Modelling Piezo-Spectroscopic Data

Andrew L. McCarren B.Sc.

Ph.D in Computer Applications

1994

Modelling Piezo-Spectroscopic Data

Andrew L. McCarren B.Sc.

Submitted in fulfilment of the requirements for a

Ph.D Degree, 1994

School of Computer Applications

Dublin City University

Dublin 9.

Supervisor: Dr. Heather Ruskin.

July 1994

Declaration

I herby certify that this material, which I now submit for assessment on the programme of study leading to the award of Ph.D is entirely my own work and has not been taken from the work of others save and to the extent that such work has been cited and acknowledged within the text of my work.

Signed: Deborah McGuinness Date: 15-9-94

Acknowledgements

I am deeply grateful to Dr. Heather Ruskin for her belief, guidance, encouragement and support over the last four years. I am also indebted to Dr. Martin Henry for his wisdom and time, to Stephen and Mal for their expert programming guidance and proof reading, and to Fiona for her patience and understanding over the last couple of years.

I would also like to thank the School of Computer Applications for providing me with the necessary opportunity and funding to undertake this research.

Finally, my great appreciation and love to my parents Andrew and Joy, and to my Uncle and Aunt, Cian and Valerie O'Kane for their support and understanding over the years.

Table of Contents

Abstract

Information from the many kinds of spectroscopy used by chemists and physicists is fundamental to our understanding of the structure of materials. Numerical techniques have an important role to play in the augmentation of the instrumentation and technology available in the laboratory, but are frequently viewed as separate from the laboratory procedures. We examine the model approaches which are currently applied in spectroscopy and determine their applicability to piezo-spectroscopic data. Typically, in piezo-spectroscopic modelling the analyses in question are required to handle large complex secular matrices, to distinguish between components in the experimental results, and to identify the transition types as rapidly and as efficiently as possible. The method proposed is based on providing a shell to the Powell or Fletcher-Reeves minimisation algorithms, and gives favourable results compared to those previously used.

Additionally, the statistical properties of the least-squares estimator used in the Powell-shell are examined and implications for nonlinear model functions are discussed. We also show that the least squares estimator performs well for piezo-spectroscopic data compared to those currently used in multi-response data analysis.

Finally we describe the development of a software tool which incorporates all features of fitting piezo-spectroscopic data.

Chapter 1

General Overview and Introduction

1.1 Introduction

The study of electronic and optoelectronic materials has been intensive for decades. This reflects the central role of materials analysis and characterisation in the improvement and development of device technology. For these materials, much of the research has been directed at obtaining a full understanding of the properties of defects. This applies equally to defects deliberately produced for device optimisation and to unwanted defects which degrade device performance. Spectroscopic techniques, of which electron paramagnetic resonance and photoluminescence are good examples, possess high sensitivity and accuracy and are widely used in the study of defects in materials. The power of spectroscopic measurements can be considerably enhanced by applying known perturbations to the system under study, Fowler (1968). Hydrostatic and uniaxial stress, and magnetic and electric fields are widely used perturbations in conjunction with optical spectroscopy. From the effects produced by perturbation, much detailed information on the physical structure and electronic state of the defect may be obtained.

The fitting of all types of spectroscopic data is however a topic which is characterised by a lack of coherent reporting in the literature. This is due both to the diversity of current experimental techniques and applications and the obvious corollary of the lack of a common forum in terms of reporting on the

analyses performed. Such considerations have ensured that the effort of reconciling common features in the underlying linear models is non-trivial, even where such models have long been employed by mathematicians, statisticians and computer scientists. Similarly, an informed discussion of the scope of the optimisation technique used is usually lacking, since the primary aim is to interpret the values obtained for the parameters defining the problem rather than to assess the efficiency of the estimation process. The wider aspects of achieving a good fit to spectroscopic data have not, of course, been entirely ignored and notable examples include Rusling and Kumosinki(1991), Femenias and Cheval (1992) and Leurgans and Ross(1992).

In this thesis, we address the central issues of fitting spectroscopic data and in particular Piezo-spectroscopic data. General spectroscopic fitting techniques are reviewed with a view to implementing them in piezo-spectroscopy, and the background to piezo-spectroscopic data generation is given. Methods for fitting piezo-spectroscopic data to date have principally involved finding the roots of a characteristic polynomial (i.e. the eigenvalues of an energy shift), but these do not always yield directly the quantities of interest, namely the parameters defining the optical transition. The alternative approach which we have developed involves the provision of a discrete shell to the Powell or Fletcher-Reeves algorithm, Walsh(1975) Press et al.(1991). The method is used both to re-examine previously published data and to analyse new data. We also investigate the appropriateness of the least squares estimator which is the basis of the Powell-shell and compare it to equivalent estimators also used for Multi-response data. An assessment of the error assumptions for the Powell-shell is also given and their value as diagnostics discussed.

1.2 General Fitting Procedures

1.2.1 Piezo-spectroscopic data

Owing to the fact that a finite number of possible point defect configurations exists for crystals, general equations, which describe the effects of an applied stress on cubic crystals, are known for all possible configurations, Kaplyanski(1963). These equations have been derived for stresses along <001>[1], <111> and <110> crystallographic directions. In practice, data is recorded for stresses along these three directions and a best fit is sought between the data and the appropriate equations. The equations are expressed in terms of several parameters (in units of energy per unit stress or shift rates) and values of these parameters which produce the best overall fit are extracted. The parameters in question are related to the bulk moduli of the solid. The values obtained for a particular defect indicate the extent to which the characteristics of the defect depart from those of the bulk solid. In addition, the symmetry properties of the electron energy states and (in favourable cases) the nature of the defect vibrational modes may be obtained.

1.2.2 Fitting Piezo-spectroscopic data

Methods used in the fitting of transitions have not been investigated in great detail and publications relating to them have been limited. As an illustration, Hughes and Runcimann(1967) describe a simplistic but still widely used fitting technique for a Tetragonal E->E example. In this case for stress measurements in the <100>, <111> and <110> directions, nine experimentally measured shifts are obtained to fit four parameters, giving (i.e.) an over-determined set of equations, which may in principle be solved for a best fit. The energy shifts

[1] *<xyz> denotes the x,y and z axes co-ordinates of the relevent stress vector.*

correspond to the eigenvalues of a matrix, known as a *secular matrix*, where each element of this matrix corresponds to the probability of a transition between two individual energy levels of the complete transition. The authors maintain however, that the error involved in the measurements does not warrant a sophisticated approach, so that in order to see if a consistent set of parameters can be obtained, the results for each direction are considered separately and each parameter determined from each of the resulting three sets of equations. A consistency check is then made between results from different stress directions, since each parameter is determined at least once.

Extensive tabulations have been prepared for the effects of uniaxial stresses on optical centres in cubic crystals, Kaplyanski (1964)a),b), Hughes and Runciman (1967), Davies and Nazaré (1980). However, these tables are relevant primarily to *absorption* spectroscopy. When the optical measurements are made using *photoluminescence* the relative intensities of the stress-split components of the transition may be very different from the intensities of the same stress-split components in absorption, Mohammed et al.(1982). These authors give equivalent limited tabulations for the effect of uniaxial stresses on the photoluminescence transitions.

For both sets of transitions the tabulations give the individual stress energy equations for an equivalent experimental energy emission. The tables provide an immense contribution in obtaining fits for data, particularly where only one transition is suspected, and in this case the approach of Hughes and Runcimann(1967) can be readily applied. Unfortunately, for real data, a number of factors may influence the complexity of the problem and there may well be more than one potential transition, see e.g. do Carmo et al.(1988). Typically the analysis is then required to handle high ordered complex secular matrices, to distinguish between components in the experimental results, and to identify the transition types as rapidly and efficiently as possible.

1.2.3 General Spectroscopy

Other methods are numerous and include multi-linear modelling of spectroscopic data, e.g. the biochemical examples of Leurgans and Ross(1992). In this case the different components of a particular chemical are again analysed using photoluminescence, but the approach is multivariate since there are numerous independent variables, including line energy, which accounts for the intensity of light being emitted during luminescence. (This contrasts with the case for piezo-spectroscopic data where stress is the single explanatory variable dominating the characterisation of the transition). Other multivariate approaches such as principal components, Windeg et al.(1991) and Gelsema et al.(1992), are also commonly applied as are methods which require deconvolution of the complete spectral profile, Seraydarian et al.(1992). A distinction may also be made between the spectral profile approach and that which involves modelling *realisations* of the spectrum e.g. Davies et al.(1990), Diaz et al.(1991). We discuss further details of these methods and their appropriateness for piezo-spectroscopic data in the next chapter.

1.3 Scope of Thesis

The organisation of the material in this thesis is as follows:

Chapter 2 "Spectroscopic Techniques "

In this chapter we discuss the current spectroscopic modelling techniques which, in particular are applied in the Physical and Chemical sciences. We group these into three different categories, namely Potential Energy Curve Fitting, Spectral Profile and Multivariate approaches. Finally, we outline the current approaches

used to analyse Piezo-spectroscopic data.

Chapter 3 "Transition Identification"

In this chapter we provide the physical background to the models and describe the steps involved in the identification of a transition.

Definitions of physical terms, and details of the generation of piezo-spectroscopic data are also given.

Chapter 4 "Method of Fitting the data"

We describe here a general methodology for the identification of transitions and common approaches to obtaining parameter estimates. We discuss alternative fitting techniques to those commonly used, and in particular recommend a method based on a "shell" supplied to Powell's or Fletcher-Reeves algorithms in order to obtain a non-linear least squares fit between the experimental data and the eigenvalues of a suspected secular matrix.

Chapter 5 "Error Analysis"

In this chapter, we evaluate the precision and accuracy of the parameter estimates obtained from the Powell-Shell through a consideration of the diagnostics of the fits. We discuss the appropriateness of other estimators and compare their relative merits to the least-squares approach of the Powell-Shell. Finally, we investigate the degree of non-linear behaviour of the model parameters and the expectation surface obtained using the Powell-Shell.

Chapter 6 "The Tool"

Specification and development of a software tool using the alternative "shell" are given. An analysis of convergence times for this algorithm is also reported on numerous machines and for problems of varying complexity.

Chapter 2

Spectroscopic Modelling Techniques

2.1 Introduction

In order to determine a method for fitting piezo-spectroscopic data and hence to identify the appropriate transition, we review the current methods used to model general spectroscopic data. These include modelling the complete spectrum using multivariate and spectral profile techniques, and modelling realisations of the spectrum, using methods such as potential energy curve fitting, Diaz et al.(1991), and electronic model identification, Kober and Meyer(1984).

2.2 Spectroscopic Data Models

Information from many kinds of spectroscopy used by chemists and physicists is fundamental to our understanding of the structure of materials. Numerical modelling techniques have an important role to play in the augmentation of the instrumentation and technology available in the laboratory, but are frequently viewed as separate from laboratory procedures. In this section we review the current literature on spectroscopic modelling. Of interest are methods of analysis which share common features with and which may be applicable to, our particular type of spectroscopic data.

Generally, the majority of work involved in modelling chemically-derived spectroscopic data is concerned with the detection of elements or components in a substance. Similarly in physics, Russ (1977) summarised the different

approaches that were in use in the late seventies and suggested that modelling techniques which are simple to interpret should be given equal weight to those that involve incredible complexity and computation. Over the years analysts

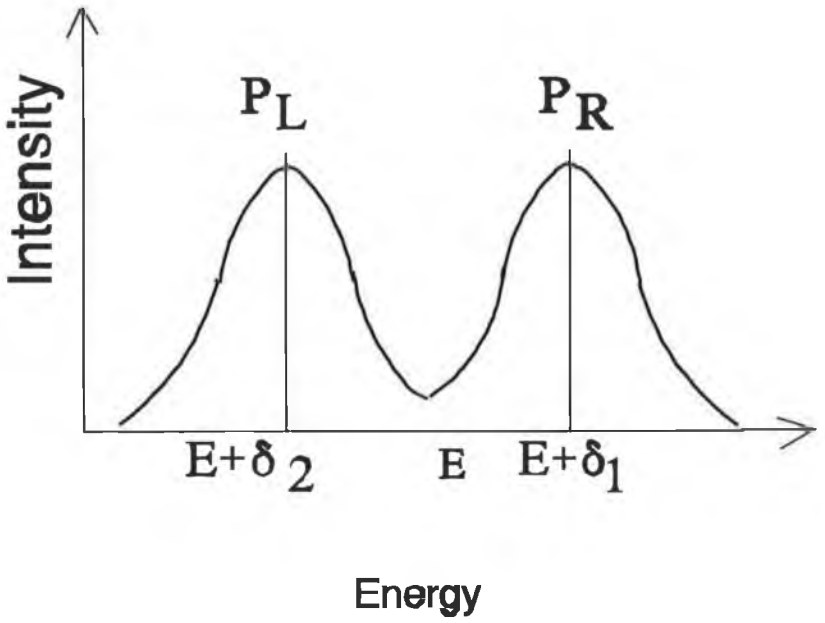


Figure 2.1 A simple vibrational spectrum with two peaks P_L and P_R.

have applied numerous techniques to spectroscopic data, depending on the required information. Generally, however, these may be broadly classified into (i) Multivariate, (ii) Spectral Profile fitting techniques and (iii) Potential energy curve fitting. With respect to the spectrum shown in figure (2.1), methods (i) and (ii) would be applied to the complete spectrum but would differ in the underlying theory. Method (iii) deals with realisations of the spectrum which in this example might correspond to the two peaks P_L and P_R.

An example of an application of a *multivariate* technique, is given by the modelling of Electron Energy Loss Spectroscopy (EELS). This spectroscopic method investigates the physical properties of a material, by detecting the presence of specific elements in a particular specimen, Gelsema et al. (1992). In addition, discriminant analysis and partial least squares regression have been used to choose samples for near infrared spectroscopy calibrations for

agricultural products, Shenk and Westerhaus (1991). Alternatively, a spectral fitting example might involve applying non-linear regression to a function of the sum of numerous gaussians in the analysis of heat deposition profiles. These profiles are crucial to all transport analysis of beam heated discharges, Seraydarian et al.(1992). The construction of potential energy curves from vibrational spectroscopic data is also an important problem since many physical properties may be calculated from an accurate knowledge of such curves. Potential energy curve fitting is based on perturbation theory and can be described as applying an appropriate numerical integration algorithm to the potential energy function such that iterative corrections are obtained until the eigenvalues of the Schrodinger equation agree with the realisations of the spectrum energy peaks, Diaz et al.(1991). Such techniques can be categorised under the heading of modelling *realisations* of the spectrum.

We now briefly discuss these general fitting techniques under our headings (i) to (iii) above.

2.2.1 Multivariate techniques

One of the most widely used group of techniques in the investigation of spectroscopic data is multivariate analysis. The approach incorporates a number of useful procedures which are well described in the literature, although not always particularly well differentiated - factor analysis and principal component analysis is a notorious example. As an illustration of multivariate techniques in spectroscopic analysis we therefore consider the factor analysis/ P.C.A approaches. Principal component analysis searches for the best fitting orthogonal axes to replace initially correlated variables, Dillon and Goldstein (1984). The vector of principal components can be described by the following:

$$\mathbf{Y} = \mathbf{P}'\mathbf{X} \qquad (2.1)$$

where **Y** corresponds to the vectors of orthogonal principal components, the

elements of **P** are contributions of the i^{th} variable to the j^{th} component and the **X** vectors are the original correlated variables. Principal components analysis (PCA) is a variance-oriented method, where the components, (typically linear combinations of the original variables), are usually determined by obtaining the eigenvalues of the covariance or correlation matrix, Chatfield and Collins(1980). The correlation matrix is used when the variables are from different scales. The benefits of such an analysis can only be seen if the deterministic variables are truly correlated. Factor analysis on the other hand is a method that accounts for the covariance between variables, i.e the factors determine the variables and can be expressed by the following:

$$X_i = \Sigma \lambda_j f_j + e_i \qquad \textbf{(2.2)}$$

where X_i corresponds to the initial variables, λ_j to the factor loadings, f_j to common factors and e_i to the unique factors. The factor loadings in the Factor Analysis (FA) can be obtained using methods such as the *"Principal Factor Method"* or *"Maximum Likelihood Estimators"*, Dillion and Goldstein (1984). The confusion between PCA and FA appears to arise from the use of the correlation matrix and from the estimation procedures used in both techniques. Methods of Factor analyses e.g. the Principal Factor Method may appear very similar to PCA which adds to the difficulty. However in the PFM each diagonal element of the correlation matrix is replaced by the respective variable's communality estimate and the loadings are calculated in a staged process. (A communality measures the amount of variance of a variable accounted for by a set of common factors). If $\mathbf{R}^*$ is the original correlation matrix with the diagonal elements substituted with communalities, then the first factor loading is as follows:

$$F_1 = \sqrt{\lambda_1}\, Y_1 \qquad \textbf{(2.3)}$$

where $\mathbf{F}_1$ is the factor loading, λ_1 is the eigenvalue of $\mathbf{R}^*$ and $\mathbf{Y}_1$ is the

corresponding eigenvector. The second factor loading is estimated using the same procedure with the augmented matrix $\mathbf{R}^*_1$, given by

$$R^*_1 = R^* - \mathbf{F}_1\mathbf{F}_1' \qquad (2.4)$$

This procedure is carried on until the elements of the augmented R* matrix are close to zero. In PCA the diagonal element measures the variable's individual variance or correlation, and is identical to the PFM method only if the communalities are set to one. In the spectroscopic literature there are numerous examples of dimensional reduction, e.g. Piepponen and Lindstrom (1989) use PCA to determine the source of the pollutants Al, As, Cd, Co, Cr, Cu, Fe, Hg, Mn, Pb, Sb, Ti, V in mussel shells from atomic absorption spectroscopy. Further, a special form of PCA called Correspondence Analysis is used for element detection in EELS, Gelsema et al.(1992). (The difference between correspondence analysis and PCA is that in correspondence analysis we deal with the orthogonal projections of the frequencies generated from a contingency table of **X** as opposed to directly orthogonalising the covariance or correlation matrix of **X**.). Other variants of PCA are also common; for example Windig et al.(1990) describe a staged multivariate technique which incorporates the PCA method in order to determine the chemical components of a substance. Subsequently, Windig et al.(1991) also discuss a method which incorporates PCA with visually orientated variance diagrams on data obtained from pyrolysis mass spectrometry of biomass feedstocks. It is difficult, however, to evaluate this particular research since the mathematical description of PCA in this paper seems to be essentially FA, (specifically the Principal Factor Method, see e.g. Chatfield and Collins(1980)).

A more elaborate form of analysis on FA principles has been put forward by Leurgans and Ross(1992), who use Multilinear models to obtain mathematical decomposition of chemical data sets. The multilinear model is an extension of a bilinear model of spectroscopic absorbance data where this is given by

$$A[i,j] = \sum_{f} e_f[i]c_f[j]L \qquad \textbf{(2.5)}$$

where A[i,j] corresponds to the absorption, $e_f[i]$ is the extinction coefficient of a chromophore[1] f at wavelength λ_i and $c_f[j]$ is the concentration of f in circumstance j, Leurgans and Ross(1992). This model is often described as a factor analysis model where the common factors correspond to the chromophores and the extinction coefficients of the chromophores relate to the loadings and is similar to the unilinear model described in equation (2.4). An additional fluorescence factor known as fluorophores can be incorporated into the above model to give a trilinear model, Leurgans and Ross(1992). This model is analogous to a complex parallel factoring model and the parameters of such a model can be obtained using nonlinear regression analysis[2]. The form of the model as a simple extension of equation (2.5) is then

$$A[i,j,k] = \sum_{f} e_f[i]q_f[j]c_f[k] \qquad \textbf{(2.6)}$$

where $c_f[k]$ is the concentration of chromophore f in circumstance k, $e_f[i]$ is the relative absorption cross section of chromophore f at wavelength λ_i^{ex} and $q_f[j]$ is the relative emission detection at wavelength λ_j^{ex}. Leurgans and Ross(1992) discuss the geometry of the model and the difficulty of obtaining good fits and estimates of the parameters in some detail. A criticism of the FA approach in general would of course be the somewhat arbitrary choice of model forms, but the authors justify this, in this in the example given above, (as in others found in spectroscopic literature), by stating that the model form is readily available from prior scientific knowledge.

---

[1] *An isolated functional group that shows absorption of a characteristic nature in the ultraviolet or visible region, is called a chromphore, Chanda(1979).*

[2] *If the derivative with respect to any parameters of a model contains any other parameters, then the model is deemed to be Non-linear, Bates and Wates(1988).*

2.2.2 Spectral Profile Fitting

The second broad category of spectroscopic analyses is Spectral fitting which specifically relates to the process of obtaining a direct functional fit to spectral data and to breaking the spectrum down into individual components using deconvolution techniques. Nonlinear regression is also used here in conjunction with Fourier deconvolution algorithms, with a view to, say, resolving severely-overlapped component bands in protein Fourier Transform infrared (FT-IR), spectra, Rusling and Kumosinki(1992). Deconvolution techniques are also used with combinations of Galatry and Voigt functions, each corresponding to an individually overlapped spectrum and fitted by least squares, Ouyang and Varghese (1989). Nonlinear modelling is also used by Beamson and Briggs(1992) to show how a comparison of spectra for a solid phase polymer and that for the corresponding small molecule in the gas phase can assist interpretation of the polymer data.

To illustrate some of the ideas involved, we discuss briefly the example given by Rusling and Kumosinki(1991). The authors describe proteins as biopolymers consisting of polypeptide chains of amino acid molecules linked in a linear fashion. In biological systems, proteins function as catalysts for life supporting chemical reactions and as structural components of living organisms. In their native state, polypeptide chains fold in a complicated manner which is essential to their biological function. Folding patterns of proteins may be characterised by periodic structures such as helices. Other structural units include a variety of turns, loops and disorder coils. Determination of the way the protein is folded is called secondary structural analysis.

Secondary protein structures can be determined by several types of instrumental methods such as X-ray crystallography, nuclear magnetic resonance and infrared spectroscopy. Since the development of commercial FT-IR spectrometers, methods for analyzing IR data are being developed which give

a high degree of accuracy for determination of global secondary structure of proteins.

Individual peaks for vibrational transitions are severely overlapped in FT-IR spectra. This overlap needs to be resolved before a complete structural analysis of the protein can be made. The spectrum is the sum of a variety of individual component bands which can be identified using a Fourier deconvolution technique, Rusling and Kumosinki(1991), and then applying a nonlinear regression to the following equation:

$$A = \sum_{j} h_j[\exp\{-(x-x_j)^2/2W_j]. \qquad \textbf{(2.7)}$$

where A is absorbance, W_j is peak width, x_j is frequency in cm^{-1}, and h_j is peak height. W_j, x_j and h_j are the parameters optimised by regression.

All the above spectral fitting techniques involve a priori resolution of the spectra whereas it is often more appropriate or convenient to work directly with principal spectral features such as peak intensities. This brings us to the final group of modelling techniques.

2.2.3 Modelling Realisations of the Spectrum

Realisation methods rely specifically on taking spectral features e.g. *peaks* over a range of values of the independent variable, deriving an appropriate model and finding its parameters using a minimisation technique.

Potential Energy Curves

A commonly occurring situation in perturbation spectroscopy, and one of the fundamental problems in vibrational, impedance and piezo spectroscopic data, is the representation of atomic or molecular behaviour in microscopic terms as a function of the experimental variables. Examples include the construction of potential energy curves from the peak intensities, Figure (2.1), Diaz et al.(1991), and identifying the transition occurring from a particular set of piezo-

spectroscopic data, e.g. Do Carmo et al.(1989).

The construction of potential energy curves as in Diaz et al.(1991) has many parallels and variants, such as determining the set of inverse-power coefficients defining the potential equation of a given molecular state, Davies et al.(1990). In this case, the experimental energy peaks are fitted to the potential equation to determine its inertial rotational constants (B_v's). Following the authors, the potential between two atoms as the sum of inverse power terms is given by

$$V(R) = D - \Sigma C_m/R \qquad \textbf{(2.8)}$$

where D is the energy at the molecular dissociation limit, the powers m are positive integers, and the electronic states of the atoms yielded on dissociation determine which terms contribute to this sum. The basis for obtaining estimates of C_m and later B_v is an iterative non-linear least-squares-fitting technique on the peaks of the spectrum.

In the Diaz et al.(1991) example, the authors describe the potential energy curve in molecular terms as a power series expansion in the reduced variable $x = (r-r_e)/r_e$:

$$V(a,x) = a_0x^2[1+\Sigma a_ix^i] \qquad \textbf{(2.9)}$$

where a_i are the parameters to be estimated, V(a,x) is the potential energy curve, r is the internuclear separation and r_e is the equilibrium distance.

