discussed in the
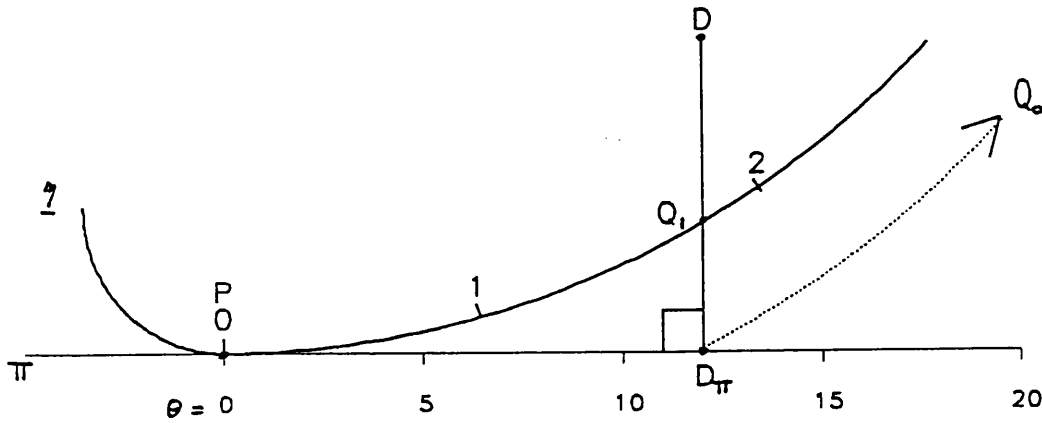
Figure 4.4.2: Solution locus for a non-linear model with 2 observations and 1 parameter (exhibiting high parameter effects curvature)

previous Section. The extension of this basic idea to the model defined by equation (4.3.1) with general $n > p \geq 1$ is relatively simple, though pictorial representation is not. Suitable re-parameterisations are sought below.

Let Q be a general point on the solution locus i.e.

$$Q = (\eta_1(\underline{\theta}) \ldots \eta_n(\underline{\theta})) \quad . \tag{4.4.2}$$

Let Q_π be the projection of Q onto the approximating tangent plane at $\underline{\eta}(\underline{\theta}_0)$. We then have

$$Q_\pi = (\eta_1(\underline{\theta}_0) \ldots \eta_n(\underline{\theta}_0)) + \sum_{k=1}^{p} (\theta_k^* - \theta_{k0}) \underline{v}_k^T \tag{4.4.3}$$

for some $\underline{\theta}*$ (in general $\neq \underline{\theta}$) where

182

$$\underline{v}_j^T = (\frac{\partial \eta_1}{\partial \theta_j}\Big|_{\underline{\theta}_0} \cdots \frac{\partial \eta_n}{\partial \theta_j}\Big|_{\underline{\theta}_0}) \quad .$$

By construction, Q_π satisfies

$$(\underline{Q} - \underline{Q}_\pi) \cdot \underline{v}_k = 0 \qquad (k = 1..p) \quad . \qquad (4.4.4)$$

Equation (4.4.4) gives us a set of (p) linear equations in the p unkowns θ_j^* , namely

$$\sum_{i=1}^{n} (\eta_i(\underline{\theta}) - \eta_i(\underline{\theta}_0) - \sum_{k=1}^{p} \frac{\partial \eta_i}{\partial \theta_k}\Big|_{\underline{\theta}_0} (\theta_k^* - \theta_{k0})) \frac{\partial \eta_i}{\partial \theta_j}\Big|_{\underline{\theta}_0} = 0$$

$$(j = 1..p) \qquad (4.4.5)$$

from which we may deduce

$$\underline{\eta}(\underline{\theta}_0) - \underline{\eta}(\underline{\theta}) + \underset{\sim}{X}(\underline{\theta}* - \underline{\theta}_0) \cdot \underset{\sim}{X}(\underline{\theta}* - \underline{\theta}_0) = 0 \qquad (4.4.6)$$

where $\underset{\sim}{X}$ (equivalently $\underset{\sim}{V}$. of Section 4.3) is the (nxp) design matrix with columns $\underline{v}_k$ (k = 1..p). Solution of the equations (4.4.5) yields

$$\underline{\theta}* = \underline{\theta}_0 + \underset{\sim}{\Lambda}^{-1} \underline{c} \qquad (4.4.7)$$

where

$$\Lambda_{(kl)} = \sum_{i=1}^{n} v_{ki} v_{li} \qquad (k,l=1..p) \qquad (4.4.8)$$

and

$$c_{(k)} = \sum_{i=1}^{n} (\eta_i(\underline{\theta}) - \eta_i(\underline{\theta}_0)) v_{ki} \qquad (k=1..p) \qquad (4.4.9)$$

183

with

$$v_{ki} = \left. \frac{\partial \eta_i}{\partial \theta_k} \right|_{\underline{\theta}_0} \quad .$$

Consider now a re-parameterisation $\underline{\phi} = G(\underline{\theta})$ i.e.

$$\phi_j = G_j(\underline{\theta}) \qquad (j = 1..p) \quad . \qquad (4.4.10)$$

Let $\underline{\xi}(\underline{\phi})$ denote the analogue of $\underline{\eta}(\underline{\theta})$ and define the (pxp) Jacobian matrix $\underset{\sim}{G}.$ via

$$(G.)_{jk} = \left. \frac{\partial \phi_j}{\partial \theta_k} \right|_{\underline{\theta}_0} \qquad (j, \ k = 1..p) \quad . \qquad (4.4.11)$$

By noting that

$$\left. \frac{\partial \xi_i}{\partial \phi_j} \right|_{\underline{\phi}_0} = \sum_k \left. \frac{\partial \xi_i}{\partial \theta_k} \right|_{\underline{\phi}_0} \left. \frac{\partial \theta_k}{\partial \phi_j} \right|_{\underline{\phi}_0} = \sum_k X_{ik} \ (G.^{-1})_{kj} \qquad (4.4.12)$$

our previous ideas lead us to seek that re-parameterisation for which

$$\underline{\eta}(\underline{\theta}_0) - \underline{\eta}(\underline{\theta}) + \underset{\sim}{X} \underset{\sim}{G}.^{-1}(\underline{\phi} - \underline{\phi}_0).\underset{\sim}{X} \underset{\sim}{G}.^{-1}(\underline{\phi} - \underline{\phi}_0) = 0;$$

$$(4.4.13)$$

by comparison with (4.4.6) we set

$$\underline{\phi} - \underline{\phi}_0 = \underset{\sim}{G}.(\underline{\theta}^* - \underline{\theta}_0)$$

or, equivalently,

184

$$\phi_j - \phi_{0j} = \sum_k \frac{\partial \phi_j}{\partial \theta_k} (\theta_k^* - \theta_{0k}) \qquad (j = 1..p),$$

which suggests the re-parameterisation

$$\phi_j = \theta_j^* \qquad (j = 1..p). \qquad (4.4.14)$$

We emphasise that the geometric interpretation of our re-parameterisation is that the analogue of any point on the solution surface is now precisely its projection onto the approximating tangent plane, or equivalently, our mapping from tangent plane to solution surface is orthogonal to the tangent plane. We note also that further simplification results by taking ϕ as the linear function of θ^* defined by

$$\phi = \Lambda^{-1} c^* \qquad (4.4.15)$$

where Λ is defined by (4.4.8) and

$$c_{(k)}^* = \sum_{i=1}^{n} \eta_i(\theta) \, v_{ki} \qquad (k = 1..p) \qquad (4.4.16)$$

The effectiveness of such a re-parameterisation will be quantified in Section 4.6; meanwhile, a trivial example of these ideas will be presented. Consider the model function

$$\eta(x,\theta) = \theta^x$$

185

with observations at $x = 1$ and $x = 2$, and initial estimate $\theta_0 = 1$. The (1x1) matrix $\underset{\sim}{\Lambda}$ and (1x1) scalar $\underline{C}*$ are 5 and $(\theta + 2\theta^2)$ respectively, suggesting the re-parameterisation

$$\phi = \frac{\theta + 2\theta^2}{5} \qquad\qquad (4.4.17)$$

with inverse

$$\theta = (\frac{1}{16} + \frac{5}{2} \phi)^{\frac{1}{2}} - \frac{1}{4} \quad .$$

To fully appreciate the geometrical significance of this re-parameterisation we refer to Figure 4.4.3 which represents the solution locus $y = x^2$ together with the projection of the general point $Q(x, x^2)$ onto the tangent line at $P(1,1)$. The co-ordinates of the projected point Q_π suggest the re-parameterisation

$$\phi^* = \frac{\theta + 2\theta^2 + 2}{5}$$

i.e. a linear function of equation (4.4.17). With data $y_1 = 1.1$, $y_2 = 1.2$ for example, linear approximation under the new model leads to $\hat{\phi} = 0.7$, corresponding to $\hat{\theta} = (\sqrt{29}-1)/4$, whereas under the original model $\hat{\theta} = 1.1$; the original residual sum of squares 1×10^{-4} is reduced to 1.7194×10^{-5}.
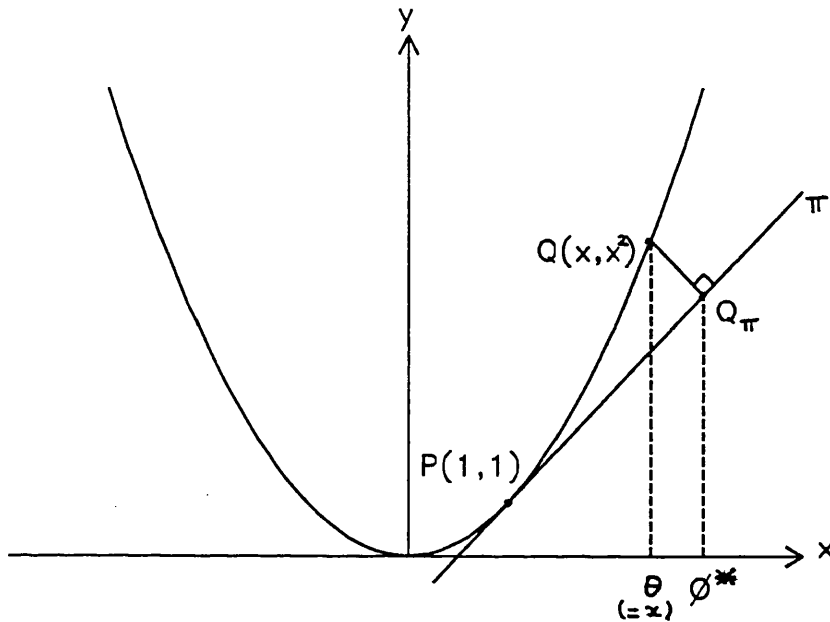
Figure 4.4.3: Diagram suggesting possible re-
parameteridation when $\eta(x,\theta) = \theta^x$
with observations at $x = 1$ and $x = 2$,
with $\theta_0 = 1$

We note next that we may use equation (4.4.12) to show

that under a 1st order Taylor series approximation, the

least squares increment in the $\underline{\phi}$ parameters - $\underline{\Delta\phi}$ (say) -

is trivially related to the original least squares

increment in the $\underline{\theta}$ parameters - $\underline{\Delta\theta}_0$ (say) - via

$$\underline{\Delta\phi} = \underset{\sim}{G} \cdot \underline{\Delta\theta}_0 \qquad\qquad (4.4.18)$$

where $\underset{\sim}{G}$ is defined by equation (4.4.11). We turn now to

the question of invertibility i.e. the recovery of the

least squares increment in the original parameters - $\underline{\Delta\theta}$

187

(say) - which satisfies

$$\underset{\sim}{\phi}_0 + \underset{\sim}{\Delta\phi} = G(\underset{\sim}{\theta}_0 + \underset{\sim}{\Delta\theta})$$

or, equivalently,

$$\underset{\sim}{\theta}_0 + \underset{\sim}{\Delta\theta} = K(\underset{\sim}{\phi}_0 + \underset{\sim}{\Delta\phi}) \qquad\qquad (4.4.19)$$

where the inverse function $\underset{\sim}{\theta} = K(\underset{\sim}{\phi})$ i.e.

$$\theta_j = K_j(\underset{\sim}{\phi}) \qquad\qquad (j = 1..p)$$

will in general be unknown.

We may proceed by expanding the right hand side of (4.4.19) in a 2nd order Taylor series to obtain

$$\Delta\theta_j = \sum_k \left. \frac{\partial\theta_j}{\partial\phi_k} \right|_{\phi_0} \Delta\phi_k + \tfrac{1}{2} \underset{\sim}{\Delta\phi}^T \underset{\sim}{S}^j \underset{\sim}{\Delta\phi} \qquad (j=1..p) \qquad (4.4.20)$$

where $\underset{\sim}{S}^j$ is the (pxp) matrix defined by

$$S^j_{kl} = \left. \frac{\partial^2\theta_j}{\partial\phi_k\partial\phi_l} \right|_{\phi_0} \qquad (k,l=1..p) \qquad . \qquad (4.4.21)$$

Using the fact that

$$0 = \frac{\partial^2\theta_j}{\partial\theta_k\partial\theta_l} \qquad (j,k,l=1..p)$$

we may deduce that $\underset{\sim}{S}^j$ satisfies

188

$$0 = \underline{G}._k{}^T \underset{\sim}{S}^j \underline{G}._l + \sum_m \left. \frac{\partial^2 \phi_m}{\partial \theta_k \partial \theta_l} \right|_{\underline{\theta}_0} (G.^{-1})_{jm} \quad (j,k,l=1..p)$$

(4.4.22)

where $\underline{G}._k$ represents the kth column of $G.$ $(k=1..p)$; equation (4.4.22) enables us to determine $\underset{\sim}{S}^j$ $(j=1..p)$. In the special case where $\underset{\sim}{G}.$ is a multiple of the (pxp) identity matrix –

$$\underset{\sim}{G}. = c \underset{\sim}{I}_p$$

– we find that

$$S_{kl}^j = - \frac{1}{c^3} \left. \frac{\partial^2 \phi_j}{\partial \theta_k \partial \theta_l} \right|_{\underline{\theta}_0}$$

whence equations (4.4.20) and (4.4.22) yield the result

$$\Delta \theta_j = \Delta \theta_{0j} - \frac{1}{2c} \Delta \theta_0{}^T \underset{\sim}{G}.._j \Delta \theta_0 \quad (j=1..p)$$

(4.4.23)

where $\underset{\sim}{G}.._j$ is the (pxp) matrix with (kl)th element

$$\left. \frac{\partial^2 \phi_j}{\partial \theta_k \partial \theta_l} \right|_{\underline{\theta}_0} .$$

We conclude this Section by considering re-parameterisation based on the criterion of arc length. Specifically, in the case of a 1-dimensional solution