Parameter estimates of this equation are given by solving the Schrödinger equation

$$\psi''(x)=Q(\underline{a},E,x)\psi(x) \qquad \textbf{(2.10)}$$

in which $Q(\underline{a},E,x)= V(\underline{a},x)-E$, E being the energy. The best values of the potential parameters $\underline{a}$ are those for which the eigenvalues $E_1, E_2,.....,E_M$ of equation(2.10) agree most closely with the corresponding *experimental* vibrational-rotational energies $E_1, E_2,.....,E_M$. More precisely, we want to obtain $a_0,a_1,.....,a_{N-1}$ $N\leq M$, so that the sum

$$S(N,M)= \Sigma\ (E_j - E_j) \qquad \textbf{(2.11)}$$

is as small as possible. If we change the potential parameters from a to $\underline{a}+\underline{\varepsilon}$, where $\underline{\varepsilon}=\{\varepsilon_0,\varepsilon_1,...,\varepsilon_{N-1}\}$, every eigenvalue will change accordingly and we can expand $E(\underline{a})$ in a Taylor series around $\underline{\varepsilon} = \underline{0}$ as follows:

$$E(\underline{a}+\underline{\varepsilon})=E(\underline{a})+\Sigma E^{(i)}\varepsilon_i+\tfrac{1}{2}\Sigma E^{(i,j)}\varepsilon_i\varepsilon_j+..... \qquad \textbf{(2.12)}$$

where the superscript(i,j,..) stands for the derivative with respect to $a_i,a_j,..$. This sort of expansion is common in perturbation theory. Terms of higher order than two will be negligible since $\varepsilon_i < 1/2$, so that on substitution of equation (2.12) into equation (2.11) we can minimise with respect to $\underline{\varepsilon}$. In this way we obtain an iterative perturbation scheme to improve the potential $V(\underline{a},x)$ systematically. Stages in the procedure are therefore : firstly, to choose a set of potential parameters $\underline{a}$, solve the Schrodinger equation and obtain the required eigenvalues and their derivatives with respect to $\underline{a}$; secondly, to minimize S(N,M) and obtain a set of corrections $\underline{\varepsilon}$ and, finally, to solve the Schrodinger equation with $\underline{a}+\underline{\varepsilon}$ instead of $\underline{a}$. The procedure is then repeated until the correction is less than a specified tolerance. In every step the roots of the system of nonlinear equations

$$\delta S(N,M)/\delta\varepsilon_i = 0, \qquad i=1,2,...,N-1 \qquad \textbf{(2.13)}$$

are obtained by means of the linear Newton-Raphson algorithm, Diaz et al.(1991). The methods of Davies et al.(1990) and Diaz et al.(1991) are thus similar in that they obtain fits of *nonlinear* equations to *realisations* of spectroscopic data. The fitting approaches differ in that Davies et al.(1990) use a direct fitting approach to the potential energy equations while Diaz et al.(1991) fit indirectly using Taylor series approximations to the eigenvalue equations.

Electronic Model Identification

An additional application which involves working with realisations of the spectrum is that of an electronic structural model for the emitting of Metal Ligand Charge Transfer excited states of $Ru(bpy)_3^{2+}$ and $Os(bpy)_3^{2+}$, Kober and Meyer(1984). Here the long term goal is the description of those features that

control the various excited-state properties. The excited-state lifetimes of both complexes are temperature dependent and increase rather dramatically as the temperature decreases. The behaviour has been successfully interpreted in terms of a thermally equilibrated Boltzmann population of several low-lying excited states whose lifetimes are temperature independent. The ultimate test of the localised hypothesis is to develop an electronic structural model which would result in explicit identification of the electron state assignments that are consistent with the various Raman spectroscopic experimental data. The localised excited-states model describes the energy shifts which can be represented by the following *spin-orbit* coupling matrix:

$$\begin{bmatrix} F+K & -\lambda/2 & -\lambda/2 \\ -\lambda/2 & -k & -\lambda/2 \\ -\lambda/2 & -\lambda/2 & G-K \end{bmatrix}$$

where F, K, G are parameters to be estimated which represent the relative energies of the singlet and triplet[3] excited states, and λ is the spin orbit coupling constant.

The eigenvalues and eigenvectors of these matrices then give the relative energies and compositions, respectively, of the excited states in terms of the pure triplet and singlet states. The method of obtaining the parameters of such a model can be considered as reasonably simplistic. Initially known values are taken for λ and K based on related analysis for similar studies, whereas no reliable predictions are available for F and G and these are treated as variables, Kober (1983). Consequently, the relative energies of the states as a function of F and G may be examined over a given range, $(0<|F,G| < 4000\ cm^{-1}$, Kober(1983)).

[3] *Singlet and Triplet states represent the dimensionality of the orbit occupied by the an electron.*

2.2 Summary

It is clear from the preceding discussion that the three general approaches of modelling spectroscopic data span a wide range of applications. It is equally clear that no single 'best' method exists or is universally acceptable. The complexities inherent to spectroscopic data as a whole, the subsidiary factors and variables affecting spectrum production in many cases, all combine to make the modelling of the principal features a non-trivial matter. The use of multivariate techniques is clearly appropriate where the particular application relies heavily on information from a large set of variables which cannot, apriori, be reduced to one or two. On the other hand, dealing directly with the spectral profile is indicated in those cases where component analysis is completed by modelling the spectrum using a Fourier series or the sum of numerous Gaussian curves. Finally, where an understanding of basic electronic properties of the material is required, (e.g. the behaviour of excited states), it is frequently advisable to work with dominant spectral features, which reflect more distinctly the changes taking place at microscopic level.

We now consider in detail the generation of Piezo-spectroscopic data and determine into which of the modelling categories the transition identification problem falls.

Chapter 3

Piezo-Spectroscopic Data

And

Transition Identification

3.1 Introduction

Piezo-Spectroscopic data is generated from the realisations of the energy peaks from stress-luminescence data. The object of applying stress to crystals is to identify the transitions taking place between energy levels of the crystal when a defect is present.

In this chapter we discuss the physical problem, define the terminology used and provide a description of the way in which piezo-spectroscopic data are generated. An outline only of the theory is provided, since this is the subject of a number of texts and papers, e.g. Rudden and Wilson(1980), Myers (1990), Huges and Runcimann (1967) and Mohammed et al.(1982). We also briefly discuss the current method, used to reconcile the experimental results with the theoretical predictions, Hughes and Runcimann (1967). Finally we evaluate the relative merits of the general spectroscopic modelling techniques with respect to the piezo-spectroscopic problem.

3.2 What is a Defect ?

An ideal crystal would consist of an array of atoms which possess a regular

lattice structure within the whole bulk of the crystal, Figure (3.1).

Any departure from such perfect regularity is a defect in the structure.

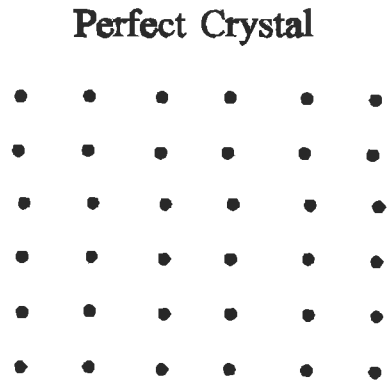


Figure 3.1 Formation of a perfect crystal.

Accordingly, a defect may be:

(i) a simple misplaced atom or group of atoms, within a crystal structure,

(ii) a missing atom, atoms or a group of atoms in an otherwise perfect crystal,

(iii) an atomic impurity or atoms inserted into an otherwise perfect crystal.

For example, point defects, such as atom vacancies, can be described as removing or introducing new atoms into the crystal structure. Figure (3.2) shows an atom being introduced (an interstitial atom) between regular lattice sites.

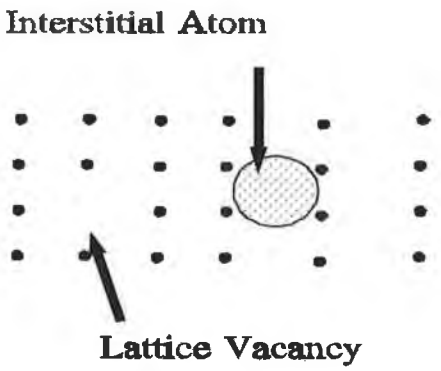


Figure 3.2 Two important point defects are the lattice vacancy and the interstitial atom.

The introduction of an atom into a perfect crystal, for example, will disturb the electronic patterns of the crystal and may cause electrons to move from one orbital to another. By studying the movement of the electrons and hence their energy emissions we can obtain information on the defect. Electron motion can be described using Wave Theory and numerous introductory texts are available in the literature (see e.g. Rudden and Wilson(1980)). We give only a brief synopsis of the principal points in what follows.

3.3 Wave Theory

The description and mathematical formulation of the time-independent Schrödinger wave equation, is given, Rudden and Wilson (1980), by

$$\frac{d^2\psi}{dx^2} + \frac{2m}{\hbar^2}.(E-V)\psi = 0. \qquad \textbf{(3.1)}$$

where m represents the mass,

E the total energy,

V the potential energy of the particle and is equivalent to that estimated in equation (2.9),

$\hbar = h/2\pi$ where h is Planck's constant.

The symbol ψ represents a complex quantity with real and imaginary parts. It is related to the requirement that it must completely represent the motion, and must specify where the electron is and what energy it has at a given time.

Typically, the solution of the Schrödinger equation, equation (3.1) for a particle in a 1-dimensional box is explained using the following assumptions and equations.

$$V(x) = \infty \text{ for } 0 > x > L \qquad \textbf{(3.2)}$$
$$= 0 \text{ for } 0 <= x <= L$$

where L is a constant equivalent to the boundary of the particle and

V is the potential function, and the choice of V(x) =0 between 0 and L is quite arbitrary and only serves as a constant reference level.

In this restricted interval the wave equation becomes

$$\frac{d^2\psi}{dx^2} + \frac{8\pi^2 m}{h^2}.E\psi(x) = 0. \qquad \textbf{(3.3)}$$

Any periodic function will satisfy this equation, and for the sake of simplicity a sine function can be chosen. Thus

$$\psi(x)=A\sin(kx) \qquad \textbf{(3.4)}$$

will be a satisfactory solution, provided that $k=n\pi/L$ where n is a positive integer, since this condition reduces the wave function to zero at both the boundaries. Substitution of this solution into equation (3.2) gives

$$E=\frac{h^2k^2}{8\pi^2m} = \frac{h^2n^2}{8mL} \qquad \textbf{(3.5)}$$

Where n is thus equivalent to the quantum number of the simple Bohr model and effectively determines the energy of the particle.

The model can be extended to 3 dimensions without much difficulty. The following wave equation is proposed

$$\frac{\delta^2\psi(x,y,z)}{\delta x^2} + \frac{\delta^2\psi(x,y,z)}{\delta y^2} + \frac{\delta^2\psi(x,y,z)}{\delta z^2} + \frac{8\pi^2 m}{h}.E\psi(x,y,z)=0 \qquad \textbf{(3.6)}$$

and yields the following solution,

$$\psi(x,y,z)=A\sin(k,x)\sin(k,y)\sin(k,z) \qquad \textbf{(3.7)}$$

where $k_1=n_1\pi/L, k_2=n_2\pi/L, k_3=n_3\pi/L$

Each number n_1, n_2 and n_3 can take an integral value from one to infinity. The energy term is therefore governed by 3 quantum numbers such that

$$E=\frac{h^2}{8mL^2}.(n_1^2+n_2^2+n_3^2) \qquad \textbf{(3.8)}$$

These quantum numbers also determine the form of the wave function which by

convention can be written as $\psi_{n1,n2,n3}$. One important consequence is that several combinations can yield *equivalent* energy values. For example, suppose one of the quantum numbers is equal to 2 the others being unity. Three possible combinations of the quantum numbers lead to the same value of energy, E.

1) $n_1=1$, $n_2=1$, $n_3=2$.

2) $n_1=1$, $n_2=2$, $n_3=1$.

3) $n_1=2$, $n_2=1$, $n_3=1$.

The corresponding wave functions are then

$$\psi_{112}=A.\sin(\pi x/L)\sin(\pi y/L)\sin(2\pi z/L) \quad \textbf{(3.9)}$$

$$\psi_{121}=A.\sin(\pi x/L)\sin(2\pi y/L)\sin(\pi z/L) \quad \textbf{(3.10)}$$

$$\psi_{211}=A.\sin(2\pi x/L)\sin(\pi y/L)\sin(\pi z/L) \quad \textbf{(3.11)}$$

but in each instance the energy is exactly the same

$$E=\frac{6h^2}{8mL^2} \quad \textbf{(3.12)}$$

That is the wave functions ψ_{112}, ψ_{121} and ψ_{211} are said to be *degenerate* in energy. The presence of degeneracy adds complications to the identification of the defect, since we may be unable to determine the number of energy shifts occurring.

The above theory can be used to explain in mathematical terms the background to a transition between two electron waves and hence the corresponding difference in energy.

3.4 Transition in Piezo-spectroscopic Data

3.4.1 General Considerations

Referring to the crystal structure and underlying theory described in the previous section, it is clear that electrons may move between orbitals, gaining or losing energy in the process. For example when an electron makes a p to s transition,

i.e. moves from a p-orbital to the single s-orbital it loses energy, which is equal to the difference in energies associated with the p and s states (orbitals). It is also possible that the electron may go from a low energy state to a higher state as a consequence of excitation. This process is known as "absorption" and is the opposite to that previously described. The pure and perfect crystal will have characteristic *emission* and *absorption* energies. The presence of a defect in the crystal structure changes the emission/absorption energy. The energy emissions can be measured using spectroscopic techniques, providing information on the defect, and on the host material.

Using wave theory we can explain in mathematical terms the background to a transition. A transition can be defined as the change in wave function and hence the difference in energy, as the electron moves from one energy level to another, and can be described in probabilistic fashion by equation (3.10),

$$p=\int \psi_b \, o \, \psi_a dV \qquad \textbf{(3.13)}$$

ψ_a, level one.

ψ_b, level two.

where p is the probability of going from level a to b, **o** is an operator which represents the stimulation or impetus required to move the electron, i.e. the perturbation employed,

V is a perturbational operator and is equivalent to the potential energy of the transition.

The difference in energy is described as E_b-E_a, where E_b and E_a correspond to the energies of the ψ_b and ψ_a wave equations respectively.

3.4.2 Electronic Transitions in Solids

For atoms in crystals, the symmetry of the crystal structure at the site of an atom plays a major role in determining the details of the electron wave functions. Group theory is fundamental to the results, and labels of electron

wave functions retain much of the nomenclature of group theory mathematics. The groups of interest here are those formed by rotations and reflections which leave the defect in the solid unchanged. One dimensional states are labelled A or B, two dimensional are labelled E and three dimensional are labelled T. As we show below, it is the effects of a reduction in symmetry which provide a powerful spectroscopic tool.

Degeneracy can occur as either *orientational* or *electronic* degeneracy. Orientational degeneracy occurs due to geometrical features within the crystal lattice. Electronic degeneracy occurs due to the fact that an electron can lie in equivalent orbitals of equal energy. By applying stress we may change the energy of one or more of the orbitals and hence remove at least some of the degeneracy.

3.5 Applying a Perturbation

3.5.1 Splitting and Shifting

The application of a perturbation such as stress on crystals is a well known technique used in the identification of the transition type in a crystal. Atoms in a particular state are characterised by a particular energy. Thus if E_a denotes the energy level of the atoms when they are in an unexcited or stable state and E_b denotes the energy level of the atoms in an excited state, then the energy emitted in a transition can be denoted by the difference in the energy levels, $E_b - E_a = E$. In the simplest case applying stress to this crystal may cause the energy levels to shift, then we arrive at the situation illustrated in Figure (3.3a).

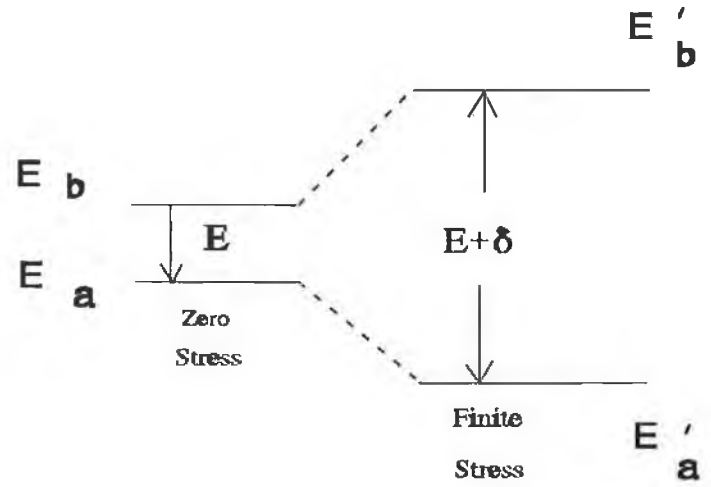


Figure 3.3a Shift in energy between the E_b and E_a levels to the E'_b and E'_a levels.

Here the two energy levels shift to E'_b and E'_a, resulting in an overall energy shift of δ:

$$E'_b - E'_a = E+\delta. \tag{3.14}$$

This shift can be described by the Energy versus Intensity diagram (spectrum) for zero stress being displaced from peak P_0 to P_1 by the application of stress. P_0 and P_1 represent the energy peaks at E and E+δ.

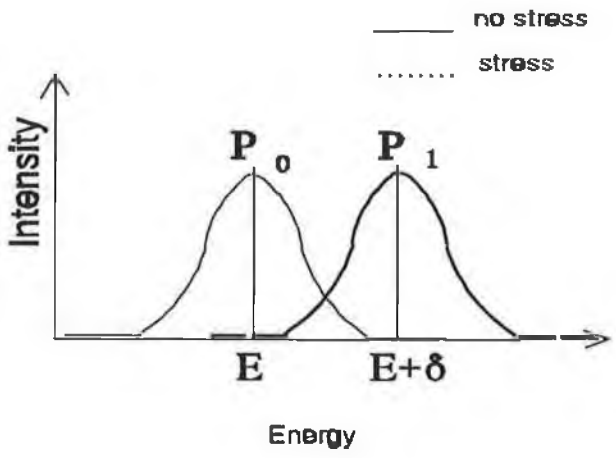


Figure 3.3b Spectrum diagram for the situation described in Figure (3.3a).

It can happen, however, that the line will split as well as shift, Figure (3.3c). In this situation the application of stress has removed some degeneracy within the transition, and the E_b level has split into two levels E'_{b1} and E'_{b2}. The equivalent spectrum will show two peaks at P_L and P_R and will have two energy shifts δ_1 and δ_2, Figure (3.3d).

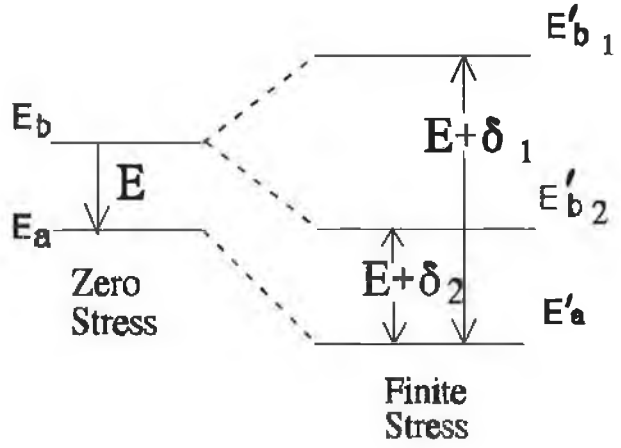


Figure 3.3c The occurrence of a level splitting as well as shifting.

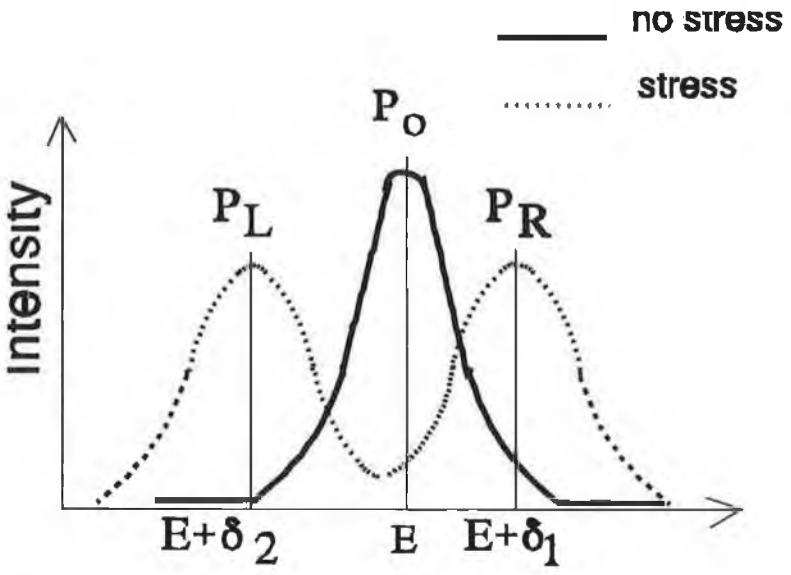


Figure 3.3d Spectrum diagram related to Figure (3.3c), where we see two peaks at P_L and P_R.

3.5.2 Mixing of States

Mixing of states may occur when we have two levels lying closely together.

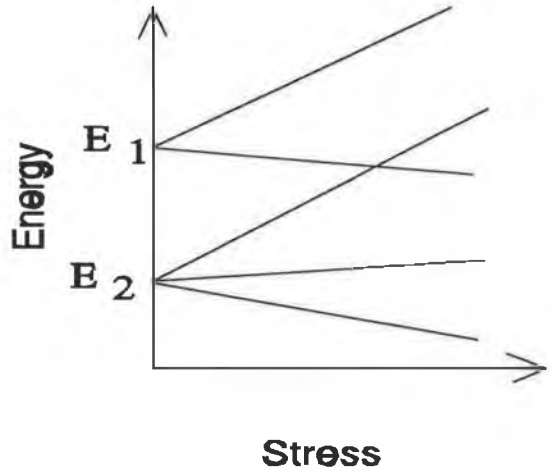


Figure 3.4 Two non-interacting transitions. Initial energies denoted E_1 and E_2.

They can also be referred to as "nearly degenerate" states. In the unstressed situation the spectrum would show more than one intensity peak, e.g. do Carmo et al. (1988). If we apply stress, one of two things may occur.

(I) The two lines might not interact with each other so they split independently. On an Energy difference versus Stress diagram (fan diagram) we would see a stress split component react linearly with increasing stress. Figure (3.4) shows this occurrence, with two sets of different lines indicating that the nearly degenerate states do not have to be the same type.

(II) Alternatively the two states might interact with each other to give non-linear behaviour such as that illustrated in Figure (3.5).

It is also possible that while one line is evident in the zero-stress situation, perturbing influence of stress may reveal that a further state is present, i.e. line-splitting occurs at some stage other than at low stress values. This is explained in terms of the application of stress causing a normally disallowed transition to occur. This is due to mixing in components of states between which

transitions are allowed.

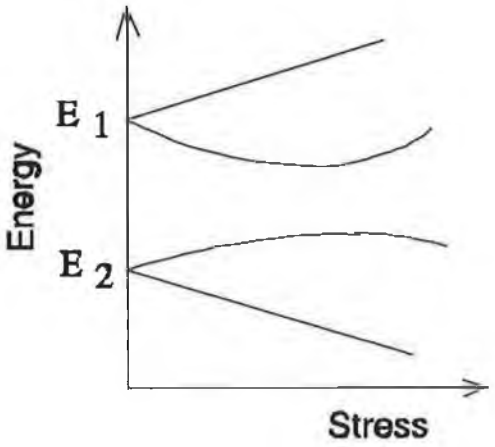


Figure 3.5 Interacting transitions. Again, zero stress positions are at E_1 and E_2.

3.6 Experiment And Theory: Reconciling the theory with Piezo-Spectroscopic Measurements

Even when we know theoretically how the transition occurs it is not necessarily the case that the defect will always give rise to the expected number of lines in an experimental situation. In order to find the experimental splittings, a general rule is to apply stress parallel to the <100>, <111> and <110> directions -shown with respect to the unit cube in Figure (3.6). By applying stress in the above directions we are able to obtain considerable information about the defect involved.

Additionally, given the geometry involved, stress may have to be applied in additional directions in order to obtain a complete description of the splitting. These additional directions however will belong to the same geometrical family as those above.

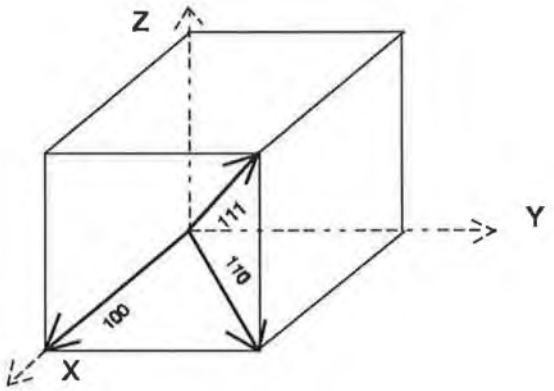


Figure 3.6

An initial description of the geometry of the defect and crystals can be found in Group Theory. Tinkham(1964) describes in tabular format each geometrical group possible. As an example, we consider the pure crystal of cubic symmetry and a trigonal[4] defect, Kaplyanskii(1964), in this crystal which has a symmetry group described by Tinkham and reproduced in Table(3.1). The column headers represent the basis transformations for a particular symmetry, the row headers describe the wave functions that are needed to describe the waves of the symmetry and the table elements correspond to the traces of the transformation matrices.

Table (3.1)

C_{3v}	E	$2C_3$	$3\delta_v$
A_1	1	1	1
A_2	1	-1	-1
E	2	-1	0

C_{3v} is a point group that describes the symmetry elements for a particular set of molecules. In geometrical terms it implies that the trigonal defect has a principal C_3 axis, an axis rotated by 120°, and 3 vertical planes labelled δ_v, which leave

[4] *Trigonal refers to the symmetrical and mathematical properties of the geometry of the defect.*

the defect apparently undistinguishable from its original position, Vincent(1977).

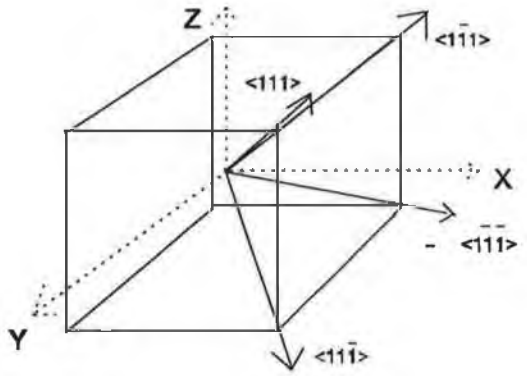


Figure 3.7 Depicts the body diagonals of a unit cube.

The C_3 axis corresponds to any of the body diagonals of the unit cube, Kaplykanskii (1964), and is illustrated in Figure (3.7).

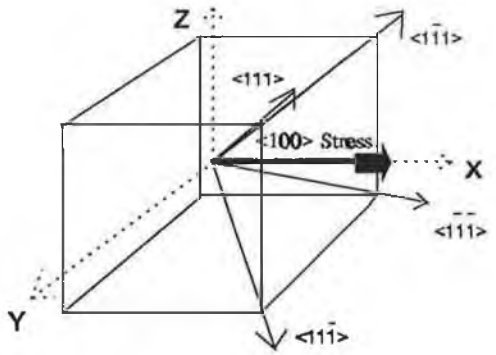


Figure 3.8 <100> Applied stress

Looking at the three families of directions it is clear that under stress from the <100> direction, Figure (3.8), we would expect the above body diagonals to be all perturbed by the same amount, and hence no splitting would occur under stress. Mathematically, by calculating the dot product[5] of the stress direction with the body diagonals, and comparing the result of each calculation, the

[5] *The dot product corresponds to the displacement ratio of one vector upon another.*

number of stress groupings can be noted. So for the <110> direction results are as follows:

$$<110>.<111> = 2/\sqrt{6},$$

$$<110>.<11\bar{1}> = 2/\sqrt{6},$$

$$<110>.<1\bar{1}\bar{1}> = 0$$

$$<110>.<1\bar{1}1> = 0$$

From the above results, Kaplyanskii(1964), it is clear that the four body diagonals split into two groups of two for a trigonal E-> A defect. Equivalently, it can be shown that for the <111> direction the axis will split into one group of one and one group of three. The effects of applying stress in the <110> and <111> directions are illustrated in Figures (3.9) and (3.10).

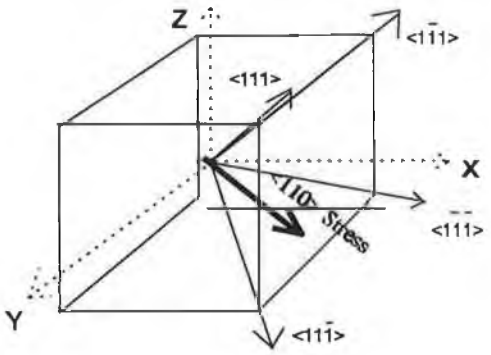


Figure 3.9 <110> Applied stress

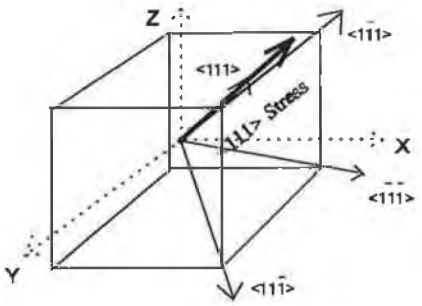


Figure 3.10 <111> Applied Stress

For each defect there is a particular set of directions such that when stress is applied, the results of the energy emissions will help in the identification of the transition. Having applied stress in all relevant directions it is possible to produce three different fan diagrams and from these three cross-sectional views, to identify which transition is occurring. It is known that each transition must belong to a set of specified types. Each defect type is characterised by a particular set of transition types, which are determined by the symmetry

properties of the defect - its "symmetry group".

3.7 Identification of a transition in Piezo-Spectroscopy

3.7.1 Setting-up the Problem

Given a particular set of experimental data we now need to use all available information on symmetry and stress features in order to identify the transition. Table (3.2), Fowler(1968), shows all the possible defects that are permitted by group theory and the transitions that are applicable within each one of these defects. Each diagram on this table describes the possible splitting for any particular transition. Hence for trigonal A->A transition we would expect to see one line in the <100> direction, 2 in the <111> direction and 3 in the <110> direction, and if the experimental fan diagrams showed the above pattern we would presume that the transition occurring was a trigonal A<->A.

A problem with this method is that each set of diagrams does not correspond to a unique transition, i.e. more than one type of transition produces very similar results. In order to obtain an identification of the appropriate transition from the experimental data, the first stage is to use Fowler's tables to identify a set of potential transitions. This is done by comparing the experimental lines to those found in the tables. Having done this it is sometimes possible to further narrow down the search by using polarisation intensities to find the particular symmetry group of the defect in question. However polarisation intensity measurements are often difficult to obtain in luminescence experiments since the light radiating out from the defects tends to get reflected on the internal surfaces of the sample and at the edges so that the polarisation information becomes somewhat scrambled.

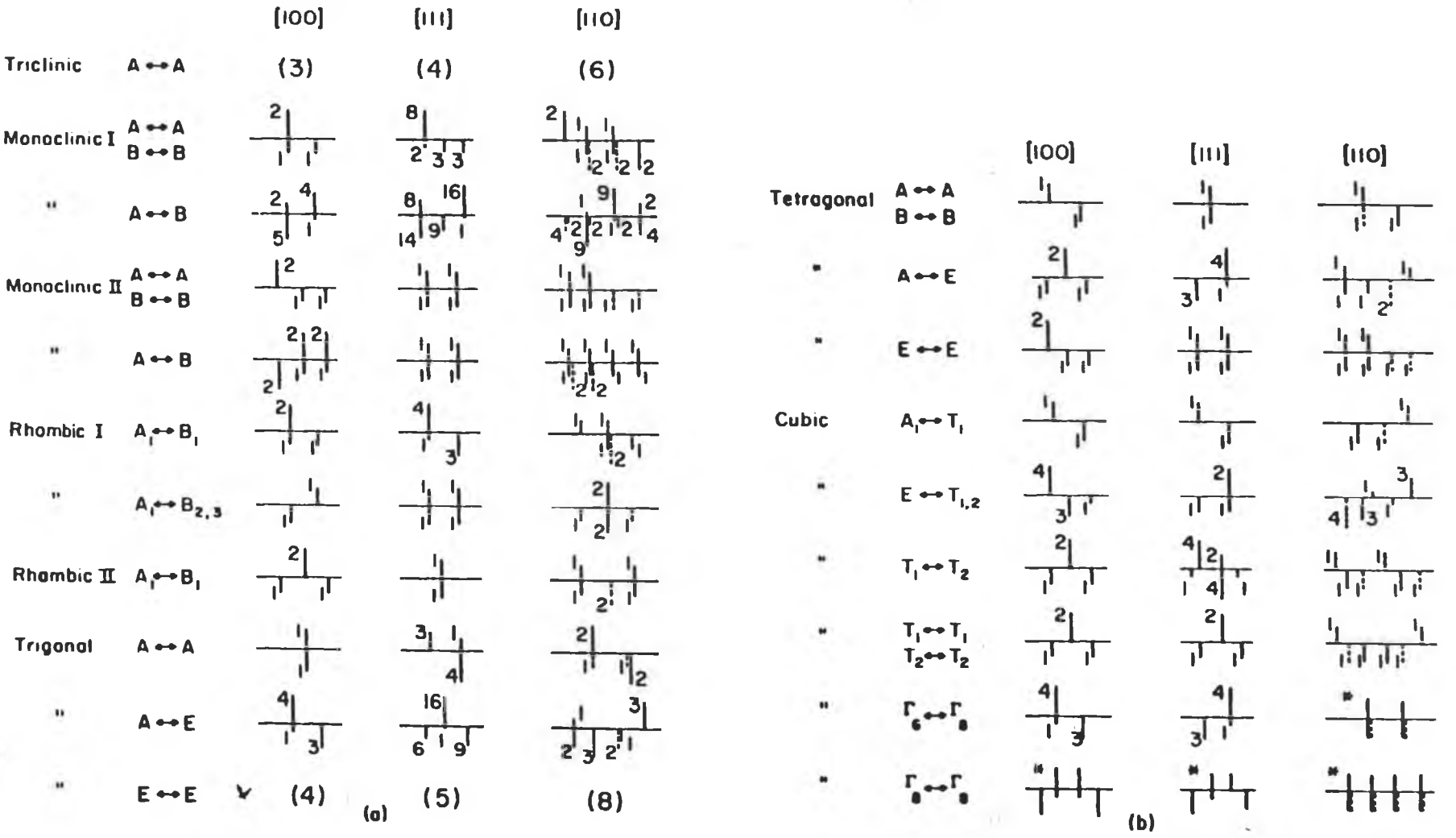


Table 3.2 Schematic splittings of zero-phonon lines under uniaxial stress for various transitions at centres belonging to the eight possible symmetry systems. The spacing of the split lines is arbitrary, and the relative intensities are indicated by the small numerals. The lines above are polarized to the stress, and those below perpendicular to it. The asterisk indicates that intensity rations are not unique.

In addition to the use of polarisation intensities we can derive the theoretical results for each transition and compare them to the experimental data. Apart from the effort involved, this is also not guaranteed to uniquely identify the transition due to experimental error. For each transition type we can assume that a particular stress Hamiltonian exists, equation (3.15). This stress Hamiltonian effectively determines the shift of the initial energy levels when stress is applied and it takes the following form:

$$\underline{H}_s = \underline{H}_o + \underline{V} \qquad \textbf{(3.15)}$$

where $\underline{H}_o$ is the Hamiltonian operator that fully describes the system in the absence of stress and $\underline{V}$ is known as the potential equation, a perturbational operator that describes the effect of stress on the transition, as defined in equation (3.10). For defects of trigonal symmetry, Hughes and Runcimann (1967), the perturbation term $\underline{V}$ is given by

$$\begin{aligned} V = {} & A_1(S_{xx}+S_{yy}+S_{zz}) + A_2(S_{yz}+S_{zx}+S_{xy})+E_x(2S_{zz}-S_{xx}-S_{yy}) \\ & +E_x'(2S_{xy}-S_{zx}-S_{yz}) +\sqrt{3}E_y(S_{xx}-S_{yy}) + \sqrt{3}E_y'(S_{yz} - S_{zx}) \end{aligned} \qquad \textbf{(3.16)}$$

where A_1, A_2, E_x, E_y, E_x', E_y' are parameters to be estimated and represent the dimensional scope of V, $S_{ij} = |P|.\cos(P,i).\cos(P,j)$, $|P|$ is the magnitude of the applied stress and Cos(P,i) is the cosine of the angle between the stress direction and the crystal axes i, McGuigan(1989).

Having decided upon the possible transition/defect type, it is necessary to derive a matrix called a *secular matrix*. Each element of this matrix corresponds to the probability of a transition between two individual energy levels of the complete transition, with its eigenvalues determining the theoretical energy shifts. The potential equation and the operator H_s are used to determine the elements of such a matrix. In the case of a transition from an E state to an A state at a trigonal centre we consider first the influence of the stress Hamiltonian on the two-dimensional E state, Hughes and Runcimann (1967), McGuigan(1989). This can be represented by the secular matrix, Table (3.3).

Table 3.3

H_s	E_x	E_y
E_x	$\langle E_x \mid V \mid E_x \rangle$	$\langle E_x \mid V \mid E_y \rangle$
E_y	$\langle E_y \mid V \mid E_x \rangle$	$\langle E_y \mid V \mid E_y \rangle$

Where E_x and E_y are the basis states of the two dimensional E state and the elements of this matrix take the form

$$\langle E_x \mid V \mid E_x \rangle = \langle E_x \mid A_1 \mid E_x \rangle (S_{xx}+S_{yy}+S_{zz}) + \langle E_x \mid A_2 \mid E_x \rangle (S_{yz}+S_{zx}+S_{xy}) +$$
$$\langle E_x \mid E_x \mid E_x \rangle (2S_{zz}-S_{xx}-S_{yy}) + \langle E_x \mid E_x' \mid E_x \rangle (2S_{xy}-S_{zx}-S_{yz}) +$$
$$\langle E_x \mid \sqrt{3}E_y \mid E_x \rangle (S_{xx}-S_{yy}) + \langle E_x \mid \sqrt{3}E_y' \mid E_x \rangle (S_{yz} - S_{zx}) \qquad \textbf{(3.17)}$$

The eigenvalues of this matrix correspond to the energies of the degenerate levels. We can monitor the effect of the stress on the levels by evaluating the eigenvalues as a function of stress. Obviously it becomes a very arduous task if an equation of form (3.14) must be substituted for each element of the secular matrix in order to obtain the eigenvalues. Fortunately the secular matrix can be simplified by applying a theorem called the "Wigner-Eckart Theorem" to each of its elements, Griffiths (1961).

3.7.2 The Wigner-Eckart Theorem

The elements of equation (3.14) can be simplified to constants or zero by the use of the Wigner-Eckart theorem. This theorem states that

$$\langle E_q \mid O\pi_r \mid E_s \rangle = [E \mid \mid O \mid \mid E] \langle \pi_r E_s \mid E_q \rangle \qquad \textbf{(3.18)}$$

Where $\langle E_q | O\pi_r | E_s \rangle$ corresponds to the probability that a state E_s will go to a state E_q when operated on by the operator $O\pi_r$. The square-bracketed term, $[E| \,|O|\, |E]$ is called the reduced matrix element which represents a generalised probability value, and $\langle \pi_r E_s | E_q \rangle$ is similar to a constant of proportionality between the above two operators, and is also known as the Clebsh-Gordon coefficient or coupling co-efficient. This theorem enables us to identify those same sub-operators ($\langle E_x | V | E_y \rangle$ for example) of V which will have a zero or non-zero contribution to the secular matrix element. If the coupling coefficient for this matrix element is zero then we need not include that sub-operator in the corresponding secular matrix element. If the coupling co-efficient is non-zero then that sub-operator must remain, McGuigan(1989).

Thus for example, the $\langle E_x | V | E_x \rangle$ element in the trigonal E->A secular matrix splits into equation (3.15). Consulting the coupling co-efficients of Griffiths (1961), the sub-operators of this equation can be broken down in the following fashion:

$$\langle E_x | A_1 | E_x \rangle = (E| \,|A_1|\, |E).1 \qquad = A_1 \qquad \textbf{(3.19)}$$

$$\langle E_x | A_1' | E_x \rangle = (E| \,|A_1'|\, |E).1 \qquad = 2A_2 \qquad \textbf{(3.20)}$$

$$\langle E_x | E_x | E_x \rangle = (E| \,|E_x|\, |E).(-1/\sqrt{2}) \qquad = -B \qquad \textbf{(3.21)}$$

$$\langle E_x | E_x' | E_x \rangle = (E| \,|E_x'|\, |E).(-1/\sqrt{2}) \qquad = -C \qquad \textbf{(3.22)}$$

$$\langle E_x | E_y | E_x \rangle = (E| \,|E_y|\, |E).1 \qquad = 0 \qquad \textbf{(3.23)}$$

$$\langle E_x | E_y' | E_x \rangle = (E| \,|E_y'|\, |E).1 \qquad = 0 \qquad \textbf{(3.24)}$$

A_1, $2A_2$, B, and C are all real parameters. Substituting the above results back into equation (3.15) :

$$\langle E_x | V | E_x \rangle = A_1(S_{xx}+S_{yy}+S_{zz})+2A_2(S_{yz}+S_{zx}+S_{xy}) - B(S_{xx}+S_{yy}-2S_{zz}) - C(S_{yz}+S_{zx}-2S_{xy}) \qquad \textbf{(3.25)}$$

If we define $\alpha = A_1(S_{xx} + S_{yy} + S_{zz})+2A_2(S_{yz} + S_{zx} + S_{xy})$,

$$ß = B\,(2S_{xx} - S_{yy} - 2S_{zz}) + C\,(2S_{yz} - S_{zx} - 2S_{xy}).$$

and $\gamma = B\sqrt{3}(S_{xx} - S_{yy}) + C\sqrt{3}(S_{yz} - S_{zx})$.

We can rewrite (3.14)

$$\langle E_x \mid V \mid E_x \rangle = \alpha\text{-}ß, \tag{3.26}$$

similarly we find

$$\langle E_y \mid V \mid E_y \rangle = \alpha+ß, \tag{3.27}$$

$$\langle E_x \mid V \mid E_y \rangle = \langle E_y \mid V \mid E_x \rangle = B\sqrt{3}(S_{xx}\text{-}S_{yy})+C\sqrt{3}(S_{yz}\text{-}S_{zx}), \tag{3.28}$$

$$\langle E_x \mid V \mid E_y \rangle = \langle E_y \mid V \mid E_x \rangle = \gamma \tag{3.29}$$

So we can rewrite the secular matrix as

Table (3.4)

H_s	E_x	E_y
E_x	α-ß	γ
E_y	γ	α+ß

3.7.3 Theoretical Energy Values

By obtaining the eigenvalues of this matrix, obtained from Table (3.4), which coincide with the theoretical energy values, and comparing them to the experimental data we can determine whether the transition is applicable or not. Hughes and Runciman(1967) describe sets of equations which are derived from the differing stress directions for the above matrix, by obtaining its determinant:

$$\begin{vmatrix} \alpha\text{-}ß\text{-}\lambda & \gamma \\ \gamma & \alpha+ß\text{-}\lambda \end{vmatrix}$$

i.e $$\lambda = \alpha \pm (\gamma^2 + \beta^2)^{1/2} \quad (3.30)$$

We need now only calculate the values of λ under the different stresses to obtain the theoretical values. Recall

$$\alpha = A_1(S_{xx} + S_{yy} + S_{zz}) + 2A_2(S_{yz} + S_{zx} + S_{xy}),$$

$$\beta = B(S_{xx} - S_{yy} - 2S_{zz}) + C(S_{yz} - S_{zx} - 2S_{xy}),$$

$$\gamma = B\sqrt{3}(S_{xx} - S_{yy}) + C\sqrt{3}(S_{yz} - S_{zx}), \text{ and}$$

$$S_{ij} = |P| .\cos(P,i).\cos(P,j),$$

So we must apply stress along 5 directions and they are as follows:

(1) Stress along the <100> axis:

$\Rightarrow$ $S_{xx} = 1$ and all other $S_{ij} = 0$

$\Rightarrow$ $$\lambda = A_1 \pm 2B \quad (3.31)$$

(2) Stress along the <111> axis :

$\Rightarrow$ $S_{ij} = 1/3$ for all i,j

$\Rightarrow$ $$\lambda = A_1 + 2A_2 \quad (3.32)$$

(3) Stress along the $<1\bar{1}1>$ axis :

$\Rightarrow$ $S_{xx} = S_{yy} = S_{zz} = S_{xz} = 1/3$

and $S_{xy} = S_{yz} = -1/3$

$\Rightarrow$ $$\lambda = A_1 - 2/3A_2 \pm 4/3C \quad (3.33)$$

(4) Stress along the <110> axis:

$\Rightarrow$ $S_{xx} = S_{yy} = S_{xy} = 1/2$

all other $S_{ij} = 0$

$\Rightarrow$ $$\lambda = A_1 + A_2 \pm (B - 2C) \quad (3.34)$$

(5) Stress along the $<1\bar{1}0>$ axis

$\Rightarrow$ $S_{xx} = Syy = 1/2$, $Sxy = -1/2$

all other $S_{ij} = 0$

$\Rightarrow$ $$\lambda = A_1 + A_2 \pm (B + 2C) \quad (3.35)$$

We now have a set of equations for each stress direction from which the theoretical eigenvalues can be calculated. However to do this, estimates of parameters A_1, A_2, B, C for the trigonal defect must be obtained.

3.8 Choosing an Appropriate Model for Piezo-Spectroscopic data

3.8.1 Multivariate and Profile Methods

We now consider the applicability of the fitting methods discussed in Chapter 2 to the problem outlined above. The questions to be addressed are, firstly, whether the information available from a knowledge of the physical theory and the experimental results support modelling through multivariate or spectral forms. For example, the need to build a tri-linear model implies that more than one independent explanatory variable has a major influence on response. For all such examples complete information of the spectrum must be generated and not just realisations. Secondly, how useful is a complete analysis of spectral components in terms of identifying piezo-spectroscopic transitions. Even if variables other than the direct perturbation (stress in this case) were considered, the model forms seem unnecessarily complex, (Leurgans and Ross (1992) and equation (2.6)), and do not yield directly solutions to equations of the form given in Section 3.7.3.. It would appear, similarly, that spectral fitting provides considerable information which is not directly relevant to a solution of the given problem. It is thus not obvious that full information is necessary or even desirable in obtaining the parameter estimates which describe the transitions.

For Piezo-spectroscopic data the key information in identification is based on the definite peak intensities and the analysis, thus falls most readily into the broad category of method (iii), (section (2.2.3), depending on limited

information or realisations of the spectrum. In support of this view, the testing of different model forms is the basis of transition identification since, due manly to degeneracy and hidden transitions, these are unlikely to be known exactly before fitting takes place.

3.8.2 Using Realisations of the Spectrum

While, the application of the more general multivariate and spectral profile analyses to piezo-spectroscopic data is unattractive, however non-linear modelling still has a role to play, e.g. Davies et al.(1990) and Diaz et al.(1991). The data described by these authors has many similarities to data generated in piezo-spectroscopy, not least multiple emissions or multi-response for a given value of the independent variable (section 3.7.3). Additionally the form of the model is generally known in outline only and must be ascertained in both cases. The parameters of the problem discussed by Davies et al.(1990) are determined directly by applying a least-squares technique, where in any iteration of the fitting procedure, it is necessary to know the partial derivatives of each of the calculated quantities with respect to each of the parameters being fitted. Additionally the authors describe how to obtain initial values for such a procedure, although this is subject to alteration depending on the application and functional form.

Diaz et al.(1991), give an alternative *indirect* fitting technique based on obtaining a Taylor series of the eigenvalues and deriving values of the increments, $\underline{\varepsilon}$, (equation (2.12)), needed for the individual parameters in order to minimise the sum of squares.

Both methods rely on minimising the sum of squares of a system of non-linear equations and in order to do this the derivatives of these equations with respect to the appropriate parameters need to be derived. Both methods could be applied to piezo-spectroscopic data if it were not for the intrinsic problems of missing

data, degeneracy and non-continuous measurements. It is with the inclusion of these aspects that both methods fail, since degeneracy implies that experimental lines have to be matched to actual lines before a sum of squares function can be calculated, which is far from trivial when missing points are taken into account since associating points to lines increase the model complexity. In the following chapter we discuss in detail the additional difficulties which arise due to these features.

Finally, the procedure used in Electronic model identification is similar to that applied in conventional piezo-spectroscopic data, Hughes and Runciman (1967), which relies on the selection of different initial parameter values for one or more particular stress directions. These values are manipulated until a reasonable fit is found. Having established a "best" fit for one particular direction, these parameter values are fitted to the secular matrix for the other stress directions and a revision of the parameter estimates is then made on the basis of the overall fit. However the Electronic model identification problem differs from that described for piezo-spectroscopic data in that no interaction terms are included i.e there is no mixing of states. This allows one to easily develop eigenvalue equations and hence the fitting is relatively straightforward. Piezo-spectroscopic data, however, frequently deals with matrices of much higher dimension and a correspondingly greater level of complexity.