189

surface, we may wish to associate the point D_π on the tangent line (at P) with the point Q on the solution surface which satisfies

arc length PQ = distance PD_π

(See Figure 4.4.4). Such a mapping has a simple interpretation and has the merit of being locally orthogonal if our tangent plane approximation is reasonable, giving rise to zero parameter effects curvature. Furthermore, such a parameterisation is globally optimum, being essentially independent of the starting point P.
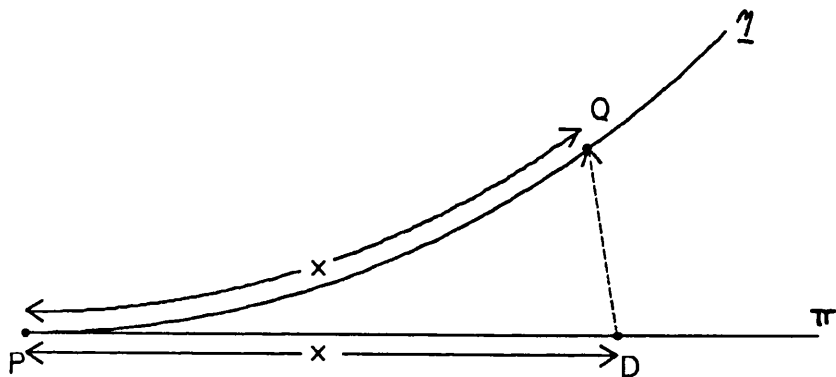
Figure 4.4.4: Mapping from tangent plane to solution
surface based on the criterion of arc
length

To illustrate the above approach we return to our previous example, in which our arc length idea now suggests the parameter estimate r (say) which satisfies

190

$$\int_1^r (1 + 4x^2)^{\frac{1}{2}} \, dx = \sqrt{5} \, \Delta\theta_0 \qquad\qquad (4.4.24)$$

where the original increment $\Delta\theta_0 = 0.1$. By using the substitution

$$x = \tfrac{1}{2} \sin h \, (u)$$

in the integral above we may deduce that r satisfies

$$5 \sin h^{-1}(2r) + 10r(1 + 4r^2)^{\frac{1}{2}} - 5 \sin h^{-1}(2) - 12\sqrt{5} = 0.$$
$$(4.4.25)$$

The solution ($r = 1.09627$ to 5 decimal places) is very similar to the value $- (\sqrt{29}-1)/4$ found under the previous method of perpendicular projection, as of course is the new residual sum of squares 1.7182×10^{-5}.

Unfortunately, there is no natural extension of the above approach that is practically feasible in the general case of p (>1) - dimensional solution surfaces. We merely note that the attractive properties of the arc length parameterisation when $p = 1$ make the quest for such a generalised approach worthwhile, though this will not be pursued here.

4.5 <u>Further Geometrical Approaches to Non-linear Parameter Estimation Based on Fitting Circles of Curvature</u>

Using the notation introduced in the previous Sections we may re-formulate our estimation problem in the following terms: given a general p-dimensional solution surface (η) - characterised by parameter θ - in n-dimensional space, and an n-dimensional data point D, we wish to find that (p-dimensional) parameter value which corresponds to the point on η nearest to D. We typically have available an initial guess at the solution, θ_0 (say), which corresponds to a point P on η.

Standard inference methods for non-linear models (where η is not planar, but a curved surface) are based on local linear approximations to the model, corresponding to imposing a tangent plane (π) at P, together with an underlying uniform co-ordinate system, as discussed in Sections 4.3. & 4.4. The appropriateness of such a co-ordinate system - which depends upon the particular parameterisation used - may be quantified by the parameter effects curvature of Section 4.3, which we shall here assume has been either completely or nearly optimised. In particular, we shall often with to use the concept of 'perpendicular projection' introduced in Section 4.4, whereby the mapping from π to η is orthogonal (to π). First, however, the underlying motivation for the ideas of

the present Section will be given.

We consider the (n =) 2-dimensional case in which the solution surface η is a circle of radius r, and apply the following stability analysis. Three possibilities arise, depending on whether the observed data point D lies on the concave/convex side of η at a distance (from the true least squares point) greater/less than r - note that we eliminate the case of D lying on the concave side of the circle at a distance greater than r. Let P_0 denote the point on the circle representing our initial guess at the least squares point, with corresponding tangent π, and P_1 our next guess. Let P_T denote the true least squares point. The results of an application of our method of perpendicular projection of the previous Section may now be represented as in Figure 4.5.1. Considered as a local analysis, for small perturbations - i.e. good starting point P_0 - cases (a) and (b) are stable whereas case (c), in which D is on the convex side of η at a distance greater than r, isn't. Consequently, we shall henceforth focus our attention on such data points lying on the convex side of η.

In practice, of course, we will not initially know the true least squares point P_T, so we will need to approximate η (which will not in general be a circle) by the appropriate circle of curvature at P_0. Let this circle have centre O; the ideas of this Section seek to
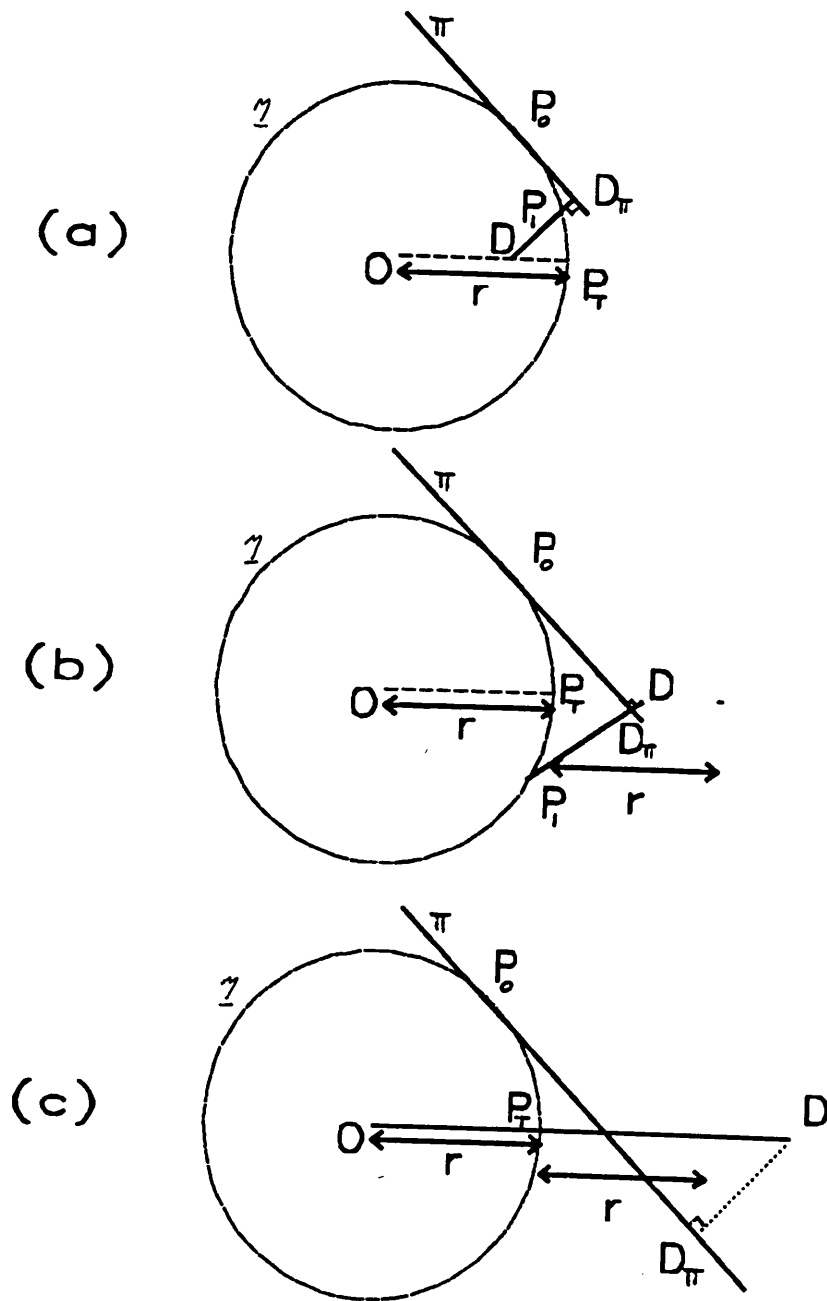
Figure 4.5.1: Stability analysis (n=2, p=1) when
the solution surface is a circle

194

utilise the direction $\overrightarrow{DO}$ rather than the standard direction $\overrightarrow{DD_\pi}$ (where D_π represents the projection of D onto π). In the general case $n > p \geq 1$, by using appropriate curvature measures, we shall wish to obtain better approximations to the solution surface - than those based on the tangent plane alone - and to modify our estimation procedure accordingly.

We shall now give a few simple examples of possible alternatives to the standard least-squares method, based on fitting appropriate circles of curvature, which may be expected not only to speed up convergence but also to work over a wider range of data values. Ideas motivated by such cases in which pictorial representation is feasible need to be extended to the general case. This is attempted later in the Section. We note that the evaluation of the appropriate radii of curvature may be carried out as shown in Section 4.3. It is assumed also that other standard quantities e.g. principal normals, principal/conjugate directions, directions of steepest descent etc may also be readily calculated.

Example 4.5.1 ($n = 2, p = 1...$plane curve)

Instead of fitting the tangent (line) at P, suppose we fit the circle of curvature with centre O along the principal normal (PN) from P (see Figure 4.5.2). The following possibilities present themselves.

195
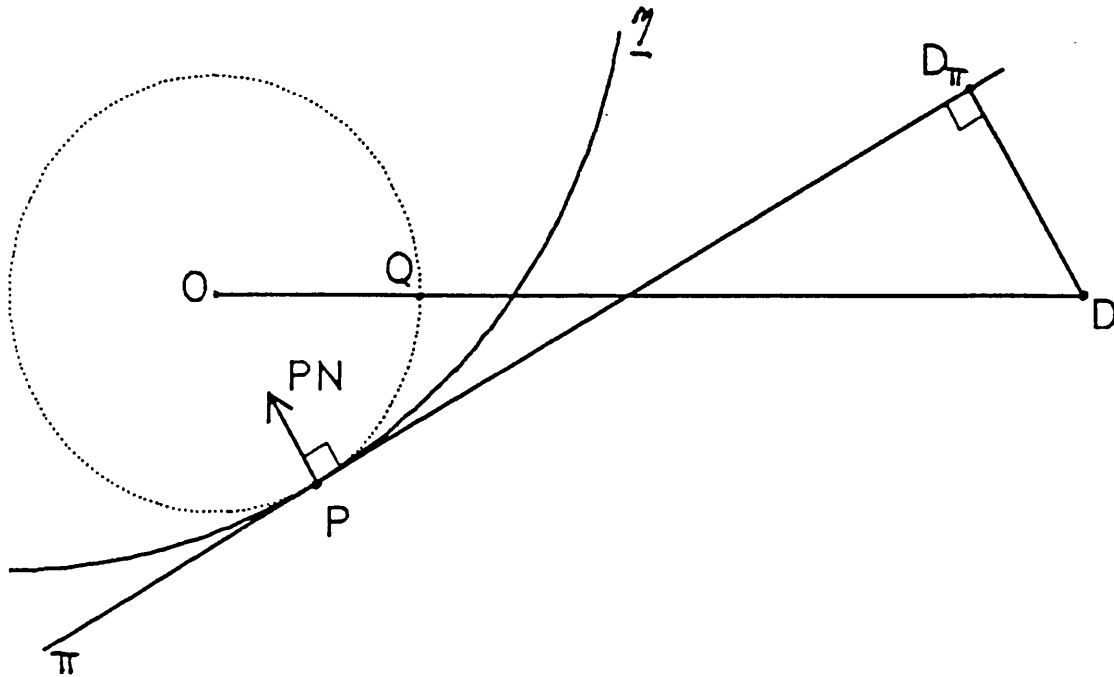
Figure 4.5.2:   Diagram   showing   possible   geometric
                approach   to   non-linear   parameter
                estimation in case where n=2, p=1

Ideally we would seek the point on the solution surface
lying on the line DO (and then repeat the procedure until
convergence achieved). For this approach, we need to know
the corresponding parameter value, which should be
possible here. Otherwise we could seek the point on π
lying on the line DO, and map to η corresponding to the
known updated parameter estimate (and then cycle). It is
readily seen that this new point on π is an improvement on
the previous value D_π (under perpendicular projection).
Alternatively, we may seek to make use of our knowledge of

the nearest point on the circle of curvature (Q) to our data point. A mapping from the circle to the solution surface is needed for this, via π is necessary. The need for a sensible mapping $\pi \rightarrow \underline{n}$ is apparent, a (near) orthogonal one being crucial in the region of the true least squares point.

The possibilities cited here may be summarised as follows (using an obvious shorthand notation):

(a)   DO $\cap$ $\underline{n}$

(b)   DO $\cap$ π ($\rightarrow \underline{n}$)

(c)   DO $\cap$ 0 = Q (($\rightarrow \pi$) $\rightarrow \underline{n}$)   .

Example 4.5.2   (n = 3, p = 1...space curve)

In this case, the principal normal (PN) and the tangent (line) π define a plane, π^* (say), through P. Let D_π^* be the projection of D onto π^* (see Figure 4.5.3). The following possibilities now present themselves.

Ideally we would seek the point on the solution surface lying in the plane DD_π^*0 (and then repeat the procedure). The same provisos as in Example 4.5.1 apply. Otherwise we would either seek the point on π lying on the line D_π^*0 and proceed as before; alternatively point Q may be utilised as before, yielding the following possibilities:
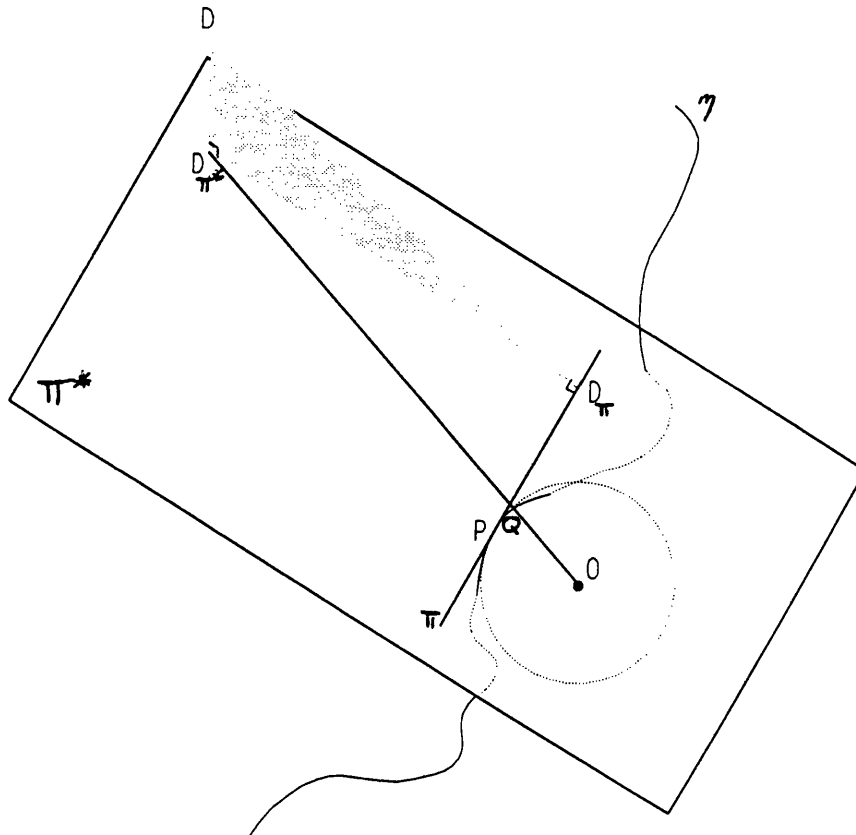
Figure 4.5.3: Diagram showing possible geometric approach to non-linear parameter estimation in the case where n=3, p=1

(a) $DD_\pi*O \cap \underline{n}$

(b) $D_\pi*O \cap \pi \ (\to \underline{n})$

(c) $D_\pi*O \cap \Theta = Q \ ((\to \pi) \to \underline{n})$.