3.9 In Summary

The approach outlined by Hughes and Runcimann (1967) has the merit of simplicity but can often in practice take many manual iterations to achieve a satisfactory fit. Furthermore, matrices of a very high order clearly involve a prohibitive number of such manipulations. It appears that the inaccuracies are very large and the mathematical framework is less than ideal.

It is clear, nonetheless, that the problems discussed by Davies et al.(1990) and Diaz et al.(1991) represent the more suitable modelling approach, rather than those utilising multivariate and spectral fitting techniques. In order to investigate in more detail the appropriateness of spectral realisation methods and hopefully to improve upon them, we must be clear as to the objectives of the analysis. Clearly a method is required which generates for piezo-spectroscopic data an overall solution for *all* the stress directions and is capable of dealing with large matrices. It is also essential to obtain in a short period of time, the solutions for *different* defect types for a specific experimental data set, since the transition will probably *not* be obvious unless the experimental data are very good. In the following chapter, we address the question of finding an appropriate model structure and method of solution which fulfil the above requirements for any transition problem. We include a brief discussion of the advantages and disadvantages of implementing the various approaches dealt with here.

Chapter 4

Method of Fitting The Data

4.1 Introduction

It is evident from the discussion of models and methods in Chapter 2 that these cover a wide range over the many fields in which spectroscopy is a common tool. There is rarely a single universally used method even for one particular application and this is also for piezo-spectroscopic analysis. The method of Hughes and Runcimann(1967), has been described briefly and is *perhaps* the most commonly used but can only be described as sufficient for problems with secular matrices of limited dimensions. In order to develop a more general approach we start by considering the limitations of this method. The primary drawback is that it requires the manipulation of the parameters with respect to each *individual* stress direction, and the review of these estimates with respect to the remaining stress directions. This manipulation involves a long tuning period when approaching an acceptable fit. Such a method which may be described as an heuristic approach, Taha (1987), neither guarantees optimality nor achieves particularly good precision.

To arrive more systematically at a global solution the method of transition identification developed should ensure the generation of an overall solution by obtaining simultaneous parameter estimates for all stress directions and have the capability of dealing with large secular matrices.

4.2 The Global Fit

The fitting procedure currently used, section (3.8.2), has been identified as an integral part of the overall transition identification process, Figure(4.1). Nevertheless its effectiveness when dealing with multiple transitions is limited. This problem can however, be resolved by the procurement and implementation of suitable fitting procedures. We now discuss the development of such piezo-spectroscopic models.

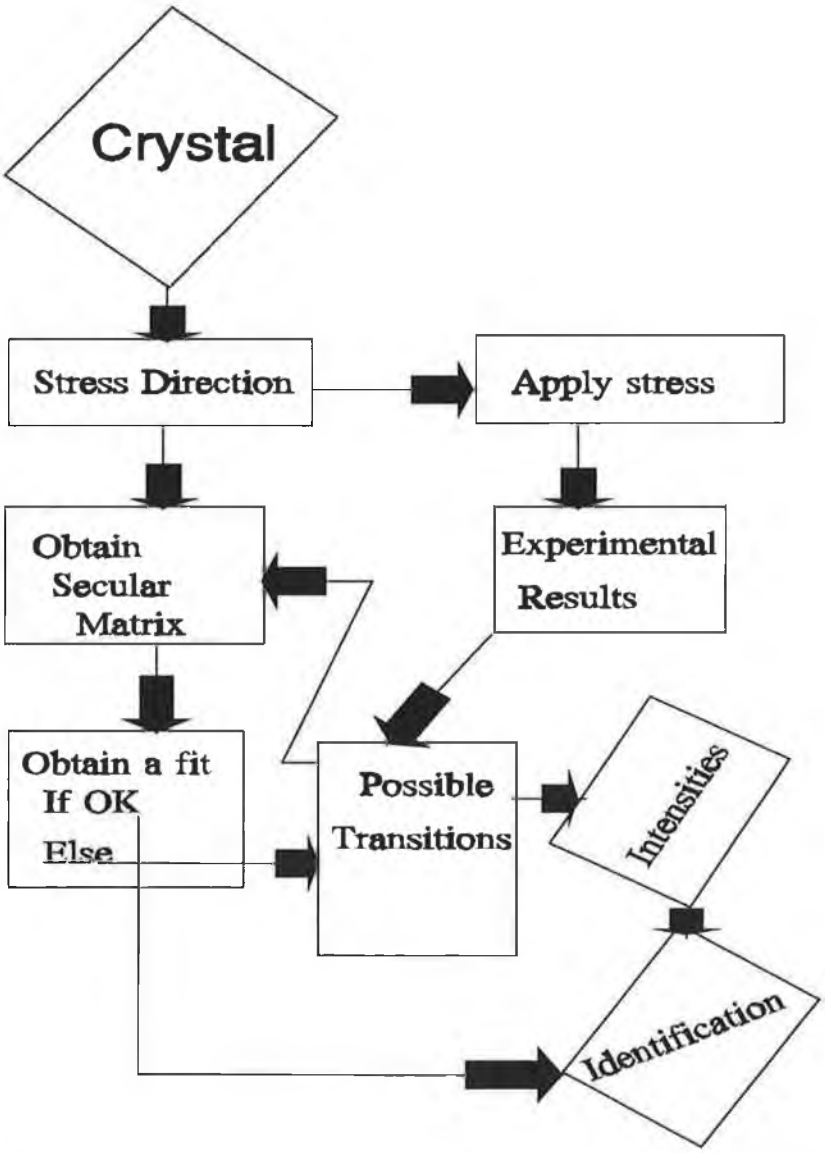


Figure 4.1

Clearly, we need to minimise the differences between the experimental energy values and the eigenvalues generated from the suspected secular matrix to achieve a good fit. One of the simplest solutions to such problems where there is a single dependent variable, is the method of Ordinary Least Squares (OLS), Neter and Wasserman(1974). However the problem outlined in the previous chapter shows one distinct difference, in that it requires a method which deals with multiple dependent variables. This feature relates to the fact that for each stress value within a particular stress direction, several energy values are possible, so that a multidimensional model, Judge et al.(1985), seems indicated. Any proposed method must not only be able minimise the Sum of Squares (SSq) between the experimental and theoretical energy values, with respect to the parameters of the secular matrix, but deal with all stress directions at the same time. The SSq function for such a model will take the following form:

$$SSq = \sum_i \sum_j \sum_k (Y_{ijk} - E(Y_{ijk}))^2 \qquad \textbf{(4.1)}$$

where Y_{ijk} coincide to the i^{th} energy value at a stress value of j *Mpa*[1] for the k^{th} direction, and $E(Y_{ijk})$ corresponds to the expected energy values or eigenvalues of the proposed matrix. Methods used to minimise equation (4.1) require the nature of the relationship between each of the Y variables to be established. Recall the matrix for a trigonal E->A symmetry, Hughes and Runciman(1967), specified in Table (3.3). We can generate the eigenvalues of this matrix by obtaining its determinant, section (3.7.3), which is equivalent to equation (3.30). Equation (3.30), reproduced for convenience, (equation (4.2)), is of appropriate form to describe the corresponding energy value for a specified stress direction.

$$\lambda = \alpha \pm (\gamma^2 + \beta^2)^{1/2} \qquad \textbf{(4.2)}$$

where

$$\alpha = A_1(S_{xx} + S_{yy} + S_{zz}) + 2A_2(S_{yz} + S_{zx} + S_{xy}),$$

$$\beta = B\,(2S_{xx} - S_{yy} - 2S_{zz}) + C(S_{yz} - S_{zx} - 2S_{xy}),$$

[1] *Mpa or Mega Pascals is a unit as a measurement of stress*

$$\gamma = B\sqrt{3}(S_{xx} - S_{yy}) + C\sqrt{3}(S_{yz} - S_{zx}), \text{ and}$$

$$S_{ij} = |P| .\cos(P,i).\cos(P,j),$$

and λ corresponds to the theoretical energy values. From section (3.7.3) there are five equations, equations (3.31) - (3.35), that are derived from equation (4.2) which describe the defect in every individual stress direction. From section (3.8.2) we see that a method which will globally solve equations (3.31) - (3.35) is required. The form of equations (3.31)- (3.35) is linear and the set can be described as a multidimensional linear model, Johnston(1984). Generally however, the form of a multidimensional linear model has the following functional specification

$$Y_1 + \beta_{12}Y_2 + \gamma_1 X_1 + \gamma_{12}X_2 = \mu_1. \qquad \textbf{(4.3)}$$

$$\beta_{12}Y_1 + Y_2 + \gamma_{21}X_1 + \gamma_{23}X_3 + \gamma_{24}X_4 = \mu_2. \qquad \textbf{(4.4)}$$

where Y_i correspond to the eigenvalues of the secular matrix,

X_i correspond to possible applicable stresses,

β and γ correspond to the parameters of the secular matrix and

μ_i corresponds to a stochastic error between the experimental energy values and the predicted eigenvalues.

Although equations (4.3)-(4.4) are not immediately comparable to equations (3.31) - (3.35) its reduced form[2] seems more appropriate, Johnston (1984).

From equations (4.3)-(4.4) the Y variables are of current interest, the X variables are pre-determined and each equation can be described as a structural relation. The general form of this model can be given by equation (4.5)

$$By_t + \Gamma x_t = u_t \quad t = 1,...,n. \qquad \textbf{(4.5)}$$

Where B is a GxG matrix of co-efficients of current variables, Γ is a GxK matrix of coefficients of predetermined variables, and y_t, x_t and u_t are column

[2]*The reduced form is obtain by solving the model so as to express each current endogenous variable solely in terms of exogenous variables and lagged endogenous variables. The terminology originates from the more usual applications of simultaneous models.*

vectors of G, K, and G elements respectively, Johnston(1984).

Assuming that B^{-1} exists, the reduced form of equation (4.5) is

$$y_t = \pi x_t + v_t \quad t = 1,...,n \qquad \textbf{(4.6)}$$

where π depends on the matrix co-efficients for y_t and x_t i.e. B and Γ respectively such that $\pi = -B^{-1}.\Gamma$ and $v_t = B^{-1}.u_t$.

In general there will be an infinity of B and Γ structures corresponding to any given π matrix. We can illustrate however with the Trigonal E->A symmetry case outlined earlier. Recall equations (3.31) - (3.35) with an extra term on each equation in order to incorporate stress and with each equation describing the individual energy shifts, y_t;

where P|| corresponds to stress parallel to a particular axis, A_1, A_2, B and C are the parameters to be estimated.

(1) P||<100>

$$\Rightarrow \quad y_1 = (A_1 - 2B).S \qquad \textbf{(4.7)}$$

$$\Rightarrow \quad y_2 = (A_1 + 2B).S \qquad \textbf{(4.8)}$$

(2) P||<111>

$$\Rightarrow \quad y_3 = (A_1 + 2A_2).S \qquad \textbf{(4.9)}$$

(3) P||<11$\bar{1}$>

$$\Rightarrow \quad y_4 = (A_1 - 2/3A_2 - 4/3C).S \qquad \textbf{(4.10)}$$

$$\Rightarrow \quad y_5 = (A_1 - 2/3A_2 + 4/3C).S \qquad \textbf{(4.11)}$$

(4) P||<110>

$$\Rightarrow \quad y_6 = (A_1 + A_2 - (B - 2C)).S \qquad \textbf{(4.12)}$$

$$\Rightarrow \quad y_7 = (A_1 + A_2 + (B - 2C)).S \qquad \textbf{(4.13)}$$

(5) P||<1$\bar{1}$0>

$$\Rightarrow \quad y_8 = (A_1 + A_2 + (B + 2C)).S \qquad \textbf{(4.14)}$$

$$\Rightarrow \qquad y_9 = (A_1 + A_2 - (B + 2C)).S \qquad \textbf{(4.15)}$$

Each equation corresponds to a particular stress direction, with each eigenvalue representing a predicted energy value. Formulating with respect to the general problem we see that the B matrix will correspond to a 9x9 identity matrix. This implies that :

$$\pi = \Gamma$$

With the only pre-determined variable being the stress component, implying K=1. Thus in the case of transition identification, equation (4.6) becomes

$$y_t = \pi.S_t + v_t \quad t=1,...,n. \qquad \textbf{(4.16)}$$

where S_t denotes the stress component, the predetermined variable ,

y_t corresponds to the current variable,

$v_t = u_t$ since B is an identity matrix, and

π takes the following form:

$$\pi^T = [a \;\; b \;\; c] \qquad \textbf{(4.17)}$$

where a,b and c take the following form;

$$a^T = [A_1+2B \;\; A_1-2B \;\; A_1+2A_2] \qquad \textbf{(4.18)}$$

$$b^T = [A_1-2/3A_2+4/3C \;\; A_1-2/3A_2-4/3C \;\; A_1+A_2+(B-2C)] \qquad \textbf{(4.19)}$$

$$c^T = [A_1+A_2-(B-2C) \;\; A_1-A_2+(B+2C) \;\; A_1-A_2-(B+2C)] \qquad \textbf{(4.20)}$$

Techniques for solving this type of problem are well known in fields such as Econometrics, Judge et al.(1985), and include variations of least squares and maximum likelihood which come under the broad headings of limited and full information estimation. For the estimation of a structural equation in limited information techniques, complete information on all other structural equations

in the model is not taken into account, whereas with full information techniques all such equations are considered together or simultaneously.

The general form outlined earlier leads to a statistical identification problem, where obtaining unique estimates of B and Γ from the estimation of π will prove to be impossible unless an appropriate number of restrictions can be incorporated, Johnston(1984). Information on the secular matrix or knowledge of the boundary conditions is required in order to uniquely identify the set of parameters. Do Carmo et al.(1988) describe a transition identification where the energy shift and the effect due to spin orbital are measured directly from the data.

Various approaches to obtaining solutions for the secular matrix are now discussed with a view to using the formulation above to obtain the set of parameter estimates.

4.3. Characteristic Polynomial

The characteristic polynomial method entails building up a system of equations for each applicable stress direction, which correspond to the characteristic polynomials of the secular matrices of each direction.

We refer in what follows to our previous example of a trigonal E->A, Kaplyanskii(1964), Table (3.3). We can define the characteristic polynomial Q for this secular matrix by obtaining the determinant, where this is given by

$$Q = \det \left| H_s - \lambda_i . I \right| \qquad \textbf{(4.21)}$$

Here I is the identity matrix and the λ_i's are the experimental energy values. From equation (4.21) the characteristic polynomial can be given as follows :

$$(\alpha-\beta-\lambda_1)(\alpha+\beta-\lambda_2)-\gamma^2 = 0 \qquad \textbf{(4.22)}$$

with λ_1 and λ_2 as the experimental energy values,

α,ß and γ denoted as in equation (4.2).

This characteristic polynomial can be manipulated in such a way as to give a regression equation with dependent variable $\gamma_1.\gamma^2$ and independent variables stress, γ_1 and γ^2, for which we can estimate the parameters of the secular matrix. Before this can happen all unique stress directions must have their characteristic polynomial derived. From Chapter 3, the defect of a trigonal transition lies along the body diagonals of a crystal cube, which as already described, implies there are five unique stress directions, 100, $1\bar{1}1$, 111, $1\bar{1}0$ and 110. Each one of these stress directions has a characteristic polynomial similar to equation (4.22). Each equation has a number of unknown parameters, A_1 A_2, B, C and three independent variables, stress, λ_1 and λ_2. It can happen that λ_1 and λ_2 have the same value due to degeneracy leading to difficulties in allocating energy values to the eigenvalues of matrices of order greater than 2. Thus in general (4.22) becomes

$$(\alpha_x-\beta_x)^2-(\alpha_x-\beta_x)(\lambda_{1x}+\lambda_{2x})-\tau^2{}_x=\lambda_{1x}\lambda_{2x} \qquad \textbf{(4.23)}$$

Where x takes the values 100, $1\bar{1}0$, 110, $1\bar{1}1$ and 111 .

From equation (4.23) we can formulate the complete system given by the expression

$$\prod_{i=1}^{d}\lambda_i = f(\text{stress},\underline{a},\lambda_i)+\underline{\varepsilon_i} \qquad \text{for } (i=1,\ldots,n) \qquad \textbf{(4.24)}$$

where λ_i are the experimental energy values,

$\underline{a}$ are the parameters to be estimated,

$\underline{\varepsilon_i}$ are the errors corresponding to each regression equation,

n is the number of stress directions applicable

and d is the dimension of the secular matrix.

Expression (4.24) describes a system of equations with common parameters. Techniques applicable to this general class of models, Judge et al.(1985), can

thus in theory be used to obtain fits on the experimental data for various transition possibilities. However, difficulties can arise in terms of associating the experimental energy values (embedded in equation (4.24)) with the theoretical values describing the transition. Further details on the equation structure are clearly desirable.

4.4. Derivation of eigenvalue equations

This method is an extension of the characteristic polynomial method, where we seek to solve the system of eigenvalue equations generated for each stress direction, instead of the characteristic polynomial of the matrix. Noting in equation (4.24) that λ_i is included in the function statement, we can reformulate the equation such that

$$\underline{\lambda_i} = R(\ \underline{stress_i}\ ,\ \underline{a}\ ) + \underline{\varepsilon_i} \qquad (i=1,..,n) \qquad \textbf{(4.25)}$$

where $\underline{\lambda_i}$, $\underline{a}$ and $\underline{\varepsilon_i}$ are defined as previously.

This is effectively an equation that gives the eigenvalues for a particular secular matrix and proves more useful than the previous method, since the problem of allocation of energy values to eigenvalues is lessened as we have more knowledge of each equation.

For a low dimensional problem, e.g. the trigonal A->E, McGuigan(1989), we can fairly readily use the eigenvalue approach to fit such a system of equations to the experimental data lines. Mohammed et al.(1982) present a detailed tabulation of the effect of uniaxial stresses on photoluminescence transitions occurring at optical centres found in cubic crystals. The elements of this table depict the eigenvalue equations that are appropriate for each transition and are

equivalent to the form outlined above. An example of the format of these tables for a monoclinic I A->A symmetry can be seen in Table (4.1).

Table 4.1

Type of Centre	S||<001>	S||<111>	S||<110>
Monoclinic I A-A	A_1	$1/3(A_1+2A_2+2A_3)$	A_2+A_3
	A_2	$1/3(A_1+2A_2-2A_3-4A_4)$	$1/2(A_1+A_2-2A_4)$
		$1/3(A_1+2A_2-2A_3+4A_4)$	$1/2(A_1+A_2+2A_4)$
			A_2-A_3

Each equation is multiplied by the stress applicable and is equivalent to the energy shift generated by the stress applied. Hence to obtain an appropriate fit, simultaneous regression techniques would be quite feasible, although recovering the desired parameters from the reduced form of the model is again a problem. Additional complexity e.g. looking at the identification of transitions which involve the mixing of states, (section (3.6)), exacerbates this difficulty. As an example of the difficulties of identification in a transition which involves mixing of states we consider the situation where we have *two* E->A trigonal states. Each individual state will have a secular matrix like that described in Table (3.3). However if the experimental data show the possibility of an interaction between the states[3], then the corresponding secular matrix that represents this situation will have to accommodate both transitions, McGuigan(1989).

Table 4.2

Hs	Ex	Ey	Ex'	Ex'
Ex	$\alpha-\beta$	γ	$\langle E_x \mid V \mid E_x' \rangle$	$\langle E_x \mid V \mid E_y' \rangle$
Ey	γ	$\alpha+\beta$	$\langle E_y \mid V \mid E_x' \rangle$	$\langle E_y \mid V \mid E_y' \rangle$
Ex'	$\langle E_x' \mid V \mid E_x \rangle$	$\langle E_x' \mid V \mid E_y \rangle$	$\alpha'-\beta'$	γ'
Ey'	$\langle E_y' \mid V \mid E_x \rangle$	$\langle E_y' \mid V \mid E_y \rangle$	γ'	$\alpha'+\beta'$

[3] *An interaction between two states is usually represented by a non-linear fan diagram.*

Table (4.2) demonstrates this situation, where we can see that the diagonal blocks are similar to that of Table (3.3). The off-diagonal blocks are similar to the diagonal blocks in functional form but will possess different parameters. They represent the amount of interaction that can occur between the two states. Hence if off-diagonal terms are zero, we expect the two states to have no interaction and behave linearly. If, however, they are not zero then we expect an interaction to occur and the stress energy relationship, will be non-linear.
In summary, the additional feature of mixing of states further complicates the use of the eigenvalue equations and characteristic polynomials as a means of fitting experimental data.

4.5. Powell-Shell : Method of Solution

To date the approaches studied in the identification of transition data have dealt specifically with generating individual relationships between the experimental energy values and the functions proposed to evaluate their corresponding eigenvalues. In the following we describe a method which enables us to obtain simultaneous parameter estimates for the defect of interest without having to derive either eigenvalue equations or characteristic polynomial equations. This reduces analytical processing time compared to that needed for the methods described in sections (4.3) and (4.4). Even with modern mathematical manipulators, the time taken to derive the characteristic polynomial for a high order matrix can be prohibitive. The approach described here also resolves the problem of parameter identification because we deal directly with the parameters within the secular matrix. We have denoted it the "Powell-Shell" algorithm.
The underlying principle is again to obtain a match for the eigenvalues of a

potential[4] secular matrix with the actual experimental energy values.

4.5.1 Methodology :-

For each transition type there is a corresponding secular matrix. However when transitions interact it is far harder to identify the transition types which are actually involved. Using the transition identification scheme illustrated in Figure (4.1) we can narrow the selection of transitions that are applicable and hence find the best fit for each potentially applicable secular matrix . Using the Powell-Shell as a fitting technique we require:

1) To match each eigenvalue generated from the secular matrix of the suspected transition type to a corresponding experimental value. This is necessary, since degenerate problems often hide many of the characteristics of the transition occurring and hence imply a non one-to-one correspondence between the experimental and theoretical domains.

2) Rank each experimental and eigenvalue (predicted value) according to the highest value and accordingly deal with the problem of missing data, since (as noted in Chapter 2) the application of stress does not always produce a measurable result corresponding to the suspected transition, (even where this transition is appropriate). An extreme case of such an occurrence is where a transition interacting with another transition is completely un-observable within the experimental stress boundaries. An example of such an occurrence is given by the data set and solution described in section (4.6.3).

3) Obtain a fit between the experimental and predicted lines of data.

4) Acquire "good" initial estimates in order to start the algorithm. (Clearly convergence is more rapid where initial estimates are closer to actual values. However, we demonstrate that even relatively poor initial estimates give

[4] *It is probable for complex cases in particular that more than one transition may be examined and hence more than one secular matrix may be fitted.*

convergence within a reasonable period using this approach).

4.5.2 Matching Problem

The presence of degeneracy complicates further the difficulty of matching an unequal amount of experimental points with the predicted eigenvalues generated for an equivalent application of stress.

As an example, we consider the problem of a trigonal centre, with a non-degenerate orbital ground state and two orbital doublet excited states, Do Carmo et al.(1988). This has an 8 x 8 secular matrix but for the <100> direction there are only four energy emissions reported. Since we are trying to obtain a fit between the estimated and actual data points, we need to assign the estimated

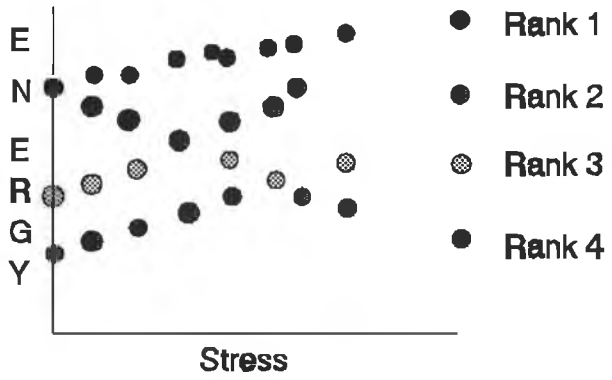


Figure 4.2

lines to the actual line emissions. Obvious difficulties arise in terms of performing this assignment. We realised, however, that the number of comparisons between both sets of lines could be limited by a consideration of the ranks of both the actual and estimated data points. By ranking in such a manner it becomes obvious that we are no longer dealing with lines but with ranked points, and associations will no longer be made according to line membership but by ranked position. Figure (4.2) shows how the actual lines look having been ranked according to highest value. We can see that the highest

ranking estimated data points can only be compared to the highest ranking actual data points. If we compared them instead to the second highest actual data points it would mean that none of the lower ranking estimated data points could be compared to the highest actual data points and hence the algorithm would tend to only estimate the lower values.

Using this technique we are able to draw up a scheme, as shown in Figure (4.3), of the comparisons that have to be checked. So for example, we check the second highest ranked estimated data points, E_2 against the first(A_1) and second(A_2) ranked actual data points and make our decision on the basis of which actual values were closest[5] to the E_2 data points. Figure (4.3) shows the allocation of actual lines to their possible estimated lines, where the E_i's are

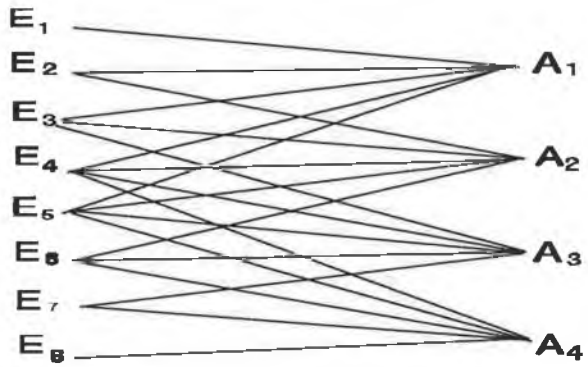


Figure 4.3

estimated energy values of the chosen secular matrix ranked from highest to lowest and A_i's are the actual experimental values. With such a scheme it is obvious that E_1 is associated with A_1, E_2 with A_1, A_2, E_3 with A_1, A_2, A_3 , E_4 with A_1, A_2, A_3, A_4, E_5 with A_1, A_2, A_3, A_4, E_6 with A_2, A_3, A_4, E_7 with A_3, A_4 and E_8 with A_4. We need to complete the above procedure for each stress direction. However for the <110> and <111> directions there are 2 directions

[5] *We determined the closeness of the estimated data points to the actual data points by using the error sum of squares between the two data sets as usual.*

within the particular families of directions. For <110> the two directions are <110> and <1$\bar{1}$0>, and for <111> they are <111> and <1$\bar{1}$1>. Within each family we generate the corresponding eigenvalues. The eigenvalues from each data set within a particular family are then ranked as if belonging to one data set, before we can complete an association to the actual experimental values.

4.5.3 Missing Values

From our description of the above matching procedure, a knowledge of the ranked position of the data values is a fundamental requirement. This however becomes a problem if :

1) some of the actual experimental data points are not measured at high stress values i.e. we have missing data values,

2) complete transitions do not appear within the experimental data.

The traditional solution to problem 1) is to visually interpolate the line to the maximum stress values and then rank accordingly. The accuracy with which this is done depends to a large extent on individual skill. In Chapter six, we describe the development of a graphical computer system which will allow an inexperienced user to obtain accurate ranking data schemes for any particular data set.

For the second difficulty we demonstrate in our system a mechanism that permits inclusion of transitions within the secular matrix that have not been experimentally detected. The basis for such an inclusion is to rank the data available in such a manner that the fitting technique will *recognise* the missing data for the un-observable transition as belonging to the transition which cannot be observed experimentally.

4.5.4 The fitting technique

The fit is obtained by checking each estimated line against each of the actual

lines proposed by the matching technique shown in section(4.5.2). The pairing with the lowest residual sum of squares is accepted and is added to the overall sum of squares (NSSq(P), where P corresponds to the parameter vector). Since we do not have an analytical expression for NSSq(P), due to the ranking involved in the matching procedure, numerical approximation of the derivatives for large and complex problems prove to be inaccurate and involve a large number of extra function evaluations. We discuss two optimisation techniques one based on non-derivative and the other on discrete derivative information. An appropriate algorithm for the optimisation of NSSq(P) when not relying on derivative information is the Powell algorithm, Walsh(1975), Press et al.(1990). Powell's algorithm is not the only one known with that property, but it was selected for its robustness and availability in numerous numerical libraries such as Press et al.(1990) and NAG(1988) and hence we decided it was a good choice here.

Alternatives using discrete derivative information are known as conjugate direction methods of which the Fletcher-Reeves and the Polak-Ribiere, Press et al.(1990), are the two most important algorithms.

4.5.5 The Minimisation Algorithm

The search for the optimum vector P is performed in a number of iterations. Each iteration, which normally involves many evaluations of NSSq(P), consists of p (sometimes p+1) one-dimensional searches in different directions in the p-dimensional space defined by the parameter coordinates, Erup and Harris(1988). The execution of each linear search is illustrated in Figure (4.4). NSSq(P) is first calculated at the starting point (point A) and at a point separated from it by a given step size (B). The initial starting points are user specified. The values of the initial points and step sizes are updated during the optimisation process. If the function value at B is lower than at A, a further step in the same direction

is taken to calculate a third point. If, as is the case Figure (4.6), the value at B is worse(larger) than at A, the third point is calculated at a point one step to the other side of A (point C).

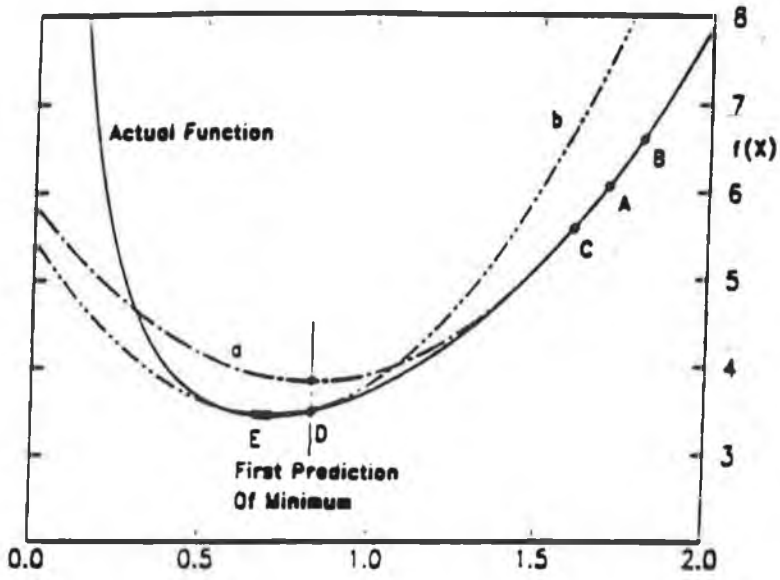


Figure 4.4

Next, a parabola is fitted through the three points (*curve a*). Assuming that this parabola has a minimum, the abscissa thereof is taken as a predictor for the minimum point on the actual curve. The function is calculated at this point (D) and a new parabola (not shown) is fitted using the last point and the best two of the previous points. A new minimum is predicted from this second parabola, a new function value calculated there, etc., until a stop criterion is satisfied, Brent (1973). If we combine the individual (u_i) linear searches, the following algorithm, Press et al.(1990), can be repeated until the function NSSq(P) stops decreasing:

·Save the starting position as P_0.

·For $i = 1,...,p$, move P_{i-1} to the minimum along direction u_i and call this point P_i.

·For $i = 1,...,p-1$, set $u_i \leftarrow u_{i+1}$.

·Set $u_p \leftarrow P_p - P_0$.

·Move P_p to the minimum along direction u_p and call this Point P_0.

where u_i corresponds to a set of N mutually conjugate directions which relate to the direction of each individual linear search and are initialised to the basis vectors.

We emphasise that this method as we said earlier is not based on derivative information. In order to examine the appropriateness of the Powell algorithm a comparison with other algorithms is necessary. Fletcher and Reeves and later Polak-Ribiere, Press et al.(1990), devised conjugate gradients techniques which rely on derivative information in order to obtain the conjugate directions. Although the sum of squares does not posses analytical derivatives a finite difference scheme can be used to calculate the gradients of the sum of squares function. In addition to this we can use the Hessian matrix of the sum of squares function to obtain covariances of the parameter estimates, (discussed further in Chapter 5). However in order to obtain the derivatives necessary for the Fletcher- Reeves algorithm we now outline the differencing scheme used.

4.5.6 Discrete Derivatives Using Finite Differencing Schemes.

The majority of optimisation routines rely on the availability of second derivatives of the model function, Bates and Watts (1988), Seber and Wild (1989) and the Levenbergg-Marquadt technique, Press et al.(1991). In our case, however, because we are unable to complete the derivation of eigenvalue equations for problems with high ordered matrices, we will be unable in each iteration to obtain the Hessian of data that is ranked since the data will be of a discontinuous nature. Additionally, in the least squares technique applied in the Powell-Shell we need to obtain the Hessian of the least-squares equation in order to obtain the covariance matrix.

The lack of information about the analytical form of the least squares equation inherent in fitting our *ranked* data implies that a differencing scheme similar to

that discussed by Smith(1985), should be used to obtain the Hessian of the least squares equation. For example for a function of a single parameter the Taylor series expansion is of the following form

$$U(x+h)=U(x)+hU'(x)+1/2h^2U''(x)+1/6h^3U'''(x)+.... \quad \textbf{(4.26a)}$$

$$U(x-h)=U(x)-hU'(x)+1/2h^2U''(x)-1/6h^3U'''(x)+.... \quad \textbf{(4.26b)}$$

$$U(x+h)+U(x-h)=2u(x)+h^2U''(x)+O(h^4) \quad \textbf{(4.27)}$$

where U(x) is the function being differentiated and h is the differencing scheme increment which is <1/2.

If the final terms (h^4 onward terms will become increasingly small if h is ≤ 0.5) tend to zero then

$$U''(x)= \{U(x+h)-2u(x)+U(x-h)\}/h^2 \quad \textbf{(4.28)}$$

this can be expanded for functions of several variables using the following notation where U(x,y) is a function of x and y.

$$U''(x)= \{U(x+h,y)-2u(x,y)+U(x-h,y)\}/h^2 \quad \textbf{(4.29a)}$$

$$U''(y)= \{U(x,y+i)-2u(x,y)+U(x,y-i)\}/i^2 \quad \textbf{(4.29b)}$$

$$U_{xy}''(x,y)=[U(x+h,y+i)+U(x-h,y-k)-2U(x,y)-U_{xx}''(x,y)h^2-U_{yy}''(x,y)i^2]/hi \quad \textbf{(4.30)}$$

where i is similar to h.

The above differencing scheme in conjunction with a conjugate gradient minimisation technique, Press et al.(1990), can be used to minimise NSSq(P) and obtain estimates of P. These methods of Feltcher and Reeves and of Polak and Ribiere rely heavily on the initial estimates given. The Powell algorithm although a relatively robust algorithm, Walsh (1979), also relies to some extent on the choice of initial estimates. We now discuss a procedure for obtaining initial estimates for both methods.

4.5.7 Initial Estimates

In order to achieve an accurate fit using the Powell algorithm the provision of reasonable initial estimates clearly speeds convergence. The initial estimates of

any mathematical model require a knowledge of the possible range of each parameter. Within this range therefore we must decide upon values which are consistent with the physical situation. It should be noted here that if a problem is more complex, requiring non-linear[6] models, it is possible to obtain initial estimates by simplifying the model to a linear form, Bates(1988), either by using a temporarily truncated experimental data set or by looking at one particular transition if mixing of states has occurred. From this we use least squares on the simplified model to obtain initial estimates and hence give a reasonable starting point. Alternatively we can look at contour plots of the sum of squares function in order to obtain parameter estimates corresponding to the sum of squares function, SURFER(1989), Bates (1988). The importance of finding good initial estimates cannot be over emphasised when dealing with non-linearity, since a wrong choice can lead to a local minimum rather than a global minimum and hence to inaccurate estimation. However, the Powell-shell tool we have developed has shown itself to be reasonably robust to even a poor initial choice of parameter estimates. When derivative information is included the minimisation is more volatile as in the case of the Fletcher-Reeves algorithm.

4.6 Analysis

In what follows we compare the results of an analysis using the three potential methods of Sections (4.3)-(4.5) for obtaining solutions to the secular matrix and ultimately identifying the transitions taking place. For the Powell-shell method discussed in Section (4.5) we also outline the performance of the Powell-shell and the least-squares estimators solutions to the problems discussed by do Carmo et al(1988), Campion et al.(1993) and Kehoe (unpublished), and discuss comparisons between the Powell and fletcher-reeves algorithms in Chapter 6.

---

[6] *Non-linear models are models where the parameters as well as the variables can take a non-linear form.*

4.6.1 Characteristic Polynomial & Eigenvalue Equation Methods

Estimating a characteristic polynomial for the physical problem described in Section (4.2) is a long and tedious task. Using the package MACSYMA(1988) to derive a characteristic polynomial for a general 8 x 8 matrix requires approximately 2 megabytes in disk space, McCarren and Ruskin(1991). Since the ultimate aim is to deal with matrices of order 24x24, the implications for consumption of CPU time are immense and not practical for on-line analysis. This is also the case for the eigenvalue equations which are even more analytically complex than the characteristic polynomial equation, and hence use a greater amount of CPU time.

As an illustration of these shortcomings, we consider the trigonal defect with a non-degenerate orbital ground state and two orbital doublet states described in section (4.5.1). We noted that this has an 8x8 secular matrix, and in the <100> stress direction only 4 energy emissions were observed, do Carmo et al.(1988). If we look at the characteristic polynomial of this equation then 8 λ terms would appear in this equation. However when we perform regression analysis, we must assign 4 experimental energy values to 8 λ variables. This leads to a problem of finding the correct allocation of experimental values to the λ (dependent) variables. Currently the only way of finding the optimum allocation is by testing each possible allocation in separate regressions of the characteristic polynomial, which is clearly unacceptable in terms of the time involved for an on-line system.

In addition the association of the matrix parameters to the parameters obtained using the characteristic polynomial and eigenvalue equations, causes considerable difficulty, since the development of both sets of equations for a matrix gives rise to the problem of unique identification of the parameters in equations (4.24) and (4.25). Consequently, having optimised the fits, we may expect that recovering the results to the physical solution is not a trivial task.

4.6.2 Analysis by the Powell-Shell Method.

An illustration of the performance of the alternative Powell-Shell method can also be given using the example cited in the previous section , do Carmo et al.(1988). The advantages of using this problem is that it involves an experiment with data that suggested a non-linear fit was required, and included imaginary components in the secular matrix due to spin orbital interaction. We can deal with such matrices relatively simply since they will at worst be Hermitian, Press et al.(1985). In order to deal with these matrices using eigenvalue routines normally based on real symmetric matrices we can apply a transformation of the Hermitian matrix into real and complex parts, Press et al(1990).

$$\mathbf{C} = \mathbf{A} + i\mathbf{B} \tag{4.31}$$

where **C** is an n x n matrix corresponding to the Hermitian matrix,

A is an n x n matrix corresponding to the real elements of **C** and

B is an n x n matrix corresponding to the complex elements of **C**.

If C is a Hermitian matrix , then

$$(\mathbf{A} + i\mathbf{B}).(\mathbf{w} + i\mathbf{z}) = \lambda(\mathbf{w}+i\mathbf{z}) \tag{4.32}$$

is equivalent to the 2n x 2n real problem

$$\begin{bmatrix} \boldsymbol{A} & -\boldsymbol{B} \\ \boldsymbol{B} & \boldsymbol{A} \end{bmatrix} \cdot \begin{bmatrix} w \\ z \end{bmatrix} = \lambda \begin{bmatrix} w \\ z \end{bmatrix} \tag{4.33}$$

where the 2n x2n matrix in equation (4.31) is symmetric since $\mathbf{A}^T = \mathbf{A}$ and $\mathbf{B}^T = -\mathbf{B}$

Corresponding to a given eigenvalue λ, we have the eigenvector

$$\begin{bmatrix} -w \\ z \end{bmatrix} \tag{4.34}$$

We can find the optimum fit using the Powell-Shell method on the symmetric real matrix **C**, in two distinct stages. In the first stage, Table (4.3) describes the process of finding initial estimates to be used on the Powell-shell method.

Each column header represents the current stage of estimation. Within each cell the corresponding value of the parameters used to find $S(\theta)$ are given, where $S(\theta)$ corresponds to a functional evaluation of sum of the squares function, and is not a full implementation of the Powell-Shell method.

In stage two the final estimates generated in stage one are used as initial estimates for the Powell-shell method with a limited number of experimental points. The estimates generated from this are then used as initial estimates for the complete data set.

Stage One:- Initial Estimates

(1) Initially we start with speculative crude estimates as shown in column (1) of Table (4.3).

(2) Tuning these estimates is achieved by using 10 points for each stress direction until the estimated points take on the shape of the experimental data, keeping the E and L parameters as constants, do Carmo et al.(1988), since these signify the energy shift between interactions and spin orbital effect respectively, and can be reasonably measured from the data set.

(3) The progress of the tuning technique can be seen through Table (4.3). We then use the estimates given in Column 12 as initial estimates for the Powell-shell method for 10 points in each direction in stage 2.

Table 4.3

Col No. / Parameter	1	2	3	4	5	6	7	8	9	10	11	12
A_1	1	0.5	-0.5	-0.5	-0.5	-0.05	-0.05	-0.005	0.005	-.0005	-0.005	-0.005
A_2	1	1	1	0.1	0.01	-0.01	0.01	0.001	0.001	-0.001	-0.001	-0.01
B_1	1	1	1	0.1	0.01	0.01	-0.01	-0.01	-0.01	-0.01	-0.01	0.01
B_3	1	1	1	0.1	0.01	0.01	0.01	0.001	0.01	0.01	0.01	-0.01
E	2.7	2.7	2.7	2.7	2.7	2.7	2.7	2.7	2.7	2.7	2.7	2.7
L	0.45	0.45	0.45	0.45	0.45	0.45	0.45	0.45	0.45	0.45	0.45	0.45
s(θ)	81212.9	74511.5	90962.9	534.546	26.45	27.0802	27.1	2.297	3.949	3.459	3.476	1.256

Stage two :- Estimation

(1) Table (4.4) describes the progress of the procedure from the choice of the column 13 in Table (4.3) as estimates, shown again in column 1 of Table (4.3). In each column header the number of experimental points that are used in the estimation procedure is described.

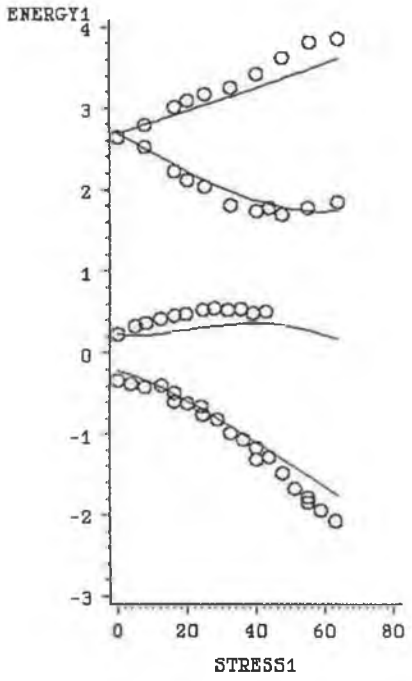
(2) Column 2 describes the estimates derived using column 1 as the initial estimates.

(3) Column 3 describes the sum of squares value, $S(\theta)$, when all the experimental points were included.

(4) Noting the values in column 3, the values described in column 4 are used as the initial estimates for the Powell-Shell method, with all the experimental points included.

(5) Column 5 describes the estimates derived by the Powell-Shell method with the complete data set. The final fits can be seen in Figures (4.7) to (4.9), noting that $S(\theta)$ is approximately 24 using the parameters described in do Carmo et al.(1988) as opposed to 9.5183 using the Powell-shell method.

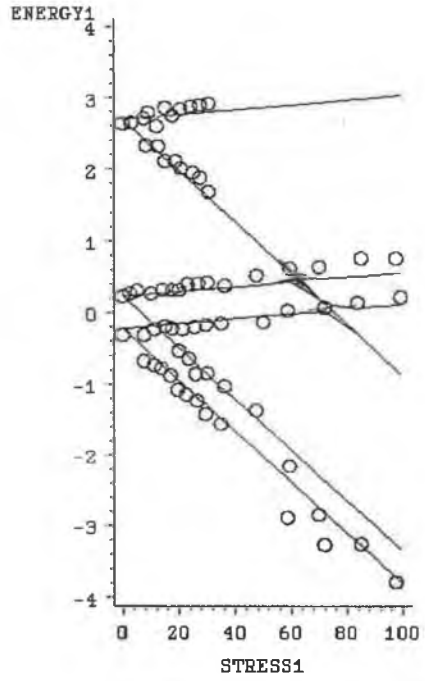
<100> Direction



o denotes actual data points
- denotes estimated lines

Figure (4.7)

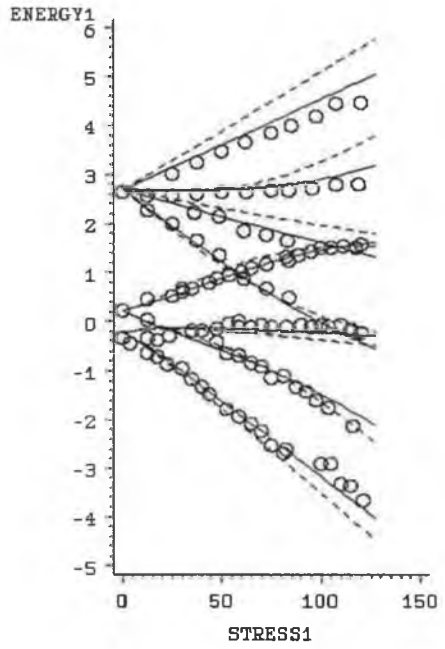
<111> Direction



o denotes actual data points
- denotes estimated lines

Figure (4.8)

<110> Direction



o denotes actual data points
- denotes estimated lines
denotes do Carmos lines

Figure (4.9)

Table 4.4 shows the path used to find final parameter estimates using the initial estimates derived .

		1	2	3	4	5
Col No.	001	10	10	33	33	33
Parameter	111	10	10	48	48	48
	110	10	10	55	55	55
A_1		-0.005	-0.00696	-0.006963	-0.006963	-0.006271
A_2		-0.01	-0.0126	-0.012608	-0.012608	-0.014872
B_1		0.01	0.0113	0.0113009	0.0113009	0.009456
B_3		-0.01	0.00003	0.000036	0.000036	-0.009456
E		2.7	2.66953	2.66953	2.7	2.70
L		0.45	0.6088	0.6088	0.45	0.45
s(θ)		1.256	0.527665	32.0266	30.8732	9.5183

Table 4.5.

Parameters	Do Carmo et al(1988)	Powell-Shell Method
A_1	-5.4 meV/Gpa	-6.27 mev/Gpa
A_2	-15.7 meV/GPa	-14.872 meV/GPa
B_1	-12.4 meV/GPa	-9.456 meV/GPa
B_3	8.7 meV/GPa	9.456 meV/GPa
E	2.7 meV	2.7 meV
L	0.45 meV	0.45 meV
s(θ)	24.2767 meV^2	9.5183 meV^2

The difference between the two sets of parameter estimates is not substantial, as can be seen from Table (4.4), but a comparison of the accuracy of the fit to the experimental data shown in Figure (4.9), is interesting. Figures (4.7) and (4.8) show the results for <100> and <111> stresses; for these cases, the fits cannot be distinguished. For <110> stresses, shown in Figure (4.9), the fit produced is clearly improved giving more precise parameter estimates and

identifying the transition type satisfactorily in a reasonable time period. Convergence times and related considerations are discussed in Chapter 5.

4.6.3 Further Examples of the Method - New Data

Other examples using Zeeman data, McGuigan(1989), and Beryllium impurities in Silicon, Campion et al.(1992), were tested. The Zeeman data was a fairly simple problem in terms of fitting since excellent initial estimates were available, McGuigan (1989), and the overall results obtained by the Powell-Shell were in agreement with those obtained by the basic method. The Beryllium impurities (previously unidentified) were successfully identified by our approach with a tetragonal defect with a E->A transition, McCarren et al.(1994). The results of both these examples are shown in Appendix A.

In addition to the above, our method has also recently been tested on Indium implanted Silicon data, Kehoe(unpublished), which is suspected of having a trigonal A->A symmetry with an additional interacting trigonal E->A state. An interesting feature of this data is that only *one* transition is observable from the experimental data. The reason that interaction of transitions is suspected is that the data in the <100> and <110> directions, become non-linear under the application of stress, Figures (4.10)-(4.12). This effect usually implies that mixing of states has occurred as noted previously, and hence indicates the presence of *at least* one additional transition. The fits, the experimental data and details of the transitions are shown in Figures (4.10)-(4.12). The secular matrix employed for this fit is shown in Table (4.6),McGuigan (private communication), and the parameters we obtained are given in Table (4.7).

A similar example described by Daly (unpublished) which includes missing transitions is also described in Appendix A.

Table (4.6) secular matrix for a Trigonal A->A interacting with a Trigonal E->A symmetry.