Example 4.5.3 (n = 3, p = 2)

Numerous possibilities arise in this case. Suppose first that we were to fit the circle of curvature at P in the direction $\overrightarrow{D_\pi P}$ (see Figure 4.5.4, in which for simplicity $\underline{n}$ is not shown). Again we would first seek the point on the solution surface lying on the line DO, with analogous

198

alternatives to the previous examples viz:
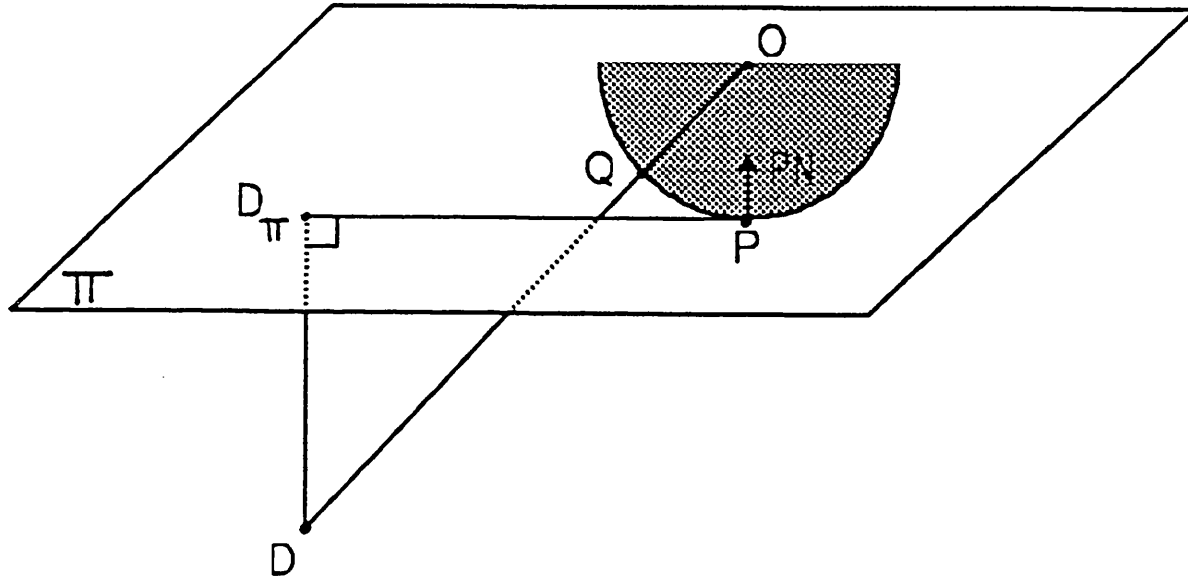
Figure 4.5.4:   Diagram showing possible geometric
                 approach to non-linear parameter
                 estimation in the case where n=3, p=2
                 (for simplicity solution surface is not
                 shown)

(a) DO ∩ $\underline{n}$

(b) DO ∩ π (→ $\underline{n}$)

(c) DO ∩ ⊙ ((→ π) → $\underline{n}$) .

However, we find that case (b) above gives a simple re-
scaling of the usual least squares parameter increment;
alternative procedures may therefore be sought which will
offer hope in cases where this direction is unpromising.
A couple of such procedures are now discussed.

The first possibility is to fit (2) circles of curvature at P in principal directions (determination of such directions may be found via e.g. [19] p. 357). D may then be projected onto the resultant planes of curvature π_1 and π_2 (say) and these points $D_{\pi 1}$ and $D_{\pi 2}$ may be joined to the corresponding centres of curvature O_1 and O_2. These lines will cross the tangent plane π along principal (orthogonal) axes centred at P and these points of intersection may be combined naturally to give the new approximation point on π. The analysis may now continue as before. An attempt at pictorially representing this procedure is given in Figure 4.5.5.

A minor modification of the above approach is to maintain the direction $\overrightarrow{D_\pi P}$ in the tangent plane and use it as one of a pair of conjugate directions. A similar procedure to that above for principal directions may be employed.

We shall now consider possible extensions to a general procedure based on the previous three examples. The first attempt would seem to be to project the data point D onto the approximating tangent plane π as before, and fit the appropriate circle of curvature in the direction $\overrightarrow{D_\pi P}$ lying in the plane through P spanned by this direction vector and the principal normal). We would then wish to solve

$$\pi' \cap \underline{\eta}$$

200
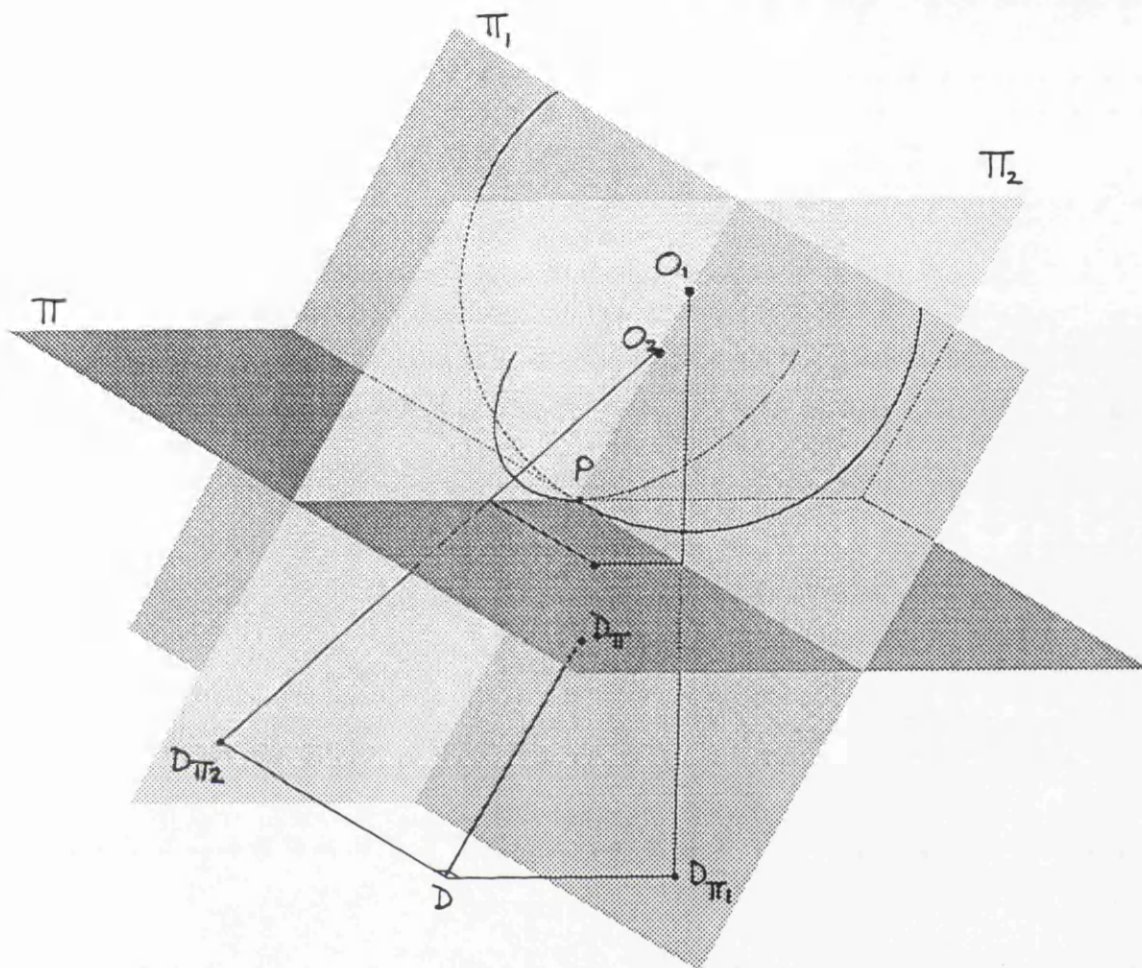
Figure 4.5.5:   Diagram showing an alternative approach
                  to non-linear parameter estimation in
                  the case where n=3, p=2 (for simplicity
                  solution surface is not shown)

where π' is an appropriate (n-p) dimensional space through

D and the resultant centre of curvature O.   This should

then give us a new approximating point on the solution

surface, corresponding to an updated parameter estimate.

The major drawbacks of such an approach, which although intuitively sound seems to be practically infeasible, may be listed as follows:

(1)   Determination of the complementary space π' (though some form of decomposition analogous to that for the acceleration vector $\underline{n}_h$ of Section 4.3 may prove useful)

(2)   Solutions of $\pi' \cap \underline{n}$

(3)   Mapping: $\underline{n} \rightarrow \underline{\theta}$ (i.e. given a new point on the solution surface, can we readily find its underlying parameter value?).

We must be careful, especially à propos (2) above, not to make the modified problem much more difficult than the original. Although such a direct approach may work in particular cases, recourse will generally need to be made to more practical methods such as the possibilities cited below.

In the light of our (illustrative) examples we may seek to reduce the general problem ($n > p \geq 1$) to a sequence of 2-dimensional problems. The obvious candidates involve first the fitting of p planes through P defined by the principal normal (PN) and p spanning vectors for π. The projection of D onto each plane in turn may be joined to the corresponding centre of curvature and we note where these lines cross the corresponding 'axes' in the tangent plane π (or we may project back from the corresponding 'Q'

if necessary). We then combine these p 'co-ordinates' to get a point on π (corresponding to a known parameter increment) and proceed as before. The case where the spanning set forms an orthonormal basis for π is outlined below:

(1)   Fit p (orthogonal) planes through P defined by (unit vector) PN and the p spanning vectors $U_1, ..U_p$ (say)

(2)   Find projection of D onto each plane ($D_{\pi j}$ j=1..p)

(3)   Find centres of curvature (O_j j=1..p)

(4)   Find $D_{\pi j}$ O_j $\cap$ π (P + λ_j U_j (say)   j=1..p)

(5)   Find parameter value corresponding to point

$$P + \sum_{j=1}^{p} \lambda_j U_j \quad (\text{on } \pi)$$

(6)   Repeat procedure until convergence achieved.

Other obvious candidates for the spanning set to use are principal directions and conjugate directions (using $D_\pi P$ as one axis), the underlying approach being exactly the same. Finally, it should be noted that in cases where the $D_{\pi j}$ lie above the corresponding axis they should be projected onto the axis either as before or via Q_j (not via the intersection with $D_{\pi j}$ O_j) as illustrated in Figure 4.5.6.
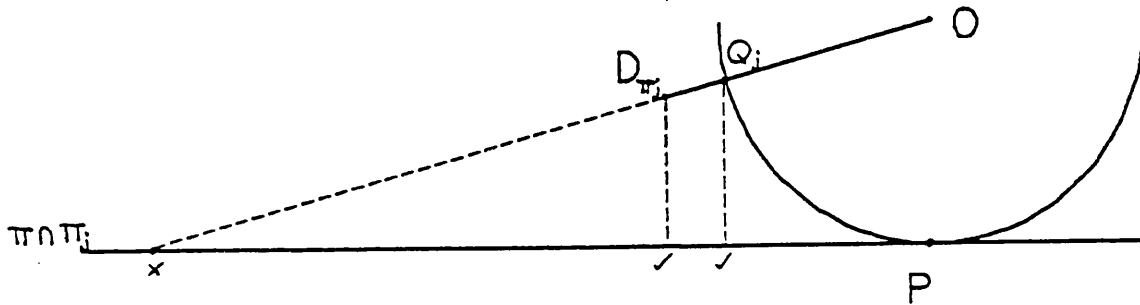
Figure 4.5.6:   Possible projections arising when $D_{\pi j}$
lies above the axis $(\pi \cap \pi_j)$

Another method conducive to this type of approach is that
of steepest descent in which we have the direction, $\underline{h}$
(say), in which to adjust the parameter estimate $\underline{\theta}_0$ but
not the appropriate step length; such a situation may be
accommodated as follows.   Simply fit the circle of
curvature at P in the direction $\underline{h}$ and join the projection
of D on the corresponding plane of curvature to its
centre;   then   pick   off   the   step   length   from   the
intersection of this line with the axis $\underline{h}$ (or from the
projection of 'Q' onto $\underline{h}$).   This should yield a nearly
optimal step length...for absolute optimality refinement
procedures (e.g. bisection) will still need to be employed
in this region.   The consequent usefulness of this
approach does therefore seem limited.   On the related

topic of re-scaling our normal least squares parameter increment, however, (as in Example 4.5.3.(b)), we note that (with data point D below π) our modified procedure will always reduce the step length. In the crystallographic setting, this will help correct for the over-estimation found in crystal structure analyses, as discussed in Section 3.5.