H_s	A	E_x	E_y
A	α	θ	ϕ
E_x	θ	$E+\alpha'-\beta'$	γ'
E_y	ϕ	γ'	$E+\alpha'+\beta'$

where

$$\alpha=A_1S_{ii}+2A_2S_{ij},$$

$$\alpha'=A_1'S_{ii}+2A_2'S_{ij},$$

$$\beta'=B'S_{\theta}+C'S_{\theta'},$$

$$\gamma'=B'\sqrt{3}S_{\varepsilon}+C'\sqrt{3}S_{\varepsilon'},$$

$$\phi=D_1\sqrt{2}S_{\varepsilon}+D_2\sqrt{3}S_{\varepsilon'},$$

$$\theta=D_1S_{\theta}+D_2S_{\theta'},$$

and

$$S_{ii}=S_{xx}+S_{yy}+S_{zz},$$

$$S_{ij}=S_{xy}+S_{yz}+S_{xz},$$

$$S_{\theta}=S_{xx}+S_{yy}-2S_{zz},$$

$$S_{\theta'}=S_{yz}+S_{zx}-2S_{xy},$$

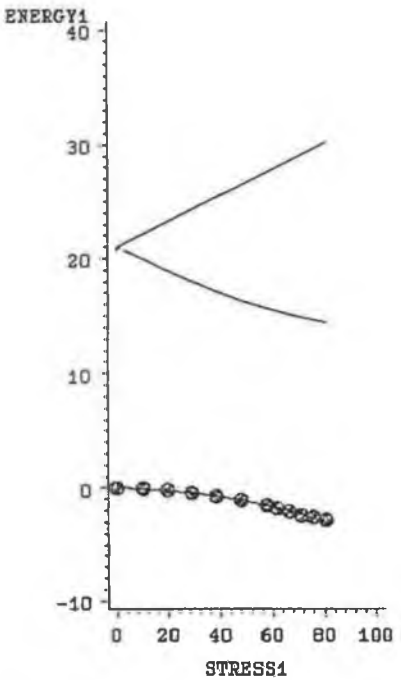
$$S_{\varepsilon}=S_{xx}-S_{yy} \text{ and}$$

$$S_{\varepsilon'}=S_{yz}-S_{zx}.$$

Table 4.7 Results for the Indium implanted Data with C and D_2 held at Zero.

Parameters	A_1	A_2	A_1'	A_2'	B	D_1	E
Results	-0.0053	0.0041	-0.006	-0.0412	-0.0169	0.0428	19.1

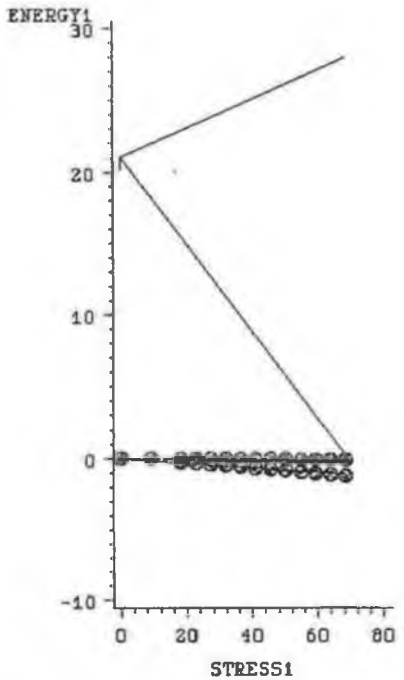
<100> Direction



Circle denote actual data points
Joint Lines denote estimated points

Figure (4.10)

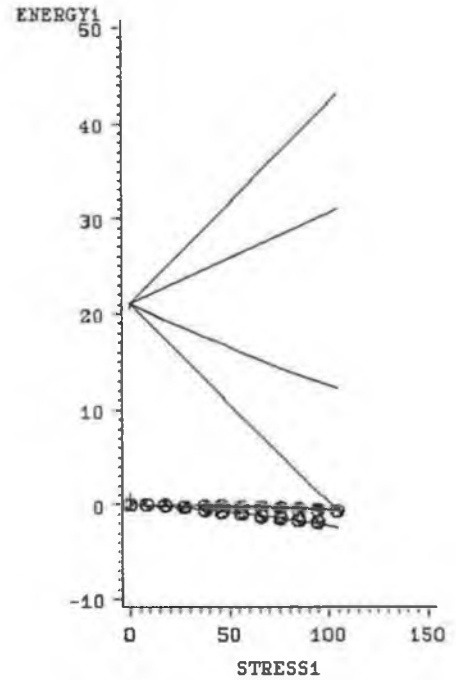
<111> Direction



Circle denote actual data points
Joint Lines denote estimated points

Figure (4.11)

<110> Direction



Circle denote actual data points
Joint Lines denote estimated points

Figure (4.12)

4.7 Summary

From the analysis in Section (4.6) the Powell-Shell method clearly performs well in terms of fitting this type of experimental data and enabling identification of transition-type for piezo-spectroscopic measurements. The data manipulation described in Section (4.6) and also in McCarren et al.(1994), while requiring some considerable initial effort, leads to good precision of estimation and systematic refinement where initial estimates are poor. The software tool we have developed (Chapter 6 and Appendix G) is thus flexible and *reasonably* comprehensive.

The Powell-shell method also effectively allows us to account for non-linearity in the form of multi-response functions, where this involves questions of bias. Where interaction effects are zero, the implications for good final parameter estimates may be negligible, but for more complex examples this is not the case. While transition identification is the ultimate goal, it is clearly important to determine the effect of increased complexity on the accuracy of the identification, on the errors involved in the fitting process, and on the speed of convergence. Clearly, these considerations are particularly important where data is missing or of poor quality.

Chapter 5

Error Analysis

5.1 Introduction

The strength of the Powell-Shell method is based upon its ability to model the experimental situation while keeping the amount of bias to a minimum. Even for relatively simple models, failures in the assumptions can result in considerable bias and lack of precision and where nonlinearity is a feature these problems can be considerably exacerbated. In particular, an iterative approach (which incorporates its own approximations and hence uncertainties) is usually necessary. Diagnostics are thus a key requirement in assessing model limitations and interpreting the analysis. In what follows, we therefore interpret the diagnostic situation of two recognised models, Campion et al.(1992), do Carmo et al.(1988), and discuss the implications for their physical interpretations. We also investigate a further problem communicated by Kehoe(unpublished), for which nonlinear effects are marked.

5.2 Multi-Response Estimation

The methodology of the Powell-Shell is equivalent to a multi-response problem measured on N experimental runs, equations (4.7)-(4.14), where the models for the m responses depend on a total of p parameters, θ and

$$Y_{nm} = f_m(Stress_n,\theta_p)+Z_{nm} \qquad \textbf{(5.1)}$$

for n=1,...,N; m=1,...,M and

where Y_{nm} is the random variable associated with the measured value of the mth energy value for the nth stress, f_m is the model function for the mth energy value (response) depending on some or all of the stress settings and on some or all of the parameters θ, and Z_{nm} is the disturbance term. In the analysis of piezo-spectroscopic data we assume that the N experimental points for stress, m=1,....,M are fixed and known, so we can form the NxM observation matrix Y with the (n,m)th element Y_{nm} and the NxM expected response matrix $H(\theta)$ with the (n,m)th element $f_m(Stress_n,\theta)$. From Y and $H(\theta)$ we create the residual matrix

$$Z(\theta) = Y\text{-}H(\theta) \qquad \textbf{(5.2)}$$

The parameter estimates θ' are given by the values of θ which optimise some criterion based on $Z(\theta)$ in the same way that the least squares estimates in uni-response parameter estimation minimises $||\underline{Z}(\theta)||^2$, the length of the error vector squared. The criterion will depend on assumptions about the disturbance. For example, if we make the stringent assumption that the Z_{nm} are normally distributed and independent with the same variance σ^2, then least squares is appropriate and we find θ' which minimises the sum of squared residuals of all NM responses. That is, the estimation criterion would be to minimize the trace of Z^tZ, $tr(Z^tZ)$, Bates and Watts (1988).

The assumptions leading to the trace criterion may not be realistic. It could be reasonable to assume that the variances of different measurements on the same response are constant, but not that variances of different responses are equal. Furthermore, the assumption of independent disturbances for different measurements in the same experimental run may not be justified. For example, in chemical experiments where the concentrations of a number of different chemical species are measured from the same sample then, for the situation where only the relative concentrations can be determined, different measurements on the same sample may be correlated, Bates and Watts (1988).

Box and Draper (1965) give the usual disturbance term for a normal distribution as

$$E[Z_{nm}]=0 \tag{5.3}$$

and

$$E[Z_{mn}Z_{ri}]=\begin{Bmatrix}\Sigma_{mi} & n=r\\ 0 & n<>r\end{Bmatrix} \tag{5.4}$$

where Σ_{mi} is a fixed MxM covariance matrix. Generally for multi-response problems, we assume that measurements from different experiments are independent but measurements from the same experiment are correlated.

5.3 Estimators : some alternatives

The underlying assumptions described in the previous section imply that the least-squares estimator is not the maximum-likelihood estimator. Bates and Watts (1988) describe the maximum-likelihood and Bayesian estimator as that which minimises the $|Z^TZ|$. However this estimator proves to be inappropriate when *missing values* are included. Bates and Watts(1988) recommend that when the number of missing data points is at a minimum we should proceed in two stages where the first is based on fixing the missing residuals at zero, and estimating the parameters, the second on reversing the procedure once the missing residuals have been estimated. Where the number of values missing is large, the parameters can be obtained by minimisation of an alternative estimator due to Stewart and Sorensen (1981) and given by,

$$S(\theta,\Sigma)=(M+1)\ln|\Sigma|+\Sigma\ln|\Sigma_n|+\Sigma\Sigma\Sigma\delta_{min}\{Y_{nm}-f_{nm}(\theta)\}\{Y_{ni}-f_{ni}(\theta)\} \tag{5.5}$$

where Σ_n is obtained from Σ by substituting

$$\delta_{mi}=\begin{Bmatrix}1 & m=i\\ 0 & m<>i\end{Bmatrix} \tag{5.6}$$

whenever the data for case (n,m) or (n,i) is missing, and the term δ_{mi} is the

(m,i)th entry of Σ^{-1}_n. This approach requires us to obtain estimates of the covariances between the responses. This is a difficult problem because of the number of parameters to be estimated, Bates and Wates(1988). An alternative approach, Seber and Wild(1988), which is based on estimating θ and Σ using a two-stage procedure is as follows

1. Minimise $||y(j)-fj(\mathbf{stress},\theta)||^2$ with respect to θ to obtain $\theta'(j)$, for each j=1,2,....,d. where j corresponds to the response.

2. Calculate e_j and obtain an estimate of $\Sigma=[\sigma_{rs}]$ using $\sigma_{rs}=e'_r e_s/n$ (r,s=1,2...d)

3. Minimise

$$T[\theta]=[y-f(\theta)]'(\Sigma^{-1} \times I_n)[y-f(\theta)]$$

w.r.t. to θ to get θ'.

This method is primarily used for complete data sets and does not perform well when data are missing. An additional drawback is the number of calculations required.

5.3.1 Derivation of Stewart and Sorensen Estimator for the Piezo-Spectroscopic Case.

The estimator described by Stewart and Sorensen(1981) is specific to multi-response cases were the m responses are correlated within each experimental run. The unbiased multivariate Normal probability density, Wilks(1962), can be given as follows :

$$p(Z_n \backslash \sum) = \prod_{n=1}^{N} 2\pi^{-m_n} |\sum_n|^{-1/2} \exp(-\frac{1}{2} Z_n \sum_n^{-1} Z_n^T) \qquad (5.7)$$

with Σ a positive definite covariance matrix. Here Z_n is as described above with dummy zeroes inserted wherever observations are missing; Σ_n is obtained from Σ by substituting elements of a unit matrix when an observation Y_{ni} or Y_{nj} is missing. Equation (5.7) may be derived for the full m-variate density by

maximising the density with respect to the missing error values, Stewart and Sorensen(1981). The joint observation density function for n experiments in any particular stress direction corresponds to substituting equation (5.1) into (5.7) to give the following :

$$p(Y_l \mid \theta,\Sigma_l)=\prod_{n=1}^{N} 2\pi^{-m_n} \mid \Sigma_n \mid^{-1/2}.\exp\left\{-\frac{1}{2}\sum_{n=1}^{N}\sum_{i=1}^{M}\sum_{j=1}^{M}\sigma^{ij}_{nl}[Y_{ni}-f_{ni}(\theta)].[Y_{nj}-f_{nj}(\theta)]\right\} \quad \textbf{(5.8)}$$

where l is the stress direction and the elements Σ^{ij}_{nl} are obtained from the matrix Σ^{-1}_{nl}, $f_{ni}(\theta)$ represents $f_i(x_n,\theta)$ and is called the likelihood function $l(\theta,\Sigma\backslash Y)$. Factoring $p(\theta,\Sigma)$, Stewart and Sorensen(1981), show that

$$p(\theta,\Sigma)=p(\theta)p(\Sigma) \quad \textbf{(5.9)}$$

with a uniform density $p(\theta)$ in the permitted region of θ. This requires some care in the reparameterisation of the model. A noninformative prior density $p(\Sigma)$ is used, Seber and Wild(1989), in the "permitted" region of Σ :

$$p(\Sigma) \propto \mid \Sigma \mid^{-(m+1)/2} \quad \textbf{(5.10)}$$

which on application of Bayes theorem then gives the following posterior density in the permitted region (θ,Σ),

$$p(\theta,\Sigma\backslash Y)=p(\theta,\Sigma)l(\theta,\Sigma\backslash Y) \quad \textbf{(5.11)}$$

Estimates of θ and Σ can be obtained by maximising the posterior density or by minimising the log-likelihood function, equation (5.4).

When all three stress directions are included together we get the joint density function for nl experiments in l=3 directions as follows :

$$P(Y_1,Y_2,Y_3\backslash\theta,\sigma)=P(Y_1\backslash\theta,\sigma).P(Y_2\backslash\theta,\sigma).P(Y_3\backslash\theta,\sigma) \quad \textbf{(5.12)}$$

where independence is explained by the fact that estimated eigenvalues are generated independently for each stress direction. Thus when using the log-likelihood function this relates to minimising the following

$$\sum_{l=1}^{3} S_l(\theta,\Sigma_l) \qquad \textbf{(5.13)}$$

where $S_l(\theta,\Sigma_l)$ is equivalent to that described in equation (5.5).

If we were to try and minimise equation (5.11) for the problem described by do Carmo et al (1989) then the number of parameters to be estimated would be so large as to be prohibitive. For example in the <110> and <111> directions of this problem the covariance matrices are 16x16, which implies that 1024 parameter estimates would be required for each direction, hence we would have more parameters to be estimated than experimental data points. This minimisation problem can be simplified somewhat if we assume that there is no real correlation between the error vectors **Z** of the differing responses, i.e.

$$\sum_{l=1}^{3} S(\theta,\Sigma)=\sum_{l=1}^{3}\left\{(M+1)\sum_{i=1}^{N_l}(n_{il}+1)\ln(\sigma_{il})+\sum_{j=1}^{N_l}\sum_{i=1}^{M_l}\sigma_{il}^{-1}[Y_{ijl}-f_{ijl}(\theta)]^2\right\} \qquad \textbf{(5.14)}$$

where n_{il} is the number of points for response i and direction l.

If the assumptions of constant variance within any particular response and constant variance between responses hold, then the above reduces to the least-squares estimator, Stewart and Sorensen(1981).

In order to examine these assumptions we consider some of the examples discussed in the previous chapter with parameter estimates obtained using our Powell-shell method.

5.4 Covariances, Correlation and Auto-correlations

5.4.1 Correlations Between Responses In Multi-Response Data.

The merit of the least squares estimator of the Powell-Shell is clearly its relative simplicity when compared to that described by Stewart and Sorenson(1981).

However, we obviously need to investigate the validity of the underlying assumptions, e.g. the properties of the errors within and between the individual responses. The underlying assumptions of least-squares, (section 5.2)), relate to the Z_{nm} normally distributed and independent with the same constant variance, σ^2, where

$$Z_{nm} = Y_{nm} - f_m(Stress_n,\theta_p) \qquad \textbf{(5.15)}$$

We look first at the correlation matrix of the errors. Under the least-squares assumption, the errors corresponding to any particular response should not be correlated with the errors of any other response. To illustrate this we look at the physical problem discussed by Campion et al.(1992), and analyse the correlation matrix of the errors, shown in Table (5.1).

Table (5.1) : Correlations between the errors of pairs of responses for the Campion et al.(1992) example.

Correlation for Direction <111>

1.0000	-0.8614
-0.8614	1.0000

Correlation for Direction <100>

1.0000	1.0000	-0.0596	0.3660
1.0000	1.0000	-0.0596	0.3660
-0.0596	-0.0596	1.0000	-0.0065
0.3660	0.3660	-0.0065	1.0000

Correlation for Direction <110>

1.0000	0.9265	0.7998	-0.4932
0.9265	1.0000	0.5002	-0.4238
0.7998	0.5002	1.0000	-0.3380
-0.4932	-0.4238	-0.3380	1.0000

Table (5.1) shows the correlations between the errors of the individual response with all other responses in the respective stress direction. On inspection of this table, we can see that no major correlation exists between the errors for direction <111>. In direction <100> we see that the errors between the first two responses and the first actual experimental line are completely correlated. This occurs because of degeneracy, since experimentally we only see three responses for this direction but in reality there are four including two equivalent responses, Appendix (A-5). For direction <110>, Appendix (A-6), we see a similar

situation to that described in the <100> direction. Once again the fits make it clear that this is due to degeneracy and again the error assumptions hold. Thus, for the case discussed by Campion et al.(1992) the assumption of no correlation of the errors between responses seems to be upheld. For the further example given in Chapter 4, Kehoe(unpublished), we have noted the occurrence of undetected or non-measurable responses. The correlation matrices for the errors in each direction are given in Table (5.2).

Table (5.2) : Correlations of the errors between pairs of responses for the Indium implanted silicon data, Kehoe(unpublished).

Correlation for Direction <100>
1.0000
Correlation for Direction <111>
1.0000 -0.2909
-0.2909 1.0000
Correlation for Direction <110>
1.0000 -0.6076
-0.6076 1.0000

The matrices for all three directions of this example do not show any serious correlation of the errors between responses. Similar results were found for the example described by do Carmo et al.(1988), Appendix B. In this last example however there are many missing values which has the effect of reducing the number of correlations that can be calculated. This is due to the fact that only points with corresponding stress values can be compared. Inevitably, some spuriously high correlations were obtained, but in all cases these related to cases where the number of points was very small[1].

5.4.2 Auto-correlation Detection

In order to investigate the variance of the errors for each response it is usual to

[1] *A correlation value was calculated between two error vectors if the number of matching points was greater than 4.*

plot the standardised error against the independent variable. For the example described by Campion et al.(1992) the error versus stress graphs for the two lines in the <111> direction are shown in figures (5.1-2), in figures (5.3-5) for the three lines in the <001> direction, and in figures (5.6-9) for the four lines the <110> direction. For this particular example the number of data points used for both lines is small, but patterns appear to be emerging within the error plots for both figures (5.1) and (5.2). This would suggest that auto-correlation and/or curvilinearity may exist in the error terms. This may occur, in part, due to the fact that the width of the spectrum peaks becomes larger as stress increases, hence the margin for error when picking a peak intensity is increased. Similar diagrams are described for the do Carmo et al.(1988), Kehoe(unpublished) and Daly(unpublished) examples in Appendix C.

Autocorrelation or serial correlation within a model can be described as the serial correlation of the error terms and this implies invalidation of the independence assumption. If we consider the nonlinear model with additive disturbances

$$Y_n=f(Stress_n;\underline{\beta})+ e_n \qquad \textbf{(5.16)}$$

which corresponds to a single response of a multi-response piezo-spectroscopic model described in section 5.2, then if autocorrelation exists, the simplest form of the error distribution can be described by a first order autoregressive model, equation (5.17).

$$e_n = \rho e_{n-1}+\mu_n, \qquad -1<\rho<1 \qquad \textbf{(5.17)}$$

where ρ is the correlation between the error terms and μ_n is IID with $N(0,\sigma^2)$. The presence of autocorrelation within the error terms of a linear model implies that the ordinary least squares regression coefficients are still unbiased, but no longer have the minimum variance property and may be quite inefficient, and that the Mean Square Error (MSE) may seriously underestimate the variance of

The Campion et al.(1992) example in the <111> direction.

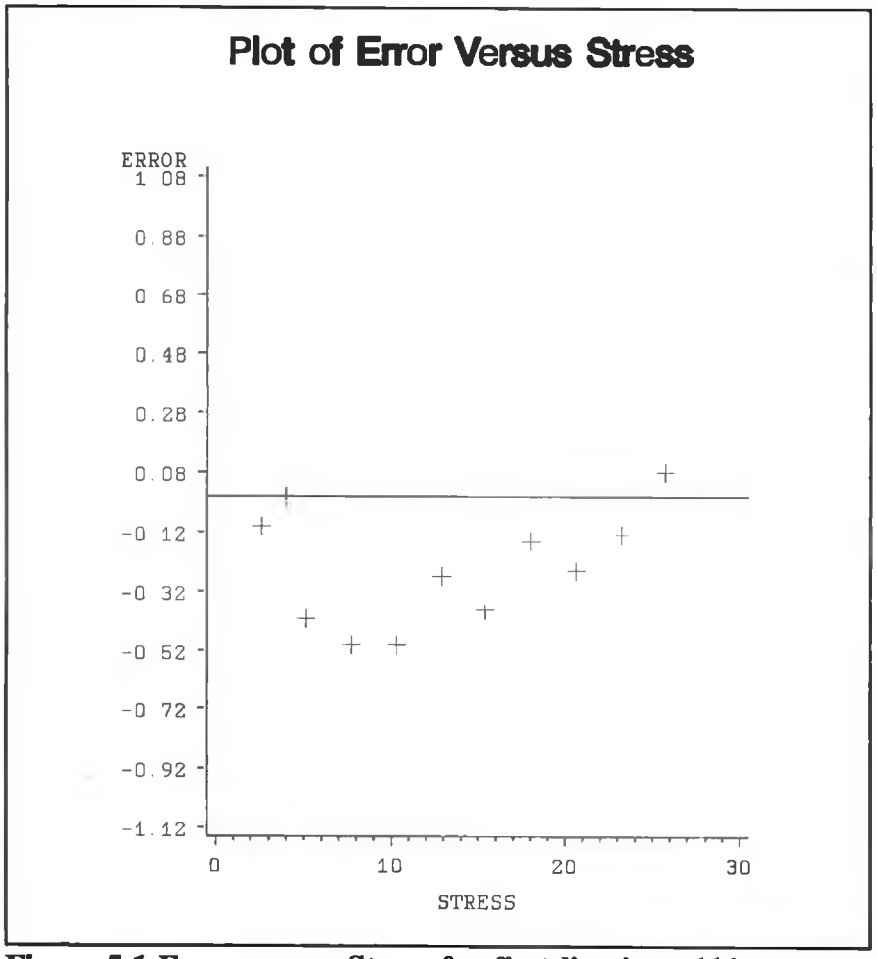


Figure 5.1 Error versus Stress for first line in <111>.

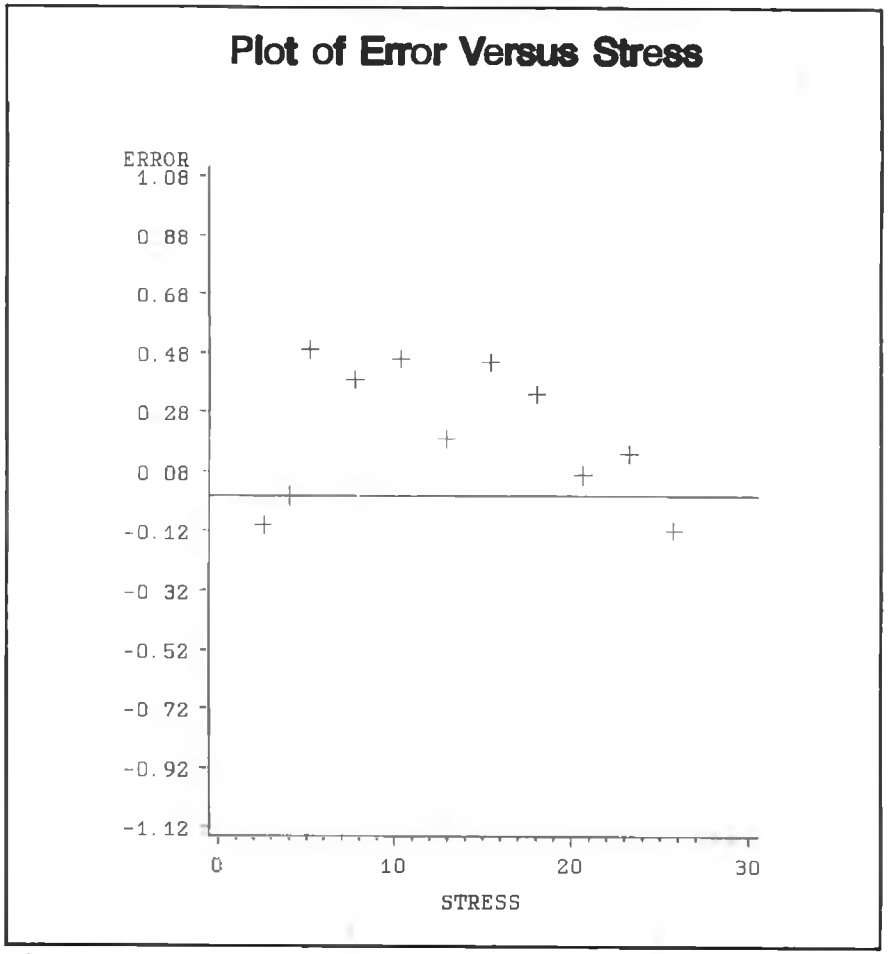


Figure 5.2 Error versus Stress for second line in <111>.