Thus far we have only considered sequences of 2-dimensional problems based on fitting circles in various planes. Suppose, however, we wished to best fit the solution surface η based on the information provided by various radii of curvature. One natural way of doing this is to approximate η by a principal ellipsoid i.e. an ellipsoid that fits the solution surface at P in the sense that its radii of curvature match those calculated along principal directions. Let the principal axes be denoted by $x_1..x_p$ and the principal normal by x_{p+1}. Then the resultant ellipsoid has equation

$$\sum_{j=1}^{p+1} \frac{x_j^2}{a_j^2} = 1 \qquad (4.5.1)$$

where P represents the point $(0....0, -a_{p+1})$ and

$$\frac{a_j^2}{a_{p+1}} = r_j \qquad (j=1..p) \quad ,$$

r_j being the radius of curvature along the x_j axis. We set a_{p+1} = min $(r_1..r_p)$. We assume the projection of our data point D = $(d_1....d_n)$ onto the plane, π^*, spanned by $x_1....x_{p+1}$ has co-ordinates D_{π^*} = $(d_1^*....d_{p+1}^*)$ with $d_{p+1}^* < -a_{p+1}$. Finding the nearest point (Q) on the ellipsoid (4.5.1) to our data point D now corresponds to minimising

$$\sum_{j=1}^{p+1} (x_j - d_j^*)^2 \qquad (4.5.2)$$

subject to (4.5.1). The method of Lagrange's undetermined multipliers yields the solution

$$x_j = \frac{d_j^*}{1 + \lambda/a_j^2} \qquad (j=1...p+1) \qquad (4.5.3)$$

where λ satisfies

$$\sum_{j=1}^{p+1} \left[\frac{d_j^*}{a_j(1 + \lambda/a_j^2)} \right]^2 = 1 \quad . \qquad (4.5.4)$$

We now have a 1-dimensional problem (solution of (4.5.4) for λ) which should not prove too intractable. Having solved this, we will still, however, need a suitable map from ellipsoid to parameter space (hence $\underline{n}$). If we do this via projection of Q onto π, perpendicular projection from π to $\underline{n}$ will return to us a point on $\underline{n}$ in the vicinity

of Q as required.

The potential problem of determining λ is somewhat alleviated if the ellipsoid becomes a (p-dimensional) sphere. The problem still remains, however, of finding the appropriate radius of the sphere. Two candidates immediately present themselves:

(1) The radius of curvature in the key direction $\overrightarrow{D\pi P}$

(2) The minimum radius of curvature in any direction. Option (2), for which convergence may be (very) slow is a conservative procedure. Option (1) is likely to be of more use. It should only be implemented, however, if the discrepancy between the maximum and minimum (principal) radii of curvature is small enough to permit such an approximation. As before, Q may now be found either by the intersection of the (n-p) -dimensional complementary space π' with the sphere or trivially via the join of $D_\pi*$ to its centre O.

Alternatively we may wish to fit the 'principal parabaloid' with equation

$$\sum_{j=1}^{p} \frac{x_j^2}{a_j^2} = 2x_{p+1} \qquad (4.5.5)$$

where P now represents $(0,\ldots 0,0)$ and $a_j^2 = r_j$ $(j=1..p)$. The method of Lagrange's undetermined multipliers now yields the solution

$$x_j = \frac{d_j^*}{1 + \lambda/a_j^2} \quad (j=1\ldots p), \quad x_{p+1} = d_{p+1}^* + \lambda \qquad (4.5.6)$$

where λ now satisfies

$$\sum_{j=1}^{p} \left(\frac{d_j^*}{a_j(1+ \lambda/a_j^2)}\right)^2 = 2(d_{p+1}^* + \lambda) . \qquad (4.5.7)$$

Other approximating surfaces based on known radii of curvature may of course be used. We would expect all to lead to improvements over methods based on the tangent plane approximation alone, which does not utilise this extra information on the curvature of the solution surface at the current approximation point.

An example of our ideas involving the fitting of an appropriate ellipsoid appears in Section 4.7.

4.6   Extension of the Ideas of Sections 4.3 and 4.4 with Particular Emphasis on the Crystallographic Model

The parameter effects curvature array defined by equation (4.3.15) for the model defined by equation (4.3.1) with p=1 reduces to the (1x1x1) array

$$A_{..} = \frac{1}{A^3} \sum_{i=1}^{n} \frac{\partial \eta_i}{\partial \theta_1}\bigg|_{\underline{\theta}_0} \frac{\partial^2 \eta_i}{\partial \theta_1^2}\bigg|_{\underline{\theta}_0} \qquad (4.6.1)$$

where

$$A = \left[\sum_{i=1}^{n} \left(\frac{\partial \eta_i}{\partial \theta_1} \bigg|_{\underline{\theta}_0} \right)^2 \right]^{\frac{1}{2}} . \qquad (4.6.2)$$

By using the re-parameterisation

$$\phi_1 = \sum_{i=1}^{n} \eta_i(\underline{\theta}) \frac{\partial \eta_i}{\partial \theta_1} \bigg|_{\underline{\theta}_0} \qquad (4.6.3)$$

motivated by the analysis of Section 4.4, we find that the (1x1) matrices G. and G.. of Section 4.3 become

$$G. = \sum_{i=1}^{n} \frac{\partial \eta_i}{\partial \theta_1} \frac{\partial \eta_i}{\partial \theta_1} \bigg|_{\underline{\theta}_0} \qquad (4.6.4)$$

and

$$G.. = \sum_{i=1}^{n} \frac{\partial^2 \eta_i}{\partial \theta_1^2} \frac{\partial \eta_i}{\partial \theta_1} \bigg|_{\underline{\theta}_0} . \qquad (4.6.5)$$

The new parameter effects array is given by equation (4.3.23) as

$$\tilde{A} = A.. - \frac{1}{A} \frac{G..}{G.} \qquad (4.6.6)$$

which reduces to zero at $\underline{\theta} = \underline{\theta}_0$.

Similarly for the 2-parameter case in which $\underline{\theta} = (\theta_1, \theta_2)$, the re-parameterisations suggested by Section 4.4 produce

zero parameter effects at $\underline{\theta} = \underline{\theta}_0$. We next consider the analogous results for the 3-parameter case $\underline{\theta}^T = (\theta_1, \theta_2, \theta_3)$, of which the model function

$$f(\underline{x}, \underline{\theta}) = \cos 2\pi \ (x_1 \ \theta_1 + x_2 \ \theta_2 + x_3 \ \theta_3) \qquad (4.6.7)$$

is an example. We shall then wish to extend our theory to cover linear combinations of distinct model functions, in order to give us some insight into how to handle the non-linear models appearing in crystallographic studies. The notation used is that of the previous Sections.

The parameter effects curvature array defined by equation (4.3.15) for the model defined by equation (4.3.1) with p=3 is the (3x3x3) array given in Appendix 2. Equation (4.4.15) leads us to consider the re-parameterisation

$$\phi_j = \sum_{k=1}^{3} \ (\gamma_{jk} \sum_{i=1}^{n} \ \eta_i(\underline{\theta}) \ v_{ki}) \quad (j=1,2,3) \qquad (4.6.8)$$

where the coefficients γ_{jk} may also be found in Appendix 2. Evaluated at $\underline{\theta}_0$, the (3x3) matrix $\underset{\sim}{G}.$ is now proportional to the (3x3) identity matrix, and the (3x3x3) array $\underset{\sim}{G}..$ has elements β_{jkl} where

$$\beta_{jkl} = \sum_{m=1}^{3} \ (\gamma_{jm} \sum_{i=1}^{n} \ v_{mi} \ \frac{\partial^2 \eta_i}{\partial \theta_k \partial \theta_l} \Bigg|_{\underline{\theta}_0}) \quad (j,k,l=1,2,3). \quad (4.6.9)$$

It may be checked that the new parameter effects array

reduces to zero. Once again, the reduction of the parameter effects curvature array is seen to quantitatively vindicate our notion of perpendicular projection. Before returning to our specific crystallographic problem, we shall now consider sums of model functions as discussed below.

Suppose our model function splits into the form

$$f_1(\underline{x}, \underline{\theta}_1) + f_2(\underline{x}, \underline{\theta}_2) \qquad\qquad (4.6.10)$$

where

$$\underline{\theta}_j = (\theta_{j1} \cdots \theta_{jp_j})^T \qquad (j=1,2),$$

with $\underline{\theta}_1$ and $\underline{\theta}_2$ containing no common parameters. Suppose further that we have available re-parameterisations $\underline{\phi}_j$ for the individual model functions $f_j(\underline{x}, \underline{\theta}_j)$ $(j=1,2)$, chosen to ensure individual zero parameter effects curvature arrays as discussed above. By letting $\underline{U}_{.j}$ represent the first p_j columns of $\underline{Q}$ in the '$\underline{Q}$ $\underline{R}$' decomposition of Section 4.3 carried through for the model function $f_j(\underline{x}, \underline{\theta}_j)$ $(j=1,2)$, the re-parameterisations $\underline{\phi}_1$ and $\underline{\phi}_2$ lead to (Appendix 2) a reduced parameter effects curvature array for our model function (4.6.10) if $\underline{U}_{.1}^T \underline{U}_{.2} = 0$. The argument may be extended to cover the sum of arbitrarily many model function involving disjoint subsets of parameters, a sufficient condition for the individual re-parameterisations being effective being that

$$\underset{\sim}{U}._j{}^T \ \underset{\sim}{U}._k \ = \ \underset{\sim}{0} \qquad (j \neq k) \ . \qquad\qquad (4.6.11)$$

Falling into this framework will be the model functions appearing in crystallography, namely (neglecting temperature factors) those of the form

$$f(\underline{x},\underline{\theta}) \ = \ \sum_{r=1}^{R} \cos 2\pi \ (x_1\theta_{1r} + x_2\theta_{2r} + x_3\theta_{3r}) \qquad (4.6.12)$$

for which condition (4.6.11) yields

$$\sum_{i=1}^{n} m_i m_i{}' \ \sin 2\pi \ (x_{1i}\theta_{1r0} + x_{2i}\theta_{2r0} + x_{3i}\ \theta_{3r0})$$

$$.\sin 2\pi \ (x_{1i}\theta_{1s0} + x_{2i}\ \theta_{2s0} + x_{3i}\ \theta_{3s0})$$

$$= 0 \ (r,s = 1..R)$$
$$(m,m'= x_1,\ x_2,\ x_3)$$

$$\qquad\qquad\qquad\qquad (4.6.13)$$

where θ_{1r0} is our present estimate of the parameter θ_{1r} etc. Equations (4.6.13) are not unreasonable in practice; after all they form the basis of the block diagonal approximations discussed in Section 3.5. Thus the individual re-parameterisations given by equation (4.6.8) for the model functions appearing in the summation term of equation (4.6.12) will be useful for crystallographic purposes, without at the same time being too unwieldy.

4.7 A Summary of Results of the Applications of Some of the Ideas from Sections 4.4 And 4.5

Throughout this Section we shall be considering the model (4.3.1) in which the model function

$$f(\underline{x},\underline{\theta}) = \eta(\underline{x},\underline{\theta}) = x_4 \cos 2\pi(x_1\theta_1 + x_2\theta_2 + x_3\theta_3), \qquad (4.7.1)$$

and shall be comparing our ideas against the standard least squares approach to non-linear parameter estimation. For the basis of our comparison we shall be using the parameter estimates at successive stages of the iterative procedures, together with the associated residual sum of squares (which we are trying to minimise),; the residual sum of squares shall henceforth be abbreviated to RSS.

The theory of Section 4.4 was primarily concerned with those cases in which the tangent plane approximation was reasonable, though the parameter effects curvature was high. By using the re-parameterisation suggested by equation (4.4.15) in conjunction with equation (4.4.23), thus enabling us to calculate the appropriate parameter increment, we would expect our ideas to be superior in such cases. This turned out to be the case in some of the limited examples studied, though similar results were obtained in general. Furthermore, in practice it was found that the high parameter effects curvature was present usually only when high intrinsic curvature also

213

pertained. In view of this, successful implementation of the techniques of Section 4.5 was deemed to be of more importance in the crystallographic setting.

Henceforth we shall be primarily concerned with data points on the convex side of the solution locus (η) sufficiently far from our initial approximating point. In a similar vein to Example 4.5.3 (b), we may seek to re-scale the usual least squares parameter increment based on fitting the appropriate circle of curvature in the least squares direction. We shall actually consider here the standard increment suggested by the projection of Q onto the tangent plane π (see Figure 4.5.4). Alternatively, we may wish to test some of the ideas at the end of Section 4.5 based on replacing the approximating tangent plane by an appropriate ellipsoid. In practice, instead of the principal ellipsoids cited previously, we shall content ourselves by fitting ellipsoids consistent with the (intrinsic) radii of curvature in the orthogonal directions determined by the '$\underset{\sim\sim}{QR}$' decomposition of Section 4.3. Such ideas are likely to prove particularly helpful in cases where the intrinsic curvature is relatively high - so that the usual tangent plane approximation is not very good. For illustrative purposes, we consider the following example, which typifies our findings.

We set n = 10 and suppose we have the data point

$$D = (-2,3,3,-7,8,4,-2,-9,5,0)$$

obtained for the following experimental settings:

$$\underline{x}_1^T = (1,2,3,4,5,6,7,8,9,10)$$

$$\underline{x}_2^T = (1,0,2,0,1,1,1,2,2,2)$$

$$\underline{x}_3^T = (4,3,2,1,4,3,2,1,2,3)$$

$$\underline{x}_4^T = (2,4,6,8,2,4,6,8,6,4) \quad .$$

We take as our initial parameter value

$$\underline{\theta}_0^T = (.3515,.7941,.0895) \ ,$$

which gives rise to an initial RSS of 67.40. A single application of the standard non-linear least squares method produces the parameter estimate (to 4 d.p.)

$$\underline{\hat{\theta}}_1^T \ (\text{say}) \ = \ (.3211,.8084,.1355),$$

which gives rise to an increased associated RSS, namely 88.08. The reason for this increase is the high intrinsic curvature: the radius of curvature of the solution locus in the least squares direction alone is found to be only 0.90. However, by fitting the appropriate circle of curvature in this direction as suggested above, our re-scaled increment gives the parameter estimate (to 4 d.p.)

$$\underline{\hat{\theta}}_2^T \ (\text{say}) \ = \ (.3464,.7965,.0971)$$

with an RSS of 62.24.

215

The alternative method based on fitting the ellipsoid yielded the parameter estimate (to 4 d.p.)

$$\hat{\underline{\theta}}_3^T \text{ (say)} = (.3405, .8130, .0946),$$

with an even smaller associated RSS of 56.23. Our results are summarised in Table 4.7.1, together with the converged estimate $\hat{\underline{\theta}}_4^T$ (say) (to 4 d.p.) and its associated RSS.

		RSS
$\underline{\theta}_0^T$:	(.3515, .7942, .0895)	67.40
$\hat{\underline{\theta}}_1^T$:	(.3211, .8084, .1355)	88.08
$\hat{\underline{\theta}}_2^T$:	(.3464, .7965, .0971)	62.24
$\hat{\underline{\theta}}_3^T$:	(.3405, .8130, .0946)	56.23
$\hat{\underline{\theta}}_4^T$:	(.3314, .8445, .0953)	53.43

Table 4.7.1:   Summary of the results of the example of this Section, for which the radius of curvature of the solution locus in the least squares direction is 0.90

As previously stated, the results above are typical of what we might expect in cases of high intrinsic curvature.

In such cases, the subsequent inapplicability of the standard least squares method may be overcome by resort to methods such as those proposed here. A cautionary note is in order, however: even these methods are not guaranteed to converge, and we may additionally require low parameter effects curvature for any mappings $\pi \rightarrow \underline{\eta}$ to be acceptable over the range of valid approximation. Our initial findings of success dictate that further such analysis of the properties of our schemes should be undertaken.

CHAPTER 5

EXAMPLES OF STATISTICAL TECHNIQUES

TO SEQUENTIAL CRYSTAL STRUCTURE DETERMINATIONS

5.1 Introduction

In this Chapter, we shall be considering the application
of some of our ideas from Chapter 3 to sequential crystal
structure determination. The example of Section 5.2 sees
us comparing a weighted least squares approach based on
(3.3.45) against alternative (Bayesian) competitors. The
example of Section 5.3 sees the implementation of the D-
optimality based ideas of Section 3.6. Details of the
structure under study - namely, anthracene - are given in
Section 5.2, together with a summary of the prior beliefs
we shall be using. Details of the specific sequential
procedures invoked will also be found in the relevant
sections, together with a summary of the results obtained.

5.2 Application of the Weighting Scheme (3.3.45) in the Structure Determination of Anthracene by Weighted Least Squares and a Comparison with Various Bayesian Methods of Structure Determination

In this Section we shall be investigating the structure of
anthracene, a product in the distillation of coal-tar
used in the manufacture of dyes.   Its unit cell is
monoclinic, that is the axes a,b,c are of unequal length
(namely 8.561, 6.036, 11.163 Å respectively) and the only

non-orthogonal relationship amongst them is that the angle between axes a and c is 124.7°.

The asymmetric unit is known to consist of 7 carbon atoms and 5 hydrogen atoms, and the symmetry relationships are as follows: for every atom at (x,y,z) there are equivalent atoms not only at $(-x,-y,-z)$ but also at both $(x+\frac{1}{2},y,z)$ and $(x,\frac{1}{2}-y,z)$. We shall have available the 240 observed structure factors as provided by the proportional counter data of Phillips used in [61], together with their associated measurement errors based on the corresponding $(\sin \theta)/\lambda$ values. The parameters we shall wish to estimate are hierarchical: of primary concern will be the (3x7) carbon atomic co-ordinates, though we shall also be estimating the (3x5) hydrogen atomic co-ordinates and (7) carbon isotropic thermal parameters (the corresponding (5) hydrogen isotropic thermal parameters being fixed). For scaling purposes, an overall scale factor will also be estimated.

The purpose of the analysis of this Section is to compare a structure determination using a weighted least squares approach based on (3.3.45) with various Bayesian procedures, including an existing scheme used in the Crystallography Unit at University College London. More shall now be said about this latter scheme which is a variant of our Bayesian approach outlined at the start of Section 3.3.

The main difference arises from the fact that the model

used is now

$$FO*(\underline{h})/FT = ft \sim N(ft, \sigma_0*^2) \qquad (5.2.1)$$

i.e. we assume the structure amplitude given the sign of the current fitted structure factor is normally distributed about the true structure factor with variance σ_0*^2 (say). The analogue of equation (3.3.4) becomes the linearised model

$$\underline{y}^{(0)}/\underline{\theta} \sim N(\underline{X}^{(0)}\underline{\Delta\theta}^{(0)}, \underline{V}*) \qquad (5.2.2)$$

where $\underline{X}^{(0)}$, $\underline{y}^{(0)}$ are defined by equations (1.5.4), (1.5.5) and $\underline{V}*$ is taken to be the diagonal matrix with elements σ_0*^2. Analysis goes through as before resulting in the equations (3.3.5) - (3.3.7), but now with $\underline{W} = [\underline{V}*]^{-1}$. With this important proviso in mind, the approach proceeds similar to before. [A discussion of the varying approaches to structure determination is deferred to later in the Section]. The formula used in the existing programme for the variances σ_0*^2 shall now be given.

We define by s^2 the quantity

$$s^2 = \underset{\underline{h}}{\Sigma} \frac{(FO*(\underline{h}) - FC(\underline{h}))^2}{n} \qquad (5.2.3)$$

where the summation is evaluated from the previous stage of our analysis (incorporating n observations). Initially, crystallographers must specify an appropriate s^2 value. We may similarly evaluate/specify the

corresponding R-value as defined by equation (4.2.16) at the previous stage. We then define

$$\sigma_0^{*2} = \sigma_0^2(\underline{h}) + \frac{s^2}{10} + \frac{(FO*(\underline{h}) - FC(\underline{h}))^2}{50}$$

$$+ (2FO*(\underline{h}))^2 \exp(2s \ R \ \frac{\sin \theta}{\lambda} \ (\underline{h})) \qquad (5.2.4)$$

where $\sigma_0^2(\underline{h})$ is the associated error variance for our measured structure amplitude. The theory behind such a scheme is briefly discussed in [58].

Although the theory proposed above is based on a different model to that discussed in Section 3.3, our ideas there centred on an improved estimate of the variances of the appropriately signed structure factors. We shall therefore consider the results obtained using the above approach in conjunction with

$$\sigma_0^{*2} = (FO*(\underline{h}) - \mu_1)^2 + \sigma_1^2 \qquad (5.2.5)$$

(where μ_1, σ_1^2 are defined by equations (3.3.43), (3.3.44)). For completeness, we shall also consider the results obtained from the Bayesian approach using our original model (3.3.30). In all three cases, the chemical information available - to be presently introduced - will be treated as observational data also. This will also be the case in our weighted least squares approach, with weights based on (3.3.45), as outlined in Section 3.3. We shall now give further details of the analysis and methods of comparison to be used.

Our initial batch of data shall comprise all data out to an S-limit of .2 (where $S = (\sin \theta)/\lambda$). The S-limit for successive cycles will be determined by the formula [58]

$$S = \frac{.2}{R+R(-1)} \qquad (5.2.6)$$

where R is the R-value just evaluated and R(-1) the R-value at the previous stage (initially set equal to .5). This is very much an ad hoc rule in keeping with the concluding comments of Section 3.2 and has been found to work well in practice. Furthermore, the initial values (although not consistent)

$$\frac{s^2}{10} = 3, \qquad s = 6$$

are used in the denominator of the right hand side of (5.2.4).

We shall compare the differing approaches by the traditional method of tabulating n and R values at the various stages, and by also giving the corresponding weighted R - values defined by

$$WR = \frac{\sum_{h} W_{\underline{h}} (FO*(\underline{h}) - FC(\underline{h}))^2}{\sum_{h} W_{\underline{h}} FO*(\underline{h})^2}$$

- where the $W_{\underline{h}}$ are the diagonal elements of the matrix $\underset{\sim}{W}$ - though we should note that the Bayesian approach doesn't purport to minimise any such criterion function. We shall

222

also list $\bar{\sigma}^2$ values, the average values of the variances of our estimates for the co-ordinates of the carbon atoms (see later discussion). The variances quoted (in units of $\mathring{A}^2$) will refer to positions in real space. We note that such variances may in fact be underestimates due to the various approximations involved, in particular those accounting for the non-linearity and any phase uncertainty. Care must consequently be taken with their interpretation. Finally, we shall give the refined (fractional) co-ordinate estimates for the 7 carbon atoms, which will serve as our best means of comparison.

For our trial structure we make use of the fact that the anthracene molecule is known to be approximately planar and of the form of Figure 5.2.1, where the numbers 1-7 represent the carbon atoms and 8-12 the hydrogen atoms. We are not, however, sure of the relation of the molecular plane to the crystal axes and unit cell and consequently impose prior beliefs for the atomic co-ordinates of the form of equation (3.2.12). The means of our prior distributions to be used for this example are summarised in Table 5.2.1, together with the analogous prior beliefs about the overall scale factor and isotropic thermal parameters. All prior variances for the atomic co-ordinates are set equal to .0025 with the remaining prior variances set equal to .01.

Our prior chemical information consists of the specification of 19 bond lengths and of 27 direction
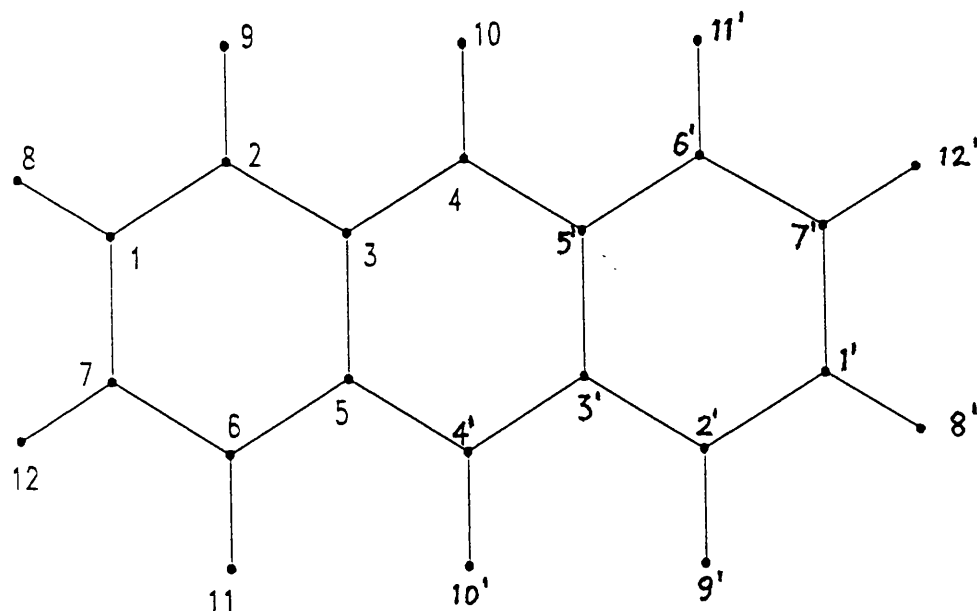
Figure 5.2.1: Basic (planar) structure of the anthracene
molecule.   Numbers   1',2'   etc   denote
symmetry related atoms

cosines as summarised in Table 5.2.2: all prior means for
direction cosines are 0 and the corresponding collinear
pairs of atoms are listed. All prior variances are .0001,
with the corresponding prior variances on bond lengths
being .001.   We note that such information does not
involve the isotropic thermal parameters.   These
parameters are furthermore not well-determined by the
diffraction data for the low angle observations that we
initially incorporate into our analysis.   The weighted
least squares approach therefore fixes these parameters in
the early stages of the refinement procedure by applying
a partial shift rule for them of the form of (3.5.20) with
$\eta=0$.   [The problem does not arise in the Bayesian case,
since the shifts are constrained by the initial variances
.01].

Parameter	Prior mean	Parameter	Prior mean
S	1.2	x_8	.08108
		y_8	-.01410
x_1	.04784	z_8	.45900
y_1	-.07001	x_9	.17210
z_1	.34979	y_9	.25000
x_2	.09298	z_9	.33960
y_2	.08680	x_{10}	.15720
z_2	.28476	y_{10}	.36375
x_3	.04600	z_{10}	.12510
y_3	.04505	x_{11}	-.14967
z_3	.14056	y_{11}	-.47907
x_4	.08952	z_{11}	-.08527
y_4	.20118	x_{12}	-.07525
z_4	.06967	y_{12}	-.43433
x_5	-.04485	z_{12}	.32271
y_5	-.15787		
z_5	.06965	B_1	4.5
x_6	-.08924	B_2	4.5
y_6	-.31429	B_3	4.6
z_6	.14309	B_4	4.7
x_7	-.04429	B_5	4.8
y_7	-.26914	B_6	4.9
z_7	.27846	B_7	5.0

Table 5.2.1: Summary of prior parameter beliefs for the anthracene example of this Section. Parameter x_r denotes the x co-ordinate of the rth atom etc; similarly B_r the isotropic thermal parameter. S represents the overall scale factor

Pair of bonded atoms		Prior mean (Å)	Collinear pairs of atoms				
1	2	1.40	1	3 ,		2	4
2	3	1.40	5	6 ,		4	7
3	4	1.40	1	2 ,		7	4
3	5	1.40	3	4 ,		7	3
5	6	1.40	1	5 ,		8	1
6	7	1.40	2	6 ,		9	2
1	7	1.40	2	6 ,	10		4
1	8	1.09	2	6 ,		6	11
2	9	1.09	3	7 ,		7	12
4	10	1.09					
6	11	1.09					
7	12	1.09					
1	3	2.42					
2	4	2.42					
4	5	2.42					
5	7	2.42					
1	5	2.80					
2	6	2.80					
3	7	2.80					

Table 5.2.2:   Summary of prior chemical information for the anthracene example of this Section

The results of our analyses are summarised in Tables 5.2.3 - 5.2.7, the latter Table comparing the resultant final estimates. As commented after (3.3.45), we maintain the original data throughout, though fixed matrices $\underset{\sim}{W}$ are used once all 240 observations are incorporated into the analyses. We shall now briefly comment on the results of our analyses.

Direct comparison of Bayesian and least squares approaches to crystal structure determination is difficult due to the differing underlying rationales - see Section 4.2 - and consequent difference in interpretation of the results. In particular, the $\bar{\sigma}^2$ values of this Section are derived from the covariance matrix of the posterior distribution for the parameters in the former case and from the covariance matrix of the sampling distribution of the least squares estimate in the latter.