The Campion et al.(1992) example in the <001> direction.

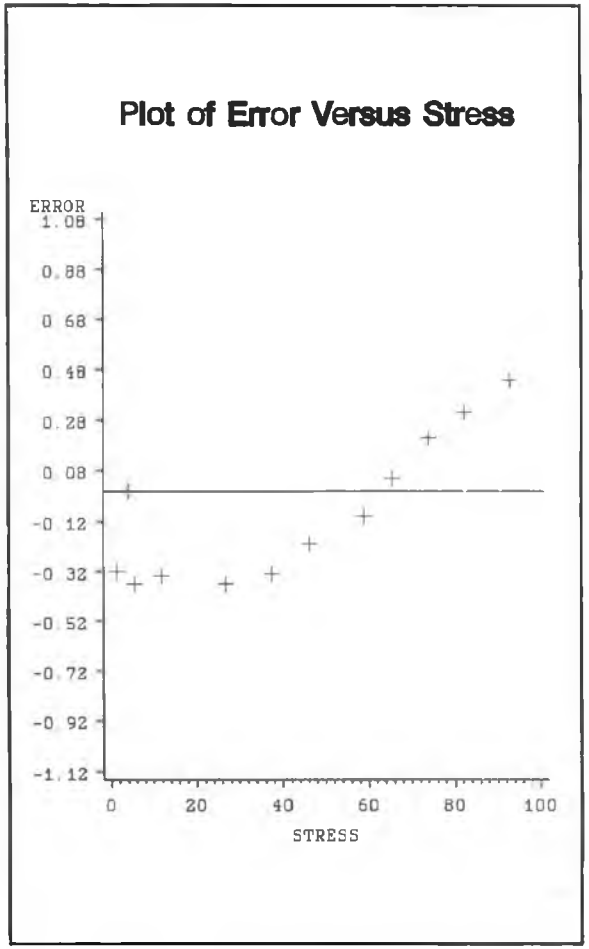


Figure 5.3 Error versus Stress for First line in <001>

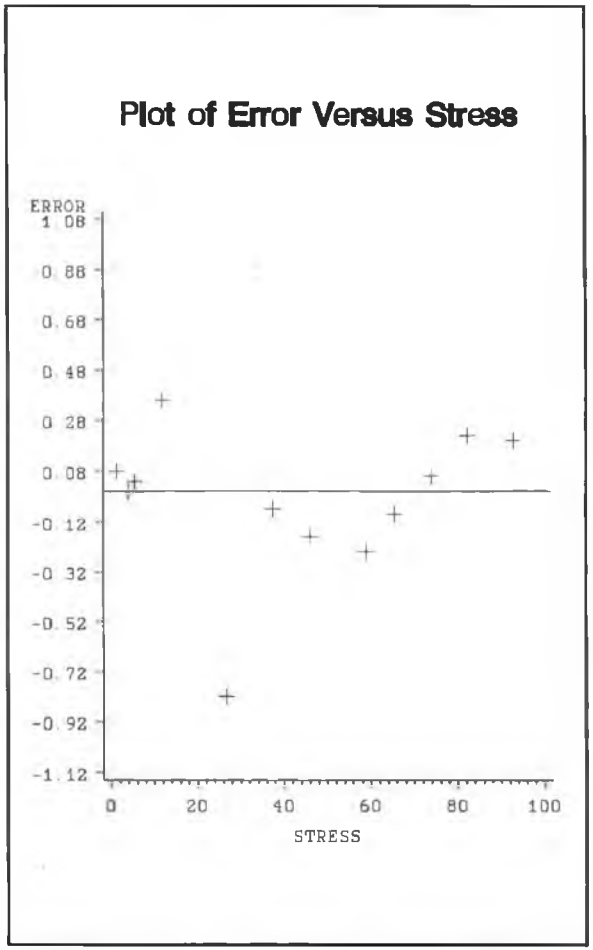


Figure 5.5 Error vs Stress for the third line in <001>

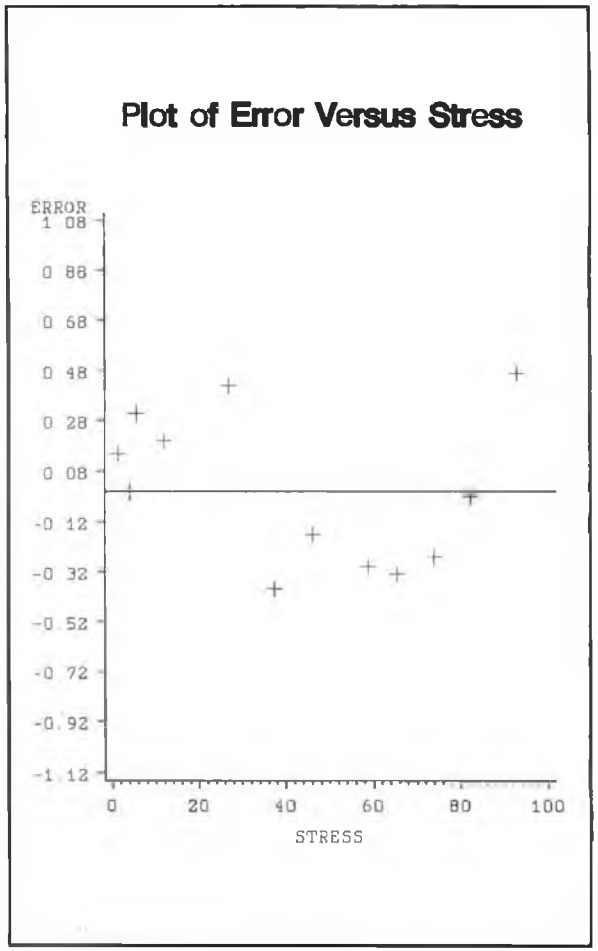


Figure 5.4 Error versus Stress for second line in <001>

The Campion et al.(1990) example in the <111> direction.

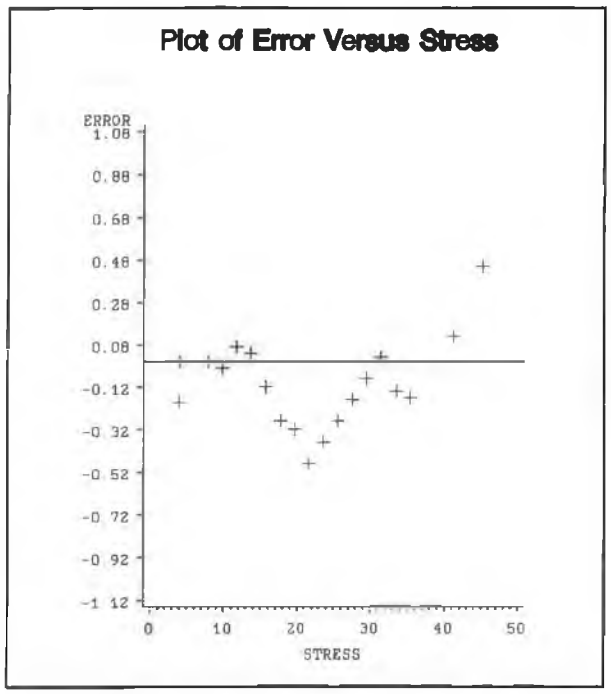


Figure 5.6 Error versus Stress for first line in <110>

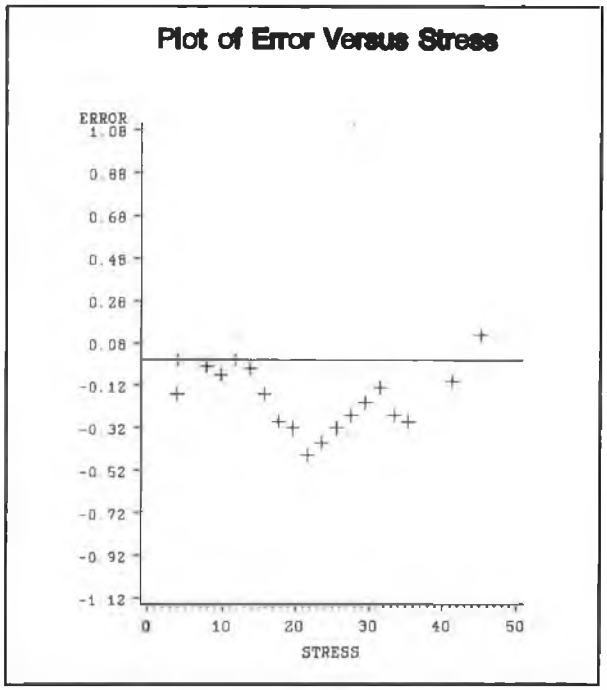


Figure 5.7 Error versus Stress for second line in <110>

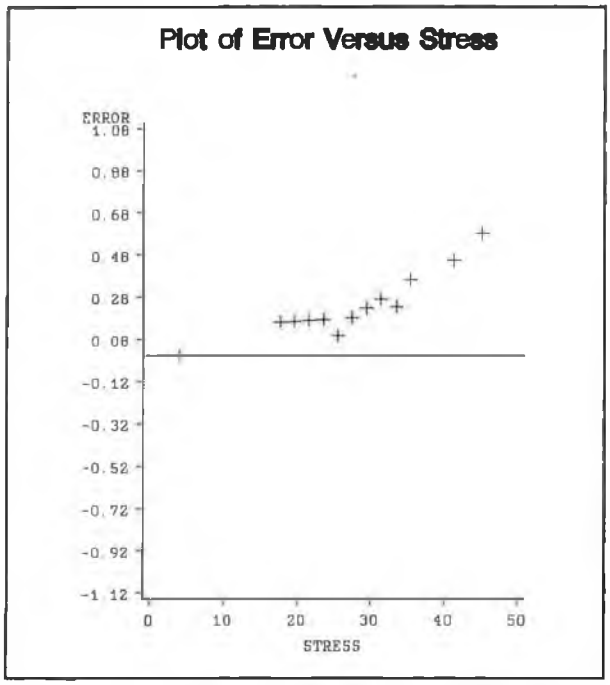


Figure 5.8 Error versus Stress for Third Line in <110>

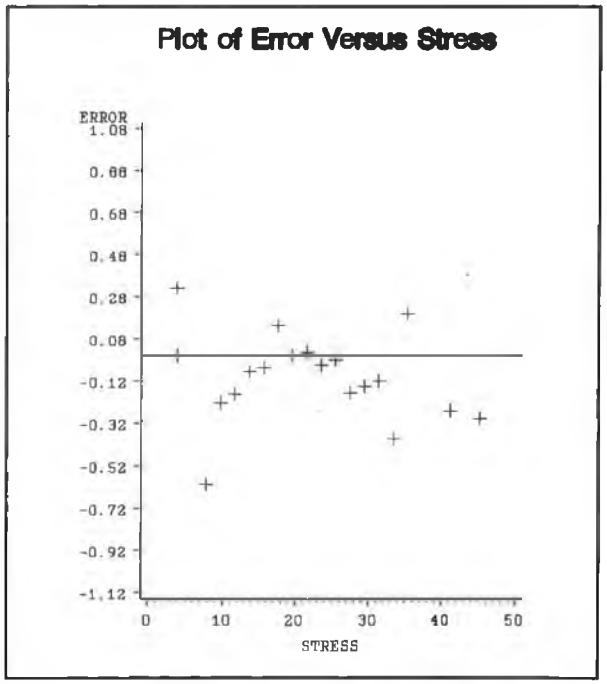


Figure 5.9 Error versus Stress for Fourth Line in <110>

the error terms, Neter et al.(1985) and Bates and Wates(1988). Also the standard deviation of the parameter estimates calculated according to ordinary least-squares procedure may seriously underestimate the true standard deviation, and the confidence intervals and tests using the t and F distributions, are not strictly applicable, Neter et al.(1985).

In order to determine if there is a significant amount of serial correlation between the error terms in a linear model we use the Durbin-Watson statistic, Neter et al(1985), of form

$$D = \frac{\sum_{t=2}^{N} (e_t - e_{t-1})^2}{\sum_{t=1}^{N} e_t^2} \qquad \textbf{(5.18)}$$

An exact procedure is not available, but Durbin and Watson have obtained lower and upper bounds d_L and d_U such that a value of D outside these bounds leads to a definite decision.

The above procedure is appropriate when the model is linear. For the case of a non-linear model, Kobayashi (1991) describes a procedure to determine if serial correlation exists for a uniresponse model. The basis of the procedure is similar to that of Durbin-Watson except that the ρ is calculated by maximising the log-likelihood conditional upon the initial disturbance e_0 under the assumption of a normal distribution. Neglecting the constant term, this may be written

$$L(\rho,\theta)=-(1/2)\Sigma[\{y_n-f(stress_n;\theta)\}-\rho\{y_{n-1}-f(Stress_{n-1};\theta)\}]^2/\sigma^2 \qquad \textbf{(5.19)}$$

The test statistic ρ is obtained by maximising with respect θ and ρ. The estimate of σ^2 is not included since it can be estimated from the errors generated as it is effectively related to ρ. Kobayashi(1991) shows that $H_o(\rho=0)$ is accepted if

$$-\infty < \rho < Z_{1-\alpha}-n^{-1/2}K \qquad \textbf{(5.20)}$$

rejected if

$$Z_{1-\alpha}+n^{-1/2}K< \rho <-\infty \quad \textbf{(5.21)}$$

and inconclusive if

$$Z_{1-\alpha}-n^{-1/2}K< \rho <Z_{1-\alpha}+n^{-1/2}K \quad \textbf{(5.22)}$$

where $Z_{1-\alpha}$ is the $(1-\alpha)$th quantile of the standardized normal distribution and K corresponds to the number of parameters.

Obtaining exact critical values for this test is complex for the piezo-spectroscopic case since it depends on obtaining the derivatives of the individual response functions which are discontinuous in our case due to missing data and ranked data points. However, an approximation of the statistic can be constructed based on the acceptance and rejection regions. The assumption here is that no correlation exists between responses, hence each response can be treated individually. Maximising the Liklihood function for the multi-response problem is equivalent to minimising the following

$$S(\rho,\theta)=(1/2)\Sigma\Sigma[\{y_{mn}-f_m(stress_n;\theta)\}-\rho\{y_{mn-1}-f(Stress_{n-1};\theta)\}]^2/\sigma^2 \quad \textbf{(5.23)}$$

This model is similar to the one described in section (5.2) with constant and equal variance assumed for each response. A problem with this method in practice is that the test often gives inconclusive results. In such situations we assume the worst case situation and model the data as if the particular $\rho > 0$.

5.4.3 Derivation of Auto-Correlation Estimators using a Two-Stage Procedure

The accuracy of parameter estimates of any function obtained using any minimisation algorithm decreases as the number of parameters increases, Walsh(1974). In relation to the minimisation techniques used in chapter 4 this is illustrated for example by the Powell which becomes ineffective for more than 15 parameters, but which could then be satisfactorily replaced by Fletcher-Reeve. In order to complete a minimisation of equation (5.23) we need to carry

out the procedure in two-stages, e.g. Huang-Huang(1991) use such a two-stage procedure to estimate linear and non-linear parameters separately. To implement a two-stage procedure on a multi-response model with autocorrelated errors, the primary task is to obtain initial estimates of the θ and ρ. This is done by using the estimates of θ, θ′, when no autocorrelation, is considered, and by estimating ρ, **r,** using the following

$$r_i = \frac{\sum_{j=2}^{N} (e_{ij} \cdot e_{ij-1})^2}{\sum_{j=1}^{N} e_{ij-1}^{\ 2}}, \qquad \textbf{(5.24)}$$

where m is the number of responses and i=1,..,m.

Using the $\mathbf{r}_i$ as initial estimates to the first stage and keeping θ constant we implement either the Powell or Fletcher-Reeves algorithms. On completion of this stage we then estimate $\underline{\theta}$, keeping **r** constant. This procedure is continued until

$$S(\underline{\theta}_j,\mathbf{r}_j)\text{-}S(\underline{\theta}_{j-1},\mathbf{r}_{j-1}) < \Delta \qquad \textbf{(5.25)}$$

where Δ is a specified tolerance and j corresponds to the current iteration.

For the example described by Campion et al.(1992) the number of estimates in the $\underline{\theta}$ vector is 4 while there are 10 estimates in the $\mathbf{r}_i$ vector. For the example discussed by do Carmo (1988) there are six θ and 40 r parameters[2]. For this example even the two stage procedure broke down, although when an r_i initial estimate was in the acceptance region discussed by Kobayashi(1991) we did not estimate for that parameter, in an endeavour to keep the number of parameters to a minimum.

[2] *There are 16 parameters in the <111> and <110> directions and 8 in the <001> direction.*

5.4.4 Incorporating Auto-Correlation Results

Having incorporated the autocorrelation terms by the method described in the previous section, the resulting error versus stress diagrams are shown for the <110> direction for illustration in Figures (5.10) -(5.13), with the remaining Figures described in Appendix D. The resulting θ and ρ estimates are shown in Table 5.3 where r_{ij} corresponds to the estimated auto-correlation parameter in direction i for response j.

Table (5.3) : Autocorrelation estimates for Campion et al.(1992).

	Parameter	Estimate
	a1	0.015219
	a2	-0.012612
	b	-0.002815
	cm	0.031976
<111>	r_{11}	0.91376
	r_{12}	0.63315
<100>	r_{21}	0.9675
	r_{22}	0.9678
	r_{23}	0.4421
	r_{24}	-0.0964
<110>	r_{31}	0.8965
	r_{32}	0.9627
	r_{33}	1.2253
	r_{34}	-0.1812

The above estimates give the fits shown in Figures (5.14)-(5.16) showing a considerable overall improvement compared to the results when no autocorrelation is included, (shown in Figures (5.17)-(5.19)). The results for the example discussed by do Carmo et al.(1988) with the autocorrelation parameters

Campion et al.(1992) with autocorrelation terms included.

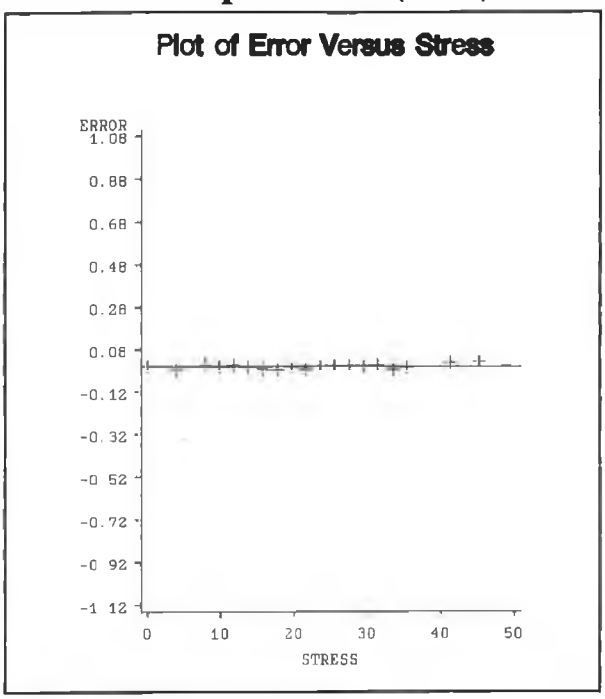


Figure 5.10 Error versus Stress for first line in <110>

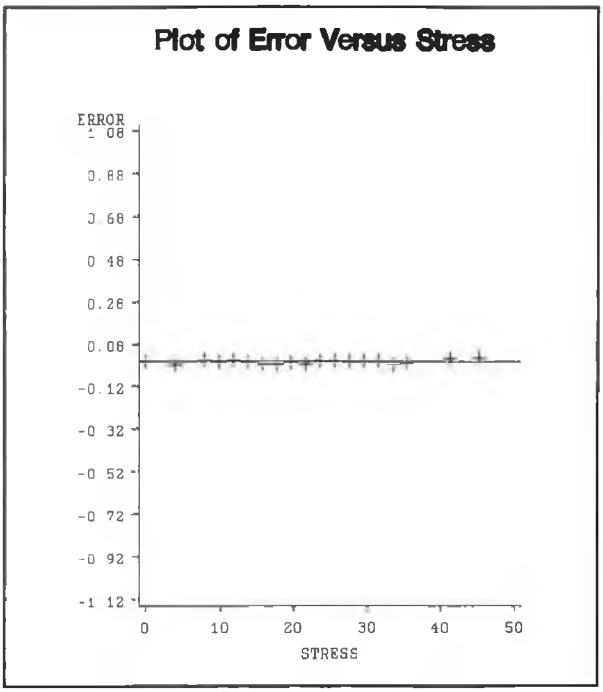


Figure 5.11 Error versus Stress for second line in <110>

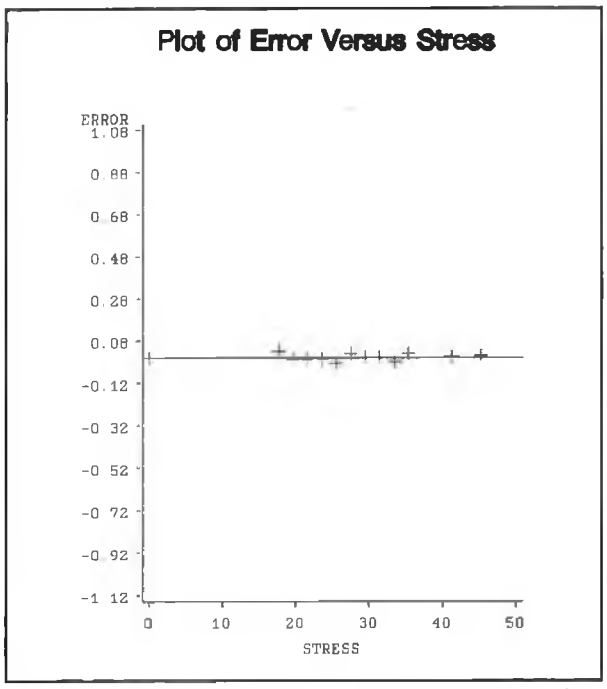


Figure 5.12 Error versus Stress for Third Line in <110>

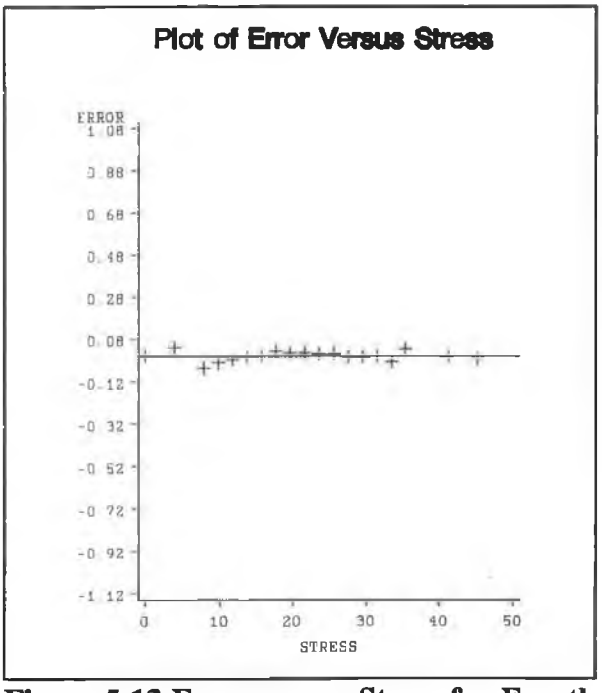
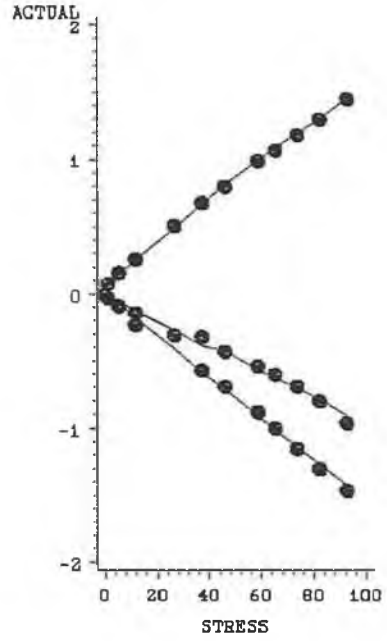


Figure 5.13 Error versus Stress for Fourth Line in <110>

Campion et al.(1992) fits with autocorrelation terms included.