We shall consider first the results obtained under the Bayesian-type approaches as summarised by Tables 5.2.3 - 5.2.5. The rationale behind Table 5.2.5 - based on the model (3.3.30) and prior information of the form of (3.2.12) - is sound as far as it goes. However, model (3.3.30) assumes no sign information; by ignoring available such information, this approach is likely to prove sub-optimal, especially in cases (unlike here) in which the initial phase discrimination is poor. The rationale behind Tables 5.2.3 and 5.2.4 attempts to correct for this by using the model (5.2.1),

227

Cycle	n	R	WR	$\bar{\sigma}^2$ (Å^2 x 10^{-5})
1	17	.550	.578	(78100)
2	15	.328	.387	202
3	26	.429	.429	319
4	48	.339	.330	111
5	47	.280	.278	101
6	86	.197	.204	85.0
7	160	.137	.158	35.7
8	240	.098	.128	13.8
9	240	.075	.100	7.28
10	240	.072	.097	5.94
11	240	.071	.095	5.56
12	240	.071	.095	5.52
13	240	.071	.095	5.51

Table 5.2.3:   Results using Bayesian approach based on (5.2.2) and (5.2.4) for the anthracene example of this Section

Cycle	n	R	WR	$\bar{\sigma}^2 (\text{Å}^2 \times 10^{-5})$
1	17	.550	.953	(78100)
2	15	.400	.568	325
3	19	.201	.165	36.6
4	93	.150	.115	36.2
5	240	.099	.031	2.04
6	240	.084	.027	1.28
7	240	.083	.027	1.23
8	240	.083	.027	1.22

Table 5.2.4:   Results using Bayesian approach based on (5.2.2) and (5.2.5) for the anthracene example of this Section

Cycle	n	R	WR	$\bar{\sigma}^2$(Å^2 x 10^{-5})
1	17	.550	.765	(78100)
2	15	.510	.544	15.9
3	15	.331	.411	35.4
4	30	.348	.365	19.6
5	69	.161	.167	5.03
6	135	.107	.097	2.29
7	240	.077	.073	1.92
8	240	.070	.069	.957
9	240	.067	.067	.939
10	240	.066	.066	.938
11	240	.066	.066	.937

Table 5.2.5:   Results using the Bayesian approach based on (3.3.30) for the anthracene example of this Section

Cycle	n	R	WR	$\bar{\sigma}^2$ ($\mathring{A}^2 \times 10^{-5}$)
1	17	.550	.953	(–)
2	15	.475	.493	497
3	15	.315	.313	65.9
4	40	.162	.126	32.3
5	160	.136	.078	5.00
6	240	.068	.032	2.21
7	240	.061	.030	1.80
8	240	.061	.029	1.78
9	240	.061	.029	1.78

Table 5.2.6: Results using weighted least squares approach based on (3.3.45) for the anthracene example of this Section

Parameter	Table 5.2.3		Table 5.2.4		Table 5.2.5		Table 5.2.6	
	Estimate	s.d. ($\times 10^{-4}$)	Estimate	s.d. ($\times 10^{-4}$)	Estimate	s.d. ($\times 10^{-4}$)	Estimate	s.d. ($\times 10^{-4}$)
x_1	.08717	8.5	.08101	5.1	.08677	3.7	.08640	5.0
y_1	.02982	10.7	.02811	4.0	.02477	4.4	.02832	5.6
z_1	.36579	7.2	.36229	3.5	.36543	3.0	.36572	4.2
x_2	.11699	8.6	.11360	4.5	.11530	3.8	.11659	5.5
y_2	.15481	10.8	.15825	4.4	.15394	4.2	.15457	5.7
z_2	.27934	6.7	.27778	3.0	.27789	2.9	.28009	4.0
x_3	.05869	8.5	.05833	4.7	.05768	3.4	.05859	4.8
y_3	.08202	11.4	.07990	4.8	.08478	4.5	.08266	6.6
z_3	.13797	7.0	.13710	3.4	.13643	2.7	.13646	3.8
x_4	.08827	7.5	.08965	4.5	.08785	3.2	.08748	4.4
y_4	.20895	10.7	.20895	4.0	.21046	4.6	.21039	6.7
z_4	.04747	6.5	.04882	3.1	.04794	2.7	.04678	3.7
x_5	-.02883	9.0	-.02801	5.2	-.03043	4.0	-.02938	5.7
y_5	-.12741	11.3	-.13078	5.0	-.12668	4.8	-.12883	7.1
z_5	.09063	6.9	.09119	3.1	.08929	2.9	.09031	3.8
x_6	-.06055	9.2	-.06110	4.8	-.05975	3.8	-.06119	4.9
y_6	-.25943	11.1	-.26083	4.2	-.26017	4.6	-.26113	6.6
z_6	.18241	7.5	.18216	3.1	.18393	3.0	.18184	3.8
x_7	-.00333	10.1	-.00797	5.5	-.00128	4.1	-.00474	5.5
y_7	-.17731	12.5	-.18189	4.6	-.18194	4.7	-.18032	6.9
z_7	.31636	7.7	.31394	3.5	.31800	3.0	.31573	3.9

Table 5.2.7: Final estimates of the (fractional) carbon atomic co-ordinates for the anthracene example of this Section for each of the refinements summarised by Tables 5.2.3 - 5.2.6. Note that the associated standard deviations (s.d.) refer to fractional co-ordinates also

though implementation of the resultant schemes raises the following additional problems. Determination of an appropriate covariance matrix ($\underset{\sim}{V}*$) is non-trivial and will almost certainly be influenced by our prior beliefs, thus raising the possibility of double-counting our prior information. Furthermore, by continuously updating this covariance matrix but maintaining the original data, Bayesian principals are being abused as discussed in Section 3.3. This may provide a degree of over-fitting whereby observed values close to their fitted values are given unduly small standard errors: this accounts for the spuriously low WR-factors of Table 5.2.4. Note, however, that this will not arise for the (diagonal) elements of $\underset{\sim}{V}*$ defined by (5.2.4), which are by definition bigger than the observational errors $\sigma_0^2(\underline{h})$ alone. Such error variances may be considered to be additive, with components due to observational error and sign uncertainty. This accounts for the larger $\bar{\sigma}^2$ values of Table 5.2.3, but is not, however, felt to be a very realistic scheme. Notice that in the limiting case in which the sign uncertainty is negligible, we should obtain the results of Table 5.2.5. Bringing the data sequentially into the analysis in appropriate shells ensures that the sign uncertainties at any stage are indeed small, so that the results of Tables 5.2.3 and 5.2.5 are comparable. Finally, direct comparison of Tables 5.2.3 and 5.2.4 clearly show that our ideas lead to a marked improvement on the existing procedure, in the

sense of quicker convergence and more realistic $\bar{\sigma}^2$ values.

We now turn our attention towards the weighted least squares approach behind Table 5.2.6. Our previous misgivings about maintaining the original data throughout apply in the evaluation of the dynamic weighting scheme. There are no such misgivings about double-counting the prior information, however, which becomes obsolete once the appropriate weights have been evaluated. The emphasis shifts instead to the observed data and finding those parameters that give the 'best' match between observed and fitted structure factors.

Despite the differing viewpoints, Tables 5.2.3 - 5.2.7 show that the various approaches of this Section yield similar results for the particular example discussed here. This leaves crystallographers free to choose between a Bayesian or least squares approach according to their own personal preference, though the fact that the refined R-value was smallest for the weighted least squares approach may tip the balance slightly in its favour. Whichever method they decide to use should be clearly stated. In practice, the choice will often be decided by the relative importance of the prior information available and the actual observational data obtained. We conclude, however, by claiming that in cases in which the initial phase discrimination is poor, our proposed weighted least squares approach is to be recommended on the grounds that it emphasises best those structure factors that are still

well determined, without placing undue emphasis on the trial structure. Further structure determinations are required, however, to support this claim.

5.3 Application of a D-optimality based Algorithm in the Structure Determination of Anthracene

In this Section we continue to consider the structure of anthracene. Using the same data and prior beliefs as in the previous Section, similar analysis to before yields Table 5.3.1 for a weighted least squares structure refinement with weights proportional to experimental errors alone. Such a scheme is in line with our original least squares theory and was used in the formulation of the theory of Section 3.6. Furthermore, since the experimental errors used with our data set increase regularly with $(\sin \theta)/\lambda$, it is not unreasonable to expect to accurately know the errors associated with any as yet unobserved structure amplitudes. Accordingly, we shall here assume such errors to be known. We stress that in the general case a reliable estimate will still be readily available as previously discussed.

Using Table 5.3.1 we see that by cycle 5 we have measured 74 reflections and have achieved a weighted R-factor of .123. Of more interest for our present purposes, however, are the associated average parameter variances for the carbon atoms, namely 5.96, 4.05, 7.21 and 5.74 ($\text{Å}^2 \times 10^{-5}$) for the x,y,z and overall co-ordinates respectively. We

Cycle	n	R	WR	$\bar{\sigma}_x^2$	$\bar{\sigma}_y^2$	$\bar{\sigma}_z^2$	$\bar{\sigma}^2$
1	17	.550	.765	(-)	(-)	(-)	(-)
2	15	.425	.503	22.1	15.5	42.1	26.6
3	19	.423	.435	57.3	14.5	49.4	40.4
4	30	.232	.260	19.0	7.12	21.0	15.7
5	74	.121	.123	5.96	4.05	7.21	5.74

Table 5.3.1:   Initial results using weighted least squares approach based on observational error alone for the anthracene example of Section 5.2. All variances are $\text{Å}^2 \times 10^{-5}$

may deem our refinement procedure to be sufficiently far advanced for the theory of Section 3.6 to apply and consider the question of which additional few observations to measure in order to best reduce these variances.

Suppose we wish to bring in an extra 6 (say) observations. The standard procedure would be to continue to bring in reflections in order of increasing $(\sin \theta)/\lambda$. However, the (diagonal approximation) theory of Section 3.6 would suggest that we should bring in those 6 observations that maximise a criterion function of the form of (3.6.25). Instead of using that approximate result, we may also base our choice on the criterion function (3.6.28). Although the D-optimality theory behind this latter scheme was based on minimising the determinant of the parameter covariance matrix, no diagonal approximation is assumed and the ensuing simultaneous reduction of atomic parameter variances should also reduce our average variances in a nearly optimal manner.   It is this D-optimality based approach that we shall pursue here.

The results of bringing in the extra 6 observations suggested by the traditional and D-optimal methods (and cycling to convergence) are shown in Table 5.3.2. Note that we start at cycle 6: the initial equivalence of parameter variances is due to the fact that they both correspond to the estimate obtained after least squares cycle 5 incorporating the original 74 observations.

Cycle	n	R	WR	$\bar{\sigma}_x^2$	$\bar{\sigma}_y^2$	$\bar{\sigma}_z^2$	$\bar{\sigma}^2$
6(a)	80	.060	.070	1.75	1.71	2.39	1.95
6(b)	80	.072	.079	1.75	1.71	2.39	1.95
7(a)	80	.045	.051	1.64	1.56	2.17	1.79
7(b)	80	.049	.056	1.64	1.41	1.88	1.64
8(a)	80	.044	.051	1.63	1.55	2.17	1.79
8(b)	80	.048	.056	1.64	1.40	1.87	1.63
9(a)	80	.044	.051	1.63	1.55	2.17	1.79
9(b)	80	.048	.056	1.64	1.40	1.87	1.63

Table 5.3.2:   Extension of Table 5.3.1 using an additional 6 observations based on (a) $(\sin \theta)/\lambda$ (b) (3.6.28). All variances are $\text{Å}^2 \times 10^{-5}$

The reduction in parameter variances achieved using the selection scheme based on (3.6.28) clearly vindicates our D-optimality theory in this example and offers encouragement for the general success of the related ideas of Section 3.6.

CHAPTER 6

CONCLUDING COMMENTS AND IDEAS FOR FURTHER STUDY

6.1 Suggestions for Further Related Work to be Undertaken

We begin this section by considering some of the theory of Section 3.3, in particular the simplifying assumptions behind the model equation (3.3.30):

$$|FO| \ / \ FT = ft \sim N \ (|ft|, \ \sigma_0^2) \quad .$$

We concentrate on the typical case where $|FO|$ is in practice taken to be the average of the square roots of observed intensities, each of which is made up of the difference of theoretically Poisson-distributed counts. The first problem arises when such differences are negative - as may be the case for small intensities - thus making the evaluation of any appropriate square root impossible. This problem may be rectified by consideration of a model function for the observed intensity ($|FO|^2$) based on the true intensity (FT^2). We may turn to the standard Bayesian technique of specifying a prior distribution for the true intensity - e.g. based on Wilson's statistics [72] - which incorporates its inherent non-negativity, and then base inferences on the posterior distribution for the true intensity obtained in the usual way. Alternatively, the posterior mean and variance for the true structure factor (square root of the true intensity) may be obtained from this distribution and

238

utilised as before.   An approach along these lines may be
found in [32] together with further comments on the areas
of normality and unbiasedness.

The generalisation of the approach outlined in Section 3.3
would involve considering the multivariate model

$$|\underline{FO}| / \underline{FT} = \underline{ft} \sim N (\underline{ft} , \underline{\Sigma}_0) \qquad\qquad (6.1.1)$$

where $|\underline{FO}|$, $\underline{FT}$ now represent vectors of observed, true
values (as opposed to individual components) and the
covariance matrix $\underline{\Sigma}_0$ is diagonal with elements $\sigma_0^2 = \sigma_0^2(\underline{h})$,
assuming independence between observational errors.   The
analogue of the prior beliefs (3.2.19) would likewise be
of the form

$$\underline{FT} \sim N (\underline{FC}, \underline{\Sigma}_T) \qquad\qquad (6.1.2)$$

where, although the diagonal elements of $\underline{\Sigma}_T$ will be $\sigma_c^2 =$
$\sigma_c^2(\underline{h})$, $\underline{\Sigma}_T$ will in general not be diagonal.   We would then
wish to use the multivariate version of Bayes' theorem for
continuous random variables - based on equations (6.1.1)
and (6.1.2) - to obtain a posterior density function for
the vector of true values $\underline{FT}$ conditional on our observed
data.   By using the fact that the (nx1) vector of random
variables

$$\underline{X} \sim N(\underline{\mu},\underline{\Sigma})$$

has a density function of the form

$$\left[\frac{1}{(2\pi)^n} \det (\underset{\sim}{\Sigma}^{-1}) \right]^{\frac{1}{2}} \exp \left(-\frac{1}{2}(\underline{x} - \underline{\mu})^T \underset{\sim}{\Sigma}^{-1} (\underline{x} - \underline{\mu}) \right)$$

we find the appropriate density to be proportional to

$$\left[\frac{1}{(2\pi)^{2n}} \det(\underset{\sim}{\Sigma}_0^{-1} \underset{\sim}{\Sigma}_T^{-1}) \right]^{\frac{1}{2}} \exp \left[-\frac{1}{2}(\underline{x} - \underline{FC})^T \underset{\sim}{\Sigma}_T^{-1} (\underline{x} - \underline{FC}) \right.$$

$$\left. - \frac{1}{2}(|\underline{FO}| - |\underline{x}|)^T \underset{\sim}{\Sigma}_0^{-1} (|FO| - |\underline{x}|) \right] \qquad (6.1.3)$$

(where, for notational convenience, $\underline{FT}$, $\underline{ft}$ are replaced by $\underline{X}, \underline{x}$). If we could then unravel the corresponding posterior mean vector and covariance matrix, we could then proceed along similar lines to before. Note that the theory of Section 3.3 pertains if we use a diagonal approximation to $\underset{\sim}{\Sigma}_T$.

Summarising our findings then, we conclude that the approach of Section 3.3 is clearly only an approximate one. In particular, further attention could be given to the formulation of a more appropriate model. For this, we would have to take into account the precise nature of the derivation of our individual $|FO(\underline{h})|$ values. Exact theory would be difficult: for example, consideration of the distribution of the difference of two Poisson random variables (counts) is a non-trivial problem and further calculations based on taking averages and square roots would only add to the general intractability. Recourse would almost certainly have to be made to some form of

approximation. Our approximating equation (3.3.30) is likely to be acceptable in most practical applications. Problems may arise in the case of negative intensities as we have discussed; in such cases we may appeal to the results of e.g. [37], which considers the treatment of 'unobserved' reflections (for which we only have available an upper bound for the structure amplitude) in the least squares adjustment of crystal structures. As we have also seen, our approach uses the available data and prior beliefs one observation at a time. Again this is only an approximation to the more general approach outlined above, whose computational tractability needs further investigation.

The above theory concentrated on the case of centrosymmetric structure determination. We shall now very briefly outline the additional problems that arise in the general non-centrosymmetric case. Perhaps the best way of proceeding is via the following argument. The prior density for any (individual) true structure factor may be represented as being concentrated around the point on the Argand diagram corresponding to the mode of our prior beliefs, namely

$$FC = |FC| \exp(i\phi)$$

(see Figure 6.1.1). The phase ϕ - previously 0 or π - may now take any value $0 \leq \phi \leq 2\pi$. Likewise, the corresponding observed structure amplitude $|FO|$ furnishes

us with a density concentrated in an anulus based around the circle centred at the origin of radius |FO|. These two densities need to be combined appropriately to get the necessary posterior density for the true structure factor. An extra dimension has been added to the problem (previously all our calculations were based on the real axis alone) and the theory is consequently an order of magnitude more difficult. The basic underlying rationale is similar, however, and the undertaking of the appropriate analysis is a major area for research in the field of crystallographic statistics.
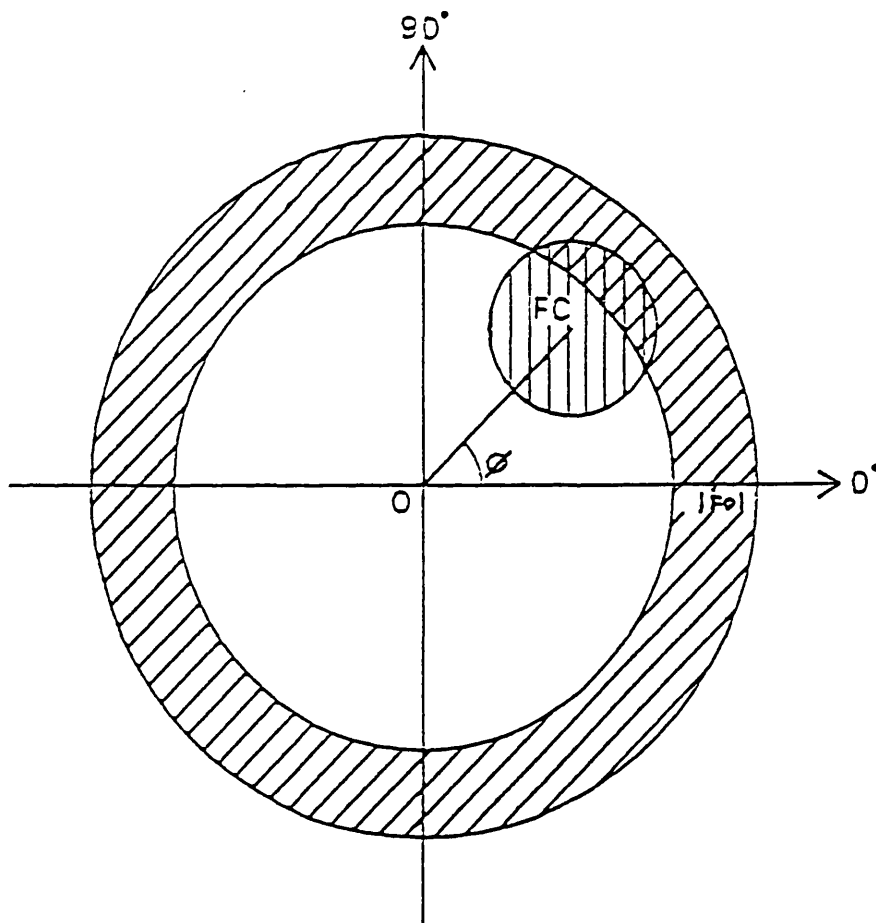

Figure 6.1.1: Pictorial representation of densities based on prior beliefs (||) and observational data (⁄⁄)

Returning to the centrosymmetric case, in which we may have symmetry stronger than centrosymmetry alone, we find that certain symmetry-related sets of reflections should yield the same structure factors. This provides a natural platform from which to apply the ideas of cross-validation, in which a structure is refined based on one set of reflections and the resultant configuration tested (validated) against further results from another set. In such cases, of course, valuable information may also be afforded us as to the experimental errors present.

It is also interesting to note the further simplifications that can be made to the centrosymmetric structure factor equation (1.2.12). For example, for structures in which for every atom at (x,y,z) there is not only an equivalent atom at $(-x,-y,-z)$ but also ones at $(x+\frac{1}{2},y,z)$ and $(x,\frac{1}{2}-y,z)$, we obtain the equations

$$
FC = \begin{cases} 4 \sum_{r=1}^{R/4} f_r \cos 2\pi(hx_r + lz_r) \cos 2\pi ky_r & h+k \text{ even} \\[2em] -4 \sum_{r=1}^{R/4} f_r \sin 2\pi(hx_r + lz_r) \sin 2\pi ky_r & h+k \text{ odd} \end{cases} \tag{6.1.4}
$$

where the summations are again taken over the asymmetric parts of the unit cell. Similar such expressions may be obtained when other symmetries pertain, and a complete list for all possible symmetry groups is given in the International Tables for X-ray Crystallography. Rather than using expressions such as (6.1.4), however, crystallographers often continue to use the original

equation (1.2.12). Although such a strategy enables a unifying approach to be made to all centrosymmetric structure determinations, the ready availability of the alternative expressions suggests that more use could and should be made of them in future analyses.

We next turn to the topic of appropriate weighting schemes. We noted in Section 2.2 that we might in general need to consider dynamic weighting schemes that vary with the current parameter estimate. For example, we might argue that structure factors with large (in modulus) calculated values but small observed values merit larger weights since they unequivocally tell us the direction of the shift needed in the fitted value. A corresponding component of the weighting scheme used may be based on e.g. $|FC|$: $|FO|$ ratios. Similarly, components may be based for example on likelihood of correct phase - sign in centrosymmetric case - allocation, or on some non-linearity criterion as suggested in Section 3.6. As is often the case in crystallography, numerous ad hoc rules of an appealing intuitive nature may readily be specified. Further work for the statistician lies in justifying the best such schemes from a theoretical viewpoint.

The rationale behind the weighting scheme (3.3.45) derived earlier was to provide a realistic compromise between the observed data and our prior (Bayesian) beliefs. Such schemes may of course be used to improve the starting data for the least squares procedure as briefly discussed in

Section 3.3.   The problem still remains of updating the appropriate weight matrix at each stage of our refinement procedure.   Such problems of adjustment are discussed in [10]: some of the pertinent points arising will now be summarised, which emphasise current areas of related research.

Suppose we wish to change the least squares weights at each iteration because of the dependence of the measurement error on the unknown parameter $\underline{\theta}$.   The effect of this is to replace the original least squares equations - c.f. (3.6.1) - by the equations

$$0 = \frac{\partial}{\partial \tau}(r) \sum_{\underline{h}} W_{\underline{h}}(\tau) \ (FC(\tau) - FO(\underline{h}))^2$$

$$(\tau = \xi, \eta, \varsigma; \ r=1...R)$$

(6.1.5)

where the weights $w_{\underline{h}}(\tau)$ now vary with the parameter estimates as discussed above.   We no longer have a least squares procedure in which a fixed sum of squares of deviations is minimised: it is instead more closely related to so-called quasi-likelihood procedures (see [56],[68]).   It is interesting to note [10] that a necessary and sufficient condition for the quasi-likelihood function to exist is that the weight of a measurement is allowed to depend only on the unknown theoretical value of that measurement, as was the case with our theory of Section 3.3.

The main purpose of [10] is to show how the problem of

245

uniqueness of solutions to the equations (6.1.5) is related to a certain curvature. The curvature in question is analagous to Efron's statistical curvature [27] and takes into account the varying weights in an appropriate way. It is suggested that such a curvature also determines the behaviour of the Gauss-Newton (least squares) iterative procedure, when weights are adjusted. Although similar results may be found in [50], and the area is investigated generally in [64], there is still plenty of scope for further related research.

The theory of Bates and Watts (see Section 4.3) that figured so prominently in Chapter 4 is related to the above work in the following sense. They conclude [5] that 'the whole concept of measuring both intrinsic and parameter-effects curvatures by an array $\underline{A}$.. in the multi-parameter situation can be extended to Efron's statistical curvature.' The other main point they raise in their discussion is the need for the design of non-linear experiments for small curvatures. This is a difficult and challenging problem and is still very much an open area for research. For our purposes, however, further investigation of the theoretical properties of our non-linear parameter estimation schemes proposed in Chapter 4 is perhaps a more pressing concern.

One of the more challenging purely statistical problems to arise from the work of this thesis is the question of the efficiency of the modified RLS algorithm of Section 3.5

for use in conjunction with non-linear models. Following Figure 3.5.1 we considered the question of whether we could justifiably omit Gauss-Newton cycles in the latter stages of the procedure. Further work lies in determining conditions under which such an approach would be acceptable, and providing further details of its theoretical properties. The results of the schemes employed in Chapter 5 - which incorporated no fixed data cycles during the sequential data acquisition stages - are promising.

As suggested in Section 1.4 we may also wish to develop alternative sequential strategies based on cost considerations and a decision theoretic approach. For the question of associated stopping rules we may wish for example (in the least squares approach) to accept as a stopping criterion that stage at which the accuracy of the fitted values is commensurate with the assessed experimental errors of the as yet unobserved reflections.

Finally, of course, much further work remains in the practical testing and development of the various schemes proposed in this thesis.

6.2 Conclusion

We conclude by stating that although in this thesis we have merely addressed some of the statistical problems that arise in X-ray crystallography, and have consequently suggested improvements to existing procedures of structure

determination, it is clear that sensible use of statistical theory still has much to offer the world of X-ray crystallography.

APPENDIX I

INTEGRATED INTENSITY AND RELATED CONCEPTS

For the rotation photographs described in Section 1.3 let Ω be the constant angular velocity at which the reciprocal lattice region is swept through the Ewald sphere. For a small crystal completely bathed in a uniform beam of intensity I_0 (expressed in units of energy per unit area per unit time) the integrated intensity I is given by

$$I = I_0 \; r^2 \; Lp \; \frac{\lambda}{\Omega} \; \left(\frac{F}{V}\right)^2 \lambda^2 v \qquad (A1.1)$$

if absorption and extinction (to be presently introduced) are neglected, where r represents the 'classical radius of an electron', v the volume of the unit cell and V the volume of the crystal. The dimensionless factors L and p representing the Lorentz and polarisation factors respectively will also be presently introduced. F represents the structure factor and λ the wavelength of the X-ray radiation used.

Since the only factors that vary between reflections are L, p and F^2, we may write

$$F^2 \propto \frac{I}{L_p} \; .$$

Most crystallographers measure I, and hence F^2, on an arbitrary, relative scale. Conversion to an approximate absolute scale may be made with the help of intensity

statistics at the start of each cycle of the structure determination. Alternatively, a scale factor may be included as an adjustable parameter in the resultant least-squares refinement.

The Lorentz and polarisation factors shall now be discussed. The Lorentz factor L expresses the fact that for a given Ω, different hkl have different 'times-of-reflection' opportunity. For rotation photographs, L has the simple form

$$L = \frac{1}{\sin 2\theta}$$

where θ is the Bragg angle.

The radiation from a normal X-ray tube is unpolarised (i.e. emits wave trains, vibrating in all direction perpendicular to the direction of propagation), but after reflection from a crystal the beam is polarised. The fraction of energy lost in this process, the polarisation factor p, depends only on the Bragg angle θ:

$$p = \frac{1 + \cos^2 2\theta}{2}$$

The Lorentz and polarisation factors may therefore conveniently be combined into a single trigonometric expression:

$$Lp = \frac{1 + \cos^2 2\theta}{2 \sin 2\theta} \ . \tag{A1.2}$$

Application of this factor is essential in order to bring

the F^2 data onto a correct relative scale.

Absorption and extinction, which have so far been assumed to be negligible, will now be discussed. All materials absorb X-rays according to an exponential law:

$$I = I_0 \exp(-\mu t)$$

where I and I_0 are the transmitted and incident intensities, μ is the linear absorption coefficient, and t is the path length through the material (comprising an incident and diffracted beam length). The transmission of the X-ray beam through a crystal is given by

$$\frac{I}{I_0} = \exp(-\mu t)$$

The intensity diffracted by the crystal as a whole is then reduced by the transmission factor

$$A = \frac{1}{V} \int \exp(-\mu t) \, dV \qquad\qquad (A1.3)$$

assuming the incident beam to have uniform intensity cross-section. We find that the actual intensity produced from the crystal is A x the intensity in the absence of absorption. In practice, crystal faces are frequently not well defined and it is necessary to resort to empirical methods for estimating the factor A.

During the latter stages of a crystal structure refinement it is often the case that the observed $|F|$ values are systematically smaller than the moduli of the

corresponding calculated F values for very strong reflections. Typically we find that this result may be characterised by the relation

$$\frac{I_0}{I_c} = \exp\,(-g\,I_c)$$

where the suffices 0 and c denote observed and calculated values, and the constant g may be estimated. This kind of systematic error is called extinction; its influence on the observed F values can be removed by multiplying the observed intensities by $\exp(-gI_c)$ and re-calculating the corrected observed F values.

GENERAL RESULTS NEEDED FOR SECTION 4.6

Following Section 4.3, the relevant $\underset{\sim}{Q}$ $\underset{\sim}{R}$ decomposition for the (nx3) matrix $\underset{\sim}{V}$. with columns $\underset{\sim}{v_1}$, $\underset{\sim}{v_2}$, $\underset{\sim}{v_3}$ yields the (3x3) matrix $\underset{\sim}{R_1} = \underset{\sim}{L}^{-1}$ and the (nx3) matrix $\underset{\sim}{U}$. defined by

$$A = L_{11}^{-1} = \left(\sum_{i=1}^{n} v_{1i}^{2} \right)^{\frac{1}{2}} \tag{A2.1}$$

$$B = L_{12}^{-1} = \frac{\sum_{i=1}^{n} v_{1i} v_{2i}}{A} \tag{A2.2}$$

$$C = L_{13}^{-1} = \frac{\sum_{i=1}^{n} v_{1i} v_{3i}}{A} \tag{A2.3}$$

$$D = L_{22}^{-1} = \left(\sum_{i=1}^{n} v_{2i}^{2} - B^{2} \right)^{\frac{1}{2}} \tag{A2.4}$$

$$E = L_{23}^{-1} = \frac{\sum_{i=1}^{n} v_{2i} v_{3i} - BC}{D} \tag{A2.5}$$

$$F = L_{33}^{-1} = \left(\sum_{i=1}^{n} v_{3i}^{2} - C^{2} - E^{2} \right)^{\frac{1}{2}} \tag{A2.6}$$

and

$$U._{i1} = \frac{v_{1i}}{A} \qquad (i=1...n) \tag{A2.7}$$

$$U \cdot_{i2} = \frac{(v_{2i} \frac{Bv_{1i}}{A})}{D} \qquad (i=1\ldots n) \qquad \text{(A2.8)}$$

$$U \cdot_{i3} = \frac{(v_{3i} + (\frac{BE-CD}{AD}) v_{1i} - \frac{E}{D} v_{2i})}{F}(i=1\ldots n) \quad . \quad \text{(A2.9)}$$

The corresponding (3x3x3) parameter effects curvature array $\underset{\sim}{A}..$ contains elements α_{jkl} given by

$$\alpha_{jkl} = \sum_{i=1}^{n} U \cdot_{ji} (W_{kl})_i \qquad \text{(A2.10)}$$

where

$$\underline{W}_{11} = \frac{1}{A^2} \underline{V}_{11} \qquad \text{(A2.11)}$$

$$\underline{W}_{12} = \underline{W}_{21} = \frac{-B}{A^2 D} \underline{V}_{11} + \frac{1}{AD} \underline{V}_{12} \qquad \text{(A2.12)}$$

$$\underline{W}_{13} = \underline{W}_{31} = \frac{(BE-CD)}{A^2 DF} \underline{V}_{11} - \frac{E}{ADF} \underline{V}_{12} + \frac{1}{AF} \underline{V}_{13} \qquad \text{(A2.13)}$$

$$\underline{W}_{22} = \frac{B^2}{A^2 D^2} \underline{V}_{11} - \frac{2B}{AD^2} \underline{V}_{12} + \frac{1}{D^2} \underline{V}_{22} \qquad \text{(A2.14)}$$

$$\underline{W}_{23} = \underline{W}_{32} = \frac{(CD-BE)}{A^2 D^2 F} B \underline{V}_{11} + \frac{(2BE-CD)}{AD^2 F} \underline{V}_{12}$$

$$- \frac{B}{ADF} \underline{V}_{13} - \frac{E}{D^2 F} \underline{V}_{22} + \frac{1}{DF} \underline{V}_{23} \qquad \text{(A2.15)}$$

$$W_{33} = \frac{(BE-CD)^2}{A^2 D^2 F^2} \underline{v}_{11} + \frac{2(CD-BE)E}{AD^2 F} \underline{v}_{12} + \frac{2(BE-CD)}{ADF^2} \underline{v}_{13}$$

$$+ \frac{E^2}{D^2 F^2} \underline{v}_{22} - \frac{2E}{DF^2} \underline{v}_{23} + \frac{1}{F^2} \underline{v}_{33} \qquad \text{(A2.16)}$$

with

$$\underline{v}_{jk} = \left. \frac{\partial^2 \eta}{\partial \theta_j \partial \theta_k} \right|_{\underline{\theta}_0} \quad .$$

The re-parameterisation given in equation (4.6.8) - namely

$$\phi_j = \sum_{k=1}^{3} \left(\gamma_{jk} \sum_{i=1}^{n} \eta_i(\underline{\theta}) \, v_{ki} \right) \qquad (j=1,2,3) -$$

contains the coefficients

$$\gamma_{11} = B^2 (E^2 + F^2) + D^2(C^2 + F^2) - 2BCDE \qquad \text{(A2.17)}$$

$$\gamma_{12} = \gamma_{21} = ACDE - AB(E^2 + F^2) \qquad \text{(A2.18)}$$

$$\gamma_{13} = \gamma_{31} = AD(BE - CD) \qquad \text{(A2.19)}$$

$$\gamma_{22} = A^2(E^2 + F^2) \qquad \text{(A2.20)}$$

$$\gamma_{23} = \gamma_{32} = -A^2 DE \qquad \text{(A2.21)}$$

$$\gamma_{33} = A^2 D^2 \quad . \qquad \text{(A2.22)}$$

The corresponding (3x3) matrix $\underline{G}$. (evaluated at $\underline{\theta}_0$) is given by $\underline{G}. = c\underline{I}_3$, where $\underline{I}_3$ is the (3x3) identity matrix and

$$c = A^2 D^2 F^2 \quad . \tag{A2.23}$$

Returning to the sum of distinct model functions given by equation (4.6.10), the relevant $\underset{\sim\sim}{QR}$ decomposition of Section 4.3 may be written

$$\left[\begin{array}{c|c|c} \underset{\sim}{U}_{\cdot 1} & \underset{\sim}{U}_{\cdot 2} & \underset{\sim}{N} \end{array} \right] \left[\begin{array}{c|c} \underset{\sim}{L}_1^{-1} & \underset{\sim}{0} \\ \hline \underset{\sim}{0} & \underset{\sim}{L}_2^{-1} \\ \hline \multicolumn{2}{c}{\underset{\sim}{0}} \end{array} \right] \tag{A2.24}$$

provided that $\underset{\sim}{U}_{\cdot 1}^{T} \underset{\sim}{U}_{\cdot 2} = 0$, where quantities with the subscript j are the analogous quantities for the individual model functions $f_j(\underline{x}, \underline{\theta}_j)$ (j=1,2). The (pxpxp) parameter effects curvature array (where $p = p_1 + p_2$) has p_1, p_2 faces of the form

$$\left[\begin{array}{c|c} \underset{\sim}{A}_{11} & \underset{\sim}{0} \\ \hline \underset{\sim}{0} & \underset{\sim}{A}_{12} \end{array} \right] \quad , \quad \left[\begin{array}{c|c} \underset{\sim}{A}_{21} & \underset{\sim}{0} \\ \hline \underset{\sim}{0} & \underset{\sim}{A}_{22} \end{array} \right] \tag{A2.25}$$

respectively, where $\underset{\sim}{A}_{jk}$ are faces of the arrays

$$[\underset{\sim}{U}_{\cdot j}^{T}] \, [\underset{\sim}{L}_k^{T} \, \underset{\sim}{V}_{\cdot \cdot k} \, \underset{\sim}{L}_k] \quad . \tag{A2.26}$$

Note that $\underset{\sim}{A}_{11}$ and $\underset{\sim}{A}_{22}$ are faces of the original parameter effects arrays. Using the re-parameterisations $\underline{\phi}_j$ (j=1,2) that yield zero individual parameter effects, note that the (pxp) array $\underset{\sim}{G}$. is of the form

$$\left[\begin{array}{c|c} \underset{\sim}{G} \cdot_1 & \underset{\sim}{0} \\ \hline \underset{\sim}{0} & \underset{\sim}{G} \cdot_2 \end{array} \right] \qquad\qquad (A2.27)$$

and the (pxpxp) array $\underset{\sim}{G}..$ has p_1, p_2 faces of the form

$$\left[\begin{array}{c|c} \underset{\sim}{G}..{}_1 & \underset{\sim}{0} \\ \hline \underset{\sim}{0} & \underset{\sim}{0} \end{array} \right] \quad , \quad \left[\begin{array}{c|c} \underset{\sim}{0} & \underset{\sim}{0} \\ \hline \underset{\sim}{0} & \underset{\sim}{G}..{}_2 \end{array} \right] \qquad (A2.28)$$

respectively (where $\underset{\sim}{G}..{}_j$ is here used to denote faces of the original $\underset{\sim}{G}..{}_j$ arrays). The new parameter effects array under this re-parameterisation has p_1, p_2 faces of the form

$$\left[\begin{array}{c|c} \underset{\sim}{0} & \underset{\sim}{0} \\ \hline \underset{\sim}{0} & \underset{\sim}{A}_{12} \end{array} \right] \quad , \quad \left[\begin{array}{c|c} \underset{\sim}{A}_{21} & \underset{\sim}{0} \\ \hline \underset{\sim}{0} & \underset{\sim}{0} \end{array} \right] \qquad (A2.29)$$

respectively, which is seen to be a definite improvement on the original array (A2.25).

REFERENCES

1. P. Armitage
   Sequential Medical Trials
   Blackwell, 2nd edition (1975)

2. Y. Bard
   New York Scientific Center
   Report 322.0902, IBM, New York (1967)

3. Y. Bard
   Non-linear Parameter Estimation
   Academic Press (1974)

4. D.M. Bates and D.G. Watts
   JRSSB 42, 1 (1980)

5. D.M. Bates and D.G. Watts
   Ann Stat 9, 1152 (1981)

6. D.M. Bates and D.G. Watts
   Non-linear Regression Analysis and its Applications
   Wiley (1988)

7. E.M.L. Beale
   JRSSB 22, 41 (1960)

8. J. Bibby and H. Toutenburg
   Prediction and Improved Estimation in Linear Models
   Wiley and Sons (1977)

9. G.W. Booth and T.I. Peterson
   IBM SHARE Program Pa. No. 687 (WLNLI)   (1958)

10. K. Borre and S.L. Lauritzen
    Geodetisk Institut Meddelelse No. 58 (1989)

11. T.L. Boullion and P.L. Odell
    Generalised Inverse Matrices
    Wiley (1970)

12. C.W. Bunn
    Chemical Crystallography
    Oxford University Press, 2nd edition (1961)

13. P. Businger and G.H. Golub
    Numerische Math 7, 269 (1965)

14. J.M. Chambers
    Biometrika 60, 1 (1973)

15. W. Cochran
    Acta Cryst 1, 138 (1948)

16. W. Cochran and M.M. Woolfson
    Acta Cryst $\underline{8}$, 1 (1955)

17. W.G. Cochran
    Sampling Techniques
    Wiley, 3rd edition (1977)

18. D.M. Collins and M.C. Mahar
    (Paper in) Crystallographic Statistics
    Indian Academy of Sciences (1982)

19. M.S.M. Coxeter
    Introduction to Geometry
    Wiley (1961)

20. D.T. Cromer
    J. Phys. Chem $\underline{61}$, 254 (1957)

21. D.W.J. Cruickshank
    Acta Cryst $\underline{2}$, 65 (1949)

22. D.W.J. Cruikshank, D.E. Pilling, A. Bujosa, F.M.
    Lovell and M.R. Truter
    Computing Methods in the Phase Problem
    Pergamon Press (1961)

23. J.L. De Boer
    (Paper in) Crystallographic Statistics
    Indian Academy of Sciences (1982)

24. N.R. Draper and H. Smith
    Applied Regression Analysis
    Wiley (1966)

25. W. Duane
    Proc. Nat. Acad. Sci. $\underline{11}$, 489 (1925)

26. J.D. Duntiz
    X-ray analysis and the Structure of Organic Molecules
    Cornell University Press (1979)

27. B. Efron
    Ann. Stat. $\underline{3}$, 1189 (1975)

28. R.A. Fisher
    The Design of Experiments
    Oliver and Boyd (1935)

29. P.D. Flanagan, P.A. Vitale and J. Mendelsohn
    Technometrics $\underline{11}$, 265 (1969)

30. S. French
    Acta Cryst $\underline{A34}$, 728 (1978)

31. S. French and S. Oatley
    (Paper in) Crystallographic Statistics
    Indian Academy of Sciences (1985)

32. S. French and K.S. Wilson
    Acta Cryst A34, 517 (1978)

33. J.P. Glusker and K.N. Trueblood
    Crystal Structure Analysis
    Oxford University Press (1972)

34. Z. Govindarajulu
    Sequential Statistical Procedures
    Academic Press (1975)

35. G. Grimett and D. Stirzaker
    Probability and Random Process
    Clarendon Press (1982)

36. S.F. Gull, A.K. Livesey and D.S. Sivia
    Acta Cryst A43, 112 (1987)

37. W.C. Hamilton
    Acta Cryst 8, 185 (1955)

38. D. Harker
    Am. Mineral 33, 764 (1948)

39. H.O. Hartley
    Trchnometrics 3, 269 (1961)

40. H. Hauptman
    Acta Cryst A31, 680 (1975)

41. H. Hauptman
    Acta Cryst A32, 877 (1976)

42. H. Hauptman and J. Karle
    Acta Cryst 6, 136 (1953)

43. R.J. Havighurst
    Proc. Nat. Acad. Sci. 11, 502 (1925)

44. P. Hougaard
    JRSSB, 44, 242 (1982)

45. E.W. Hughes
    J. Am. Chem. Soc 63, 1737 (1941)

46. E.W. Hughes
    Acta Cryst 2, 34 (1949)

260

47. J.W. Jeffery
    Methods in X-ray Crystallography
    Academic Press (1971)

48. J. Karle
    (Paper in) Structure and Statistics in
    Crystallography
    Adenine Press (1985)

49. J. Karle and H. Hauptman
    Acta Cryst 6, 131 (1953)

50. T. Krarup
    (Paper in) Daar heb ik veertig jaar over
    nagedacht ... Feestbundel ter gelegenheid van de 65
    ste verjaardag van Professor Baarda, Delft, 145
    (1982)

51. M.F.C. Ladd and R.A. Palmer
    Structure Determination by X-ray Crystallography
    Plenum Press, 2nd edition (1985)

52. D.V. Lindley
    Bayesian Statistics
    SIAM, Philadelphia, Pennsylvania (1971)

53. D.V. Lindley and A.F.M. Smith
    JRSSB 34, 1 (1972)

54. H. Lipson and W. Cochran
    The Determination of Crystal Structures
    G. Bell and Sons, 3rd edition (1966)

55. D.W. Marquardt
    SIAM J. 11, 431 (1963)

56. P. McCullagh
    Ann. Stat. 11, 59 (1983)

57. M. Mendes and C. De Polignac
    Acta Cryst A29, 1 (1973)

58. H.J. Milledge, M.J. Mendelssohn, C.M. O'Brien and
    G.I. Webb
    (Paper in) Structure and Statistics in
    Crystallography
    Adenine Press (1985)

59. G.J.S. Ross
    Appl. Statist 19, 205 (1970)

60. S.D. Silvey
    Optimal Design
    Chapman and Hall (1980)

61. R.A. Sparks
    PHd Thesis, University of California (1958)

62. R. Srinivasan and S. Parthasarathy
    Some Statistical Applications in X-ray
    Crystallography
    Pergamon Press (1976)

63. D.M. Steinberg and W.G. Hunter
    Technometrics 26, 71 (1984)

64. P.J.G. Teunissen
    Netherlands Geodetic Commision
    Publications on Geodesy
    New Series. Vol 8, No. 1 (1985)

65. A. Wald
    Sequential Analysis
    Wiley and Sons (1947)

66. A. Wald
    Statistical Decision Functions
    Wiley (1950)

67. J. Waser
    Acta Cryst 16, 1091 (1963)

68. R.W.M. Wedderburn
    Biometrika 61, 439 (1974)

69. G.B. Wetherill
    Sampling Inspection and Quality Control
    Methuen and Co. Ltd. (1969)

70. G. B. Wetherill and K.D. Glazebrook
    Sequential Methods in Statistics
    Chapman and Hall, 3rd edition (1986)

71. A.J.C. Wilson
    Nature (London) 150, 151 (1942)

72. A.J.C. Wilson
    Acta Cryst 2, 318 (1949)

73. R.L. Winkler
    Introduction to Bayesian Inference and Decision
    Holt, Rinehart and Winston, Inc (1972)