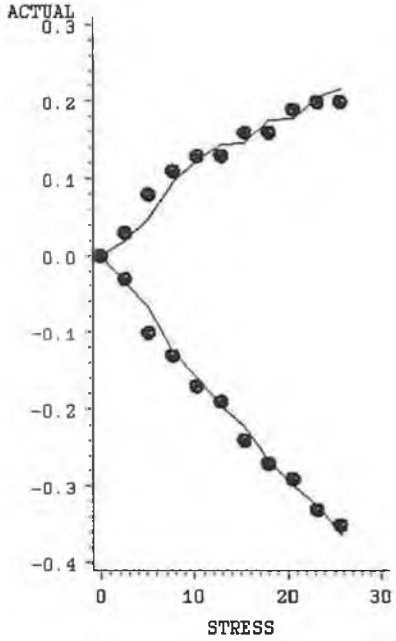
<001> Direction



Circle denote actual data points
Joint Lines denote estimated points

Figure 5.14

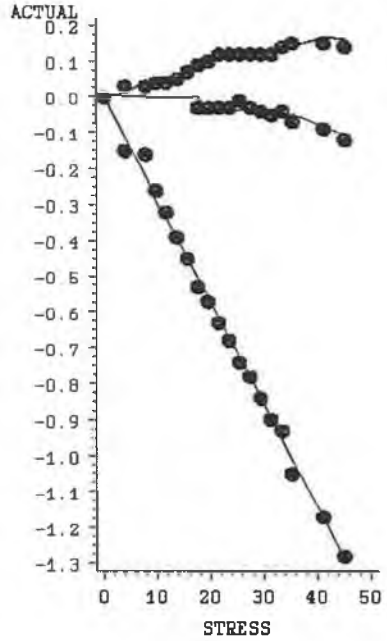
<111> Direction



Circle denote actual data points
Joint Lines denote estimated points

Figure 5.15

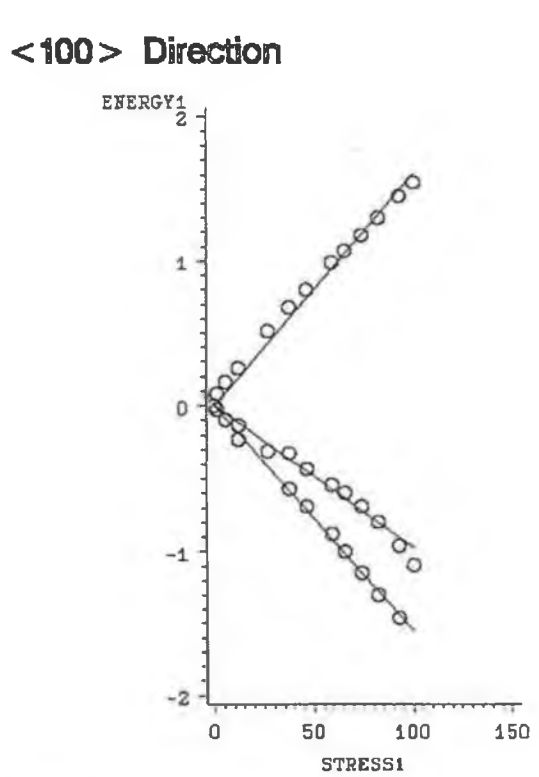
<110> Direction



Circle denote actual data points
Joint Lines denote estimated points

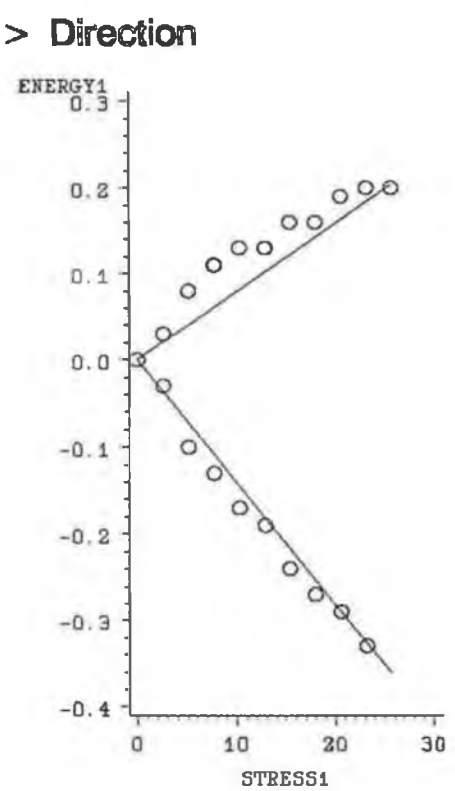
Figure 5.16

Campion et al.(1992) fits with no autocorrelation terms included.



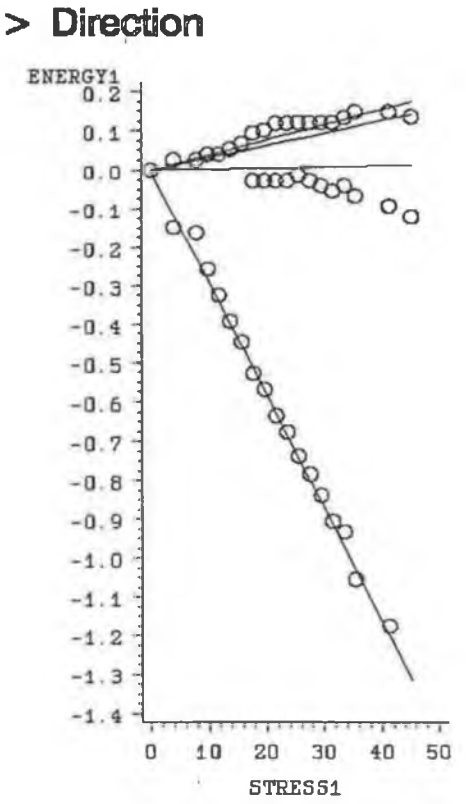
Source: Source Kevin McGuigans data
For Stress in <001> direction

Figure 5.17



Source: Source Kevin McGuigans data
For Stress in <111> direction

Figure 5.18



Source: Source Dr do Carmo
For Stress in <110> direction

Figure 5.19

included are shown in Appendix E and show a marked improvement on Figures (4.7-4.9). Additionally we give amended fits for the Kehoe(unpublished) and Daly(unpublished) examples in Appendix E.

5.5 Variance and Outliers

Given our evidence that the errors for the individual responses are independent and having dealt with the individual autocorrelations for each response we now need to examine the assumption of a constant variance between the responses. If the variance of individual responses were significantly different then the least-squares estimator would not be valid. Several formal tests are available for studying whether or not m responses have equal variance, Neter et al.(1985). In addition to the assumption of constant variance and often distorting the picture is the presence of outliers within the data set. We now outline Bartletts test and implement it on the examples described by Campion et al.(1992) and do Carmo et al.(1988), and then use it in conjunction with the error diagrams of the examples to identify possible outliers. In addition, when a particular data set shows non-constant behaviour of the variance, we look at the implementation of a general multi-response estimator, Stewart and Sorensen (1981), Section (5.3).

5.5.1 Bartletts test

The basic principle underlying the Bartlett test is simple. If we let $s_1^2,,,,,,s_r^2$ denote the sample variances from m normal populations of spectroscopic peak intensities, and let df_i denote the degrees of freedom associated with the sample variance s_i^2 taken from the experimental data, then the weighted arithmetic average of the sample variances using the associated degrees of freedom df_i as weights is the mean square error:

$$MSE = \frac{1}{df_T}\sum_{i=1}^{M} df_i s_i^2 \qquad (5.26)$$

where

$$df_i = \sum_{i=1}^{M} df_i \qquad (5.27)$$

Similarly, the weighted geometric average of the s_i^2, denoted by GMSE, is:

$$GMSE = \left[(s_1^2)^{df_1}.(s_2^2)^{df_2}.....(s_r^2)^{df_M}\right]^{\frac{1}{df_i}} \qquad (5.28)$$

The two averages are equal if all s_i^2 are equal; the greater the variation between the s_i^2, then the further apart the two averages will be. Hence, if the ratio MSE/GMSE is close to 1, we have evidence that the population variances are equal. If the ratio is large, it indicates that the population variances are unequal. The same conclusions follow if we consider log(MSE/GMSE)=logMSE-logGMSE.

Bartlett has shown that a function of logMSE-logGMSE approximately follows a chi-squared distribution with m-1 degrees of freedom for large sample sizes, when the population variances are equal. The test statistic is;

$$B = \frac{df_T}{C}(\log_e MSE - \log_e GMSE) \qquad (5.29)$$

where

$$C = 1 + \frac{1}{3(r-1)}\left[\left(\sum_{i=1}^{r}\frac{1}{df_i}\right) - \frac{1}{df_T}\right] \qquad (5.30)$$

The term C is always greater than 1, Neter et al(1985). For deciding between:

$$H_o: \sigma_1^2=\sigma_2^2=...=\sigma_M^2$$

$$H_a: \text{not all } \sigma_i^2 \text{ are equal}$$

we calculate test statistic B. The appropriate decision for this test is then:

$$\text{if } B <= \chi^2_{(1-\alpha;r-1)}, \text{ conclude } H_0$$

$$\text{if } B > \chi^2_{(1-\alpha;r-1)}, \text{ conclude } H_a$$

For the models described in Campion et al.(1992) and do Carmo et al.(1988) the results of Bartletts test are given in Table (5.4).

Table (5.4) : Bartlett test results for the examples described by Campion et al.(1992) (Cam.) and do Carmo et al.(1988)(do C.) where P.C.V is the approximate probability above the critical value for the chi-squared distribution.

Direction	Chi-square		Degrees of Freedom		P.C.V	
	do C.	Cam.	do C.	Cam.	do C.	Cam.
<100>	14.66	4.94	7	3	@5%	@25%
<111>	64.74	0.17	15	1	@0%	@50%
<110>	43.44	12.83	15	3	@0%	@0.5%

From Table (5.4) it is clear that the we cannot reject H_o for the three directions of the example describe by Campion et al.(1992), although the result for the <110> direction appears significant. We would consider on balance that the performance of the least squares estimator is not completely satisfactory in this case. For the example described by do Carmo et al.(1988) we see that at the 1% level we have a significant result for directions <111> and <110>, implying that the application of the least-squares estimator is still less good. We have, therefore, also investigated the possibility of outliers and the distortion that these may give rise to. From Appendices C and D we see that at high stress the errors for this more complex example tend to increase for certain responses even when

complex do Carmo et al. case even with the improvement effected by the omission of probable outliers. We therefore further investigate for this example the possible influence of unequal variance on the parameter estimates by implementing the Stewart and Sorenson estimator (section 5.3).

5.5.3 Implementation of Stewart and Sorenson for the do Carmo et al.(1988) Example

Minimisation of the Stewart and Sorensen estimator can only be completed using the Powell algorithm with the assumption that the error covariance matrix, Σ equation (5.4), is a diagonal matrix in each direction. Here we estimate the elements of the Σ^{-1} in each function evaluation by substituting the current parameter estimates, θ', into $S(\theta,\Sigma)$, then estimating Σ accordingly. The parameter results on minimisation and their standard errors, (in brackets) are given in Table (5.6).

Table (5.6) Parameter Results

Parameter	Stewart & Sorenson using Powell	Least-Squares using Powell
A_1	-0.005698(0.000055)	-0.005672(0.001)
A_2	-0.015231(0.00002)	-0.015327(0.001)
B_1	0.008613(0.000063)	0.008819(0.0014)
B_3	-0.006720(0.000068)	-0.006844(0.003)

Using results of the Fletcher-Reeves minimisation algorithm on the do Carmo et al. example with the outliers excluded and the autocorrelated terms included, as initial estimates, Table (5.6) shows that the final parameter estimates for both the Stewart and Sorensen and the least-squares estimators are similar. However the covariance matrix of the Stewart and Sorensen estimator shows less variance

within each of the parameter estimates. On the other hand, it seems clear that the improvement in the estimates obtained using the Stewart and Sorensen method is not sufficient to make its use worthwhile when compared with the simpler least squares approach. A choice of either estimator should also lead to identification of the correct transition on the basis of the percentage differences. The difference in the variances, however, could be due to non-linearity in the expectation surface or in the parameters, which must be independently assessed for least-squares. We go on to do this in what follows.

5.6 Covariance Matrix of Least-Squares Estimator

On completion of the minimisation procedure for the least-squares estimator we can use the final hessian matrix for estimation of the covariance matrix,

$$\text{cov}(\theta)=\Omega^{-1} \tag{5.31}$$

where Ω is the hessian matrix of the sum of squares function with respect to the parameter vector θ, Stewart and Sorensen(1981). Additionally

$$(\theta-\hat{\theta})^T\Omega(\theta-\hat{\theta})\leq\chi^2_k(\alpha) \tag{5.32}$$

roughly approximates the 100(1-α) percent Highest Posterior Density region for any k-dimensional subset of the estimated parameters, conditional on the values θ_r for all other parameters. The intervals

$$|\theta_r-\hat{\theta}_r|<Z_{\alpha\backslash 2}(\Omega^{rr})^{-1/2} \tag{5.33}$$

roughly approximate the 100(1-α) percent HPD intervals of the marginal distributions for the individual estimated parameters. Here $Z_{\alpha/2}$ is the critical argument for the one-sided standard normal distribution at the $\alpha/2$ significance level.

The results for the correlations of the parameters for the Campion et al.(1992) example are given in Table (5.7) and for do Carmo et al.(1988) are given in

Table (5.8).

Table (5.7) : Parameter Correlation Matrix for Campion et al.(1992) data.

Correlation	A_1	A_2	B	C
A_1	1.0	0.74	0.79	0.59
A_2	0.74	1.0	0.89	0.68
B	0.79	0.89	1.0	-0.73
C	0.59	0.68	0.73	1.0

Table (5.8) : Parameter Correlation Matrix for do Carmo et al.(1988) data.

Correlation	A_1	A_2	B_1	B_3
A_1	1.0	-0.18	-0.11	0.48
A_2	0.18	1.0	0.22	-0.52
B_1	-0.11	0.22	1.0	-0.22
B_3	0.48	-0.52	-0.22	1.0

Both Tables (5.7) and (5.8) we can see that correlation between the parameters for each for the proposed models is not excessively high, i.e < 0.9.

Table (5.9) : Parameter standard errors for Campion et al.(1992) and do Carmo et al.(1988)

Do C.	A_1=-0.00575	A_2=-0.01525	B_1=0.00875	B_3=-0.00662
$\sigma(\theta)$	0.001	0.001	0.001414	0.002
Cam.	A_1=0.0152	A_2=-0.0126	B=-0.0028	C=0.0319
$\sigma(\theta)$	0.002	0.0014	0.0024	0.0014

Table (5.9) gives the parameter estimates for each of the problems and the appropriate estimated standard errors $\sigma(\theta)$ for each parameter. For the Campion et al. example there are 42 independent points from which we have 41 degrees of freedom for each parameter. For the case described by do Carmo et al. we have 136 independent stress values. Using the standard normal as a close

approximation for both examples we see that the B parameter of the Campion et al. example is not significantly different from 0. For the do Carmo example all the parameter estimates are significantly different from zero using the marginal HPD intervals, and are also significantly different from zero using a Joint HPD interval on the complete parameter set. This implies that we can tentatively accept that there is a significant relationship between the data and the model specified by do Carmo et al.(1988). Often however the parameter and variance estimates can become biased using a multiresponse model. This could, as in the Campion et al. example severely over estimate the parameter variances and hence imply that individual parameters do not contribute significantly, whereas they may be important for a multi-response fit.

5.7 Assessing the Non-linearity

To apply tests normally appropriate for linear regression, such as t-tests or F-tests, to a problem where a degree of nonlinearity is anticipated we need to decide how nonlinear the model actually is. We can identify the nonlinearity of a function by inspecting the sum of squares surface plot . In the linear case we would expect the surface to be smooth and to exhibit an obvious global minimum, Figure (5.20), with the equivalent contour plots of this surface displaying elliptical patterns, Figure (5.21). For a model with some degree of nonlinearity, e.g. Campion et al.(1992), the surface plot and contour plots will bend and show possibly many local minima, Beale(1960) and Goldberg et al.(1983) and Figures (5.22) and (5.23). To determine the amount of nonlinearity present in a specific problem (if any) we need a basis for comparison. This is usually the linear model and we can therefore refer to any departure from the simple form in terms of "closeness to linearity".

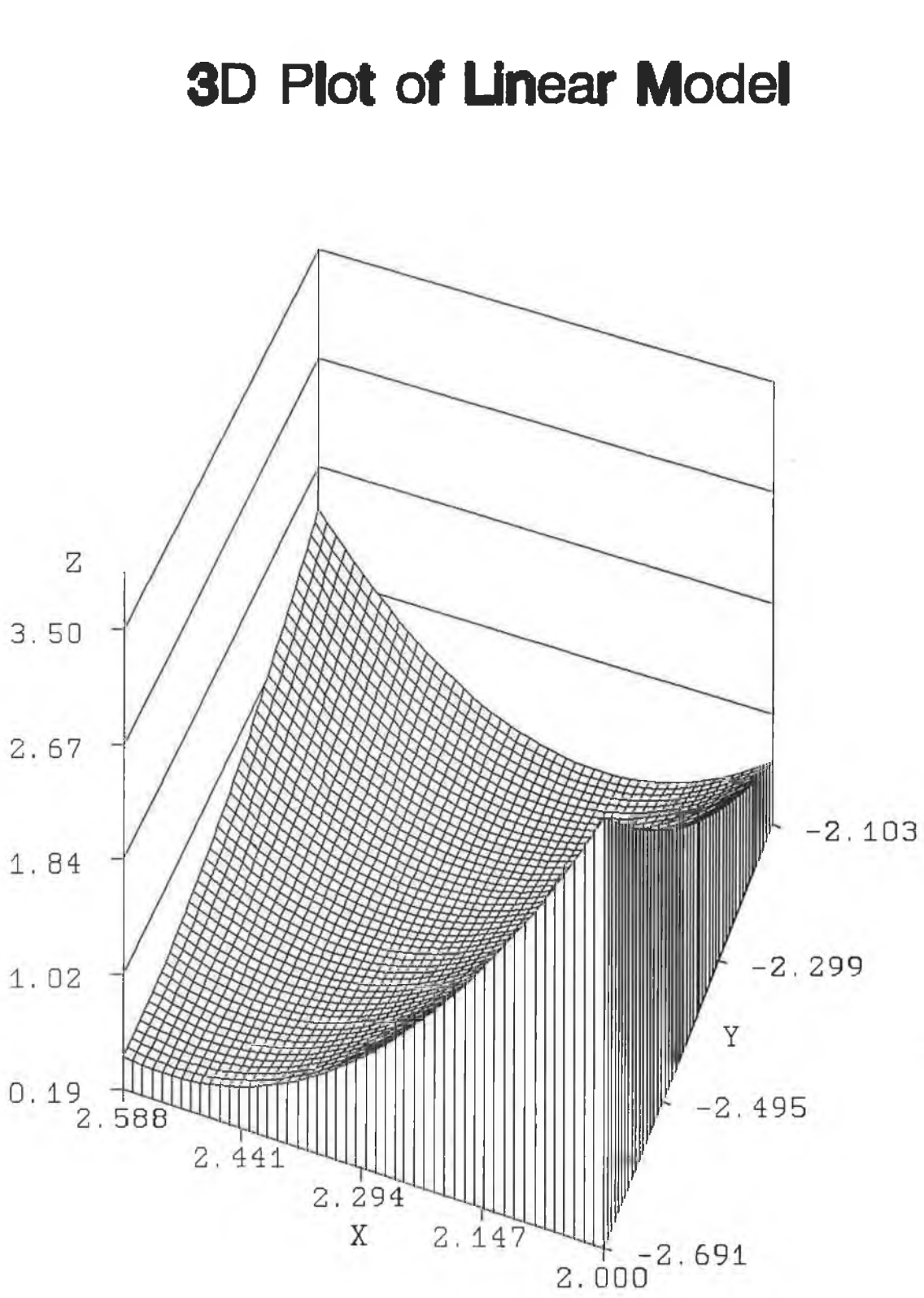


Figure 5.20 3-D surface plot of an artifically derived linear sum of squares. X and Y are linear parameters and Z is the sum of squares.

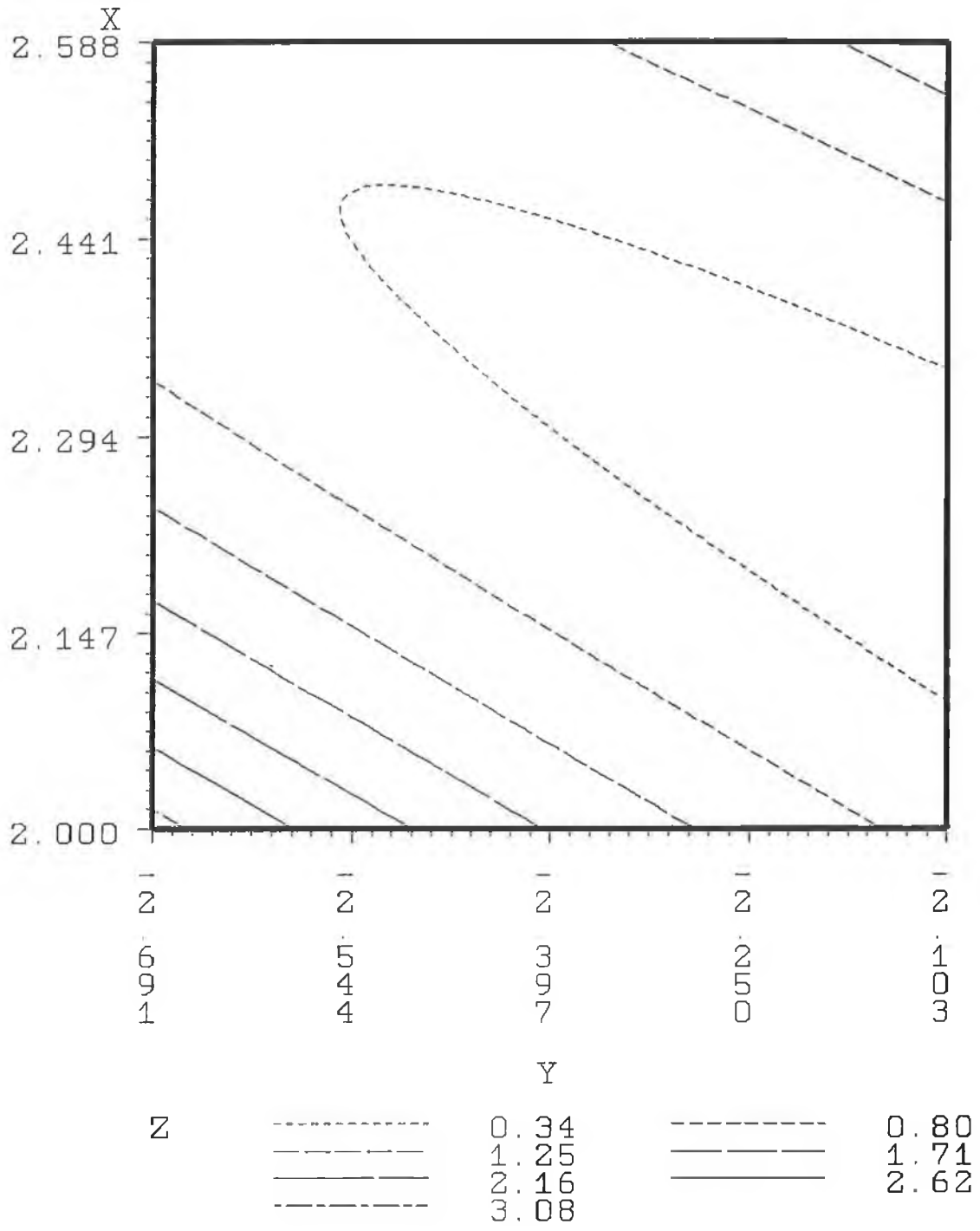


Figure 5.21 Contour plot of an artifically derived linear sum of squares. X and Y are the linear parameters and Z is the sum of squares.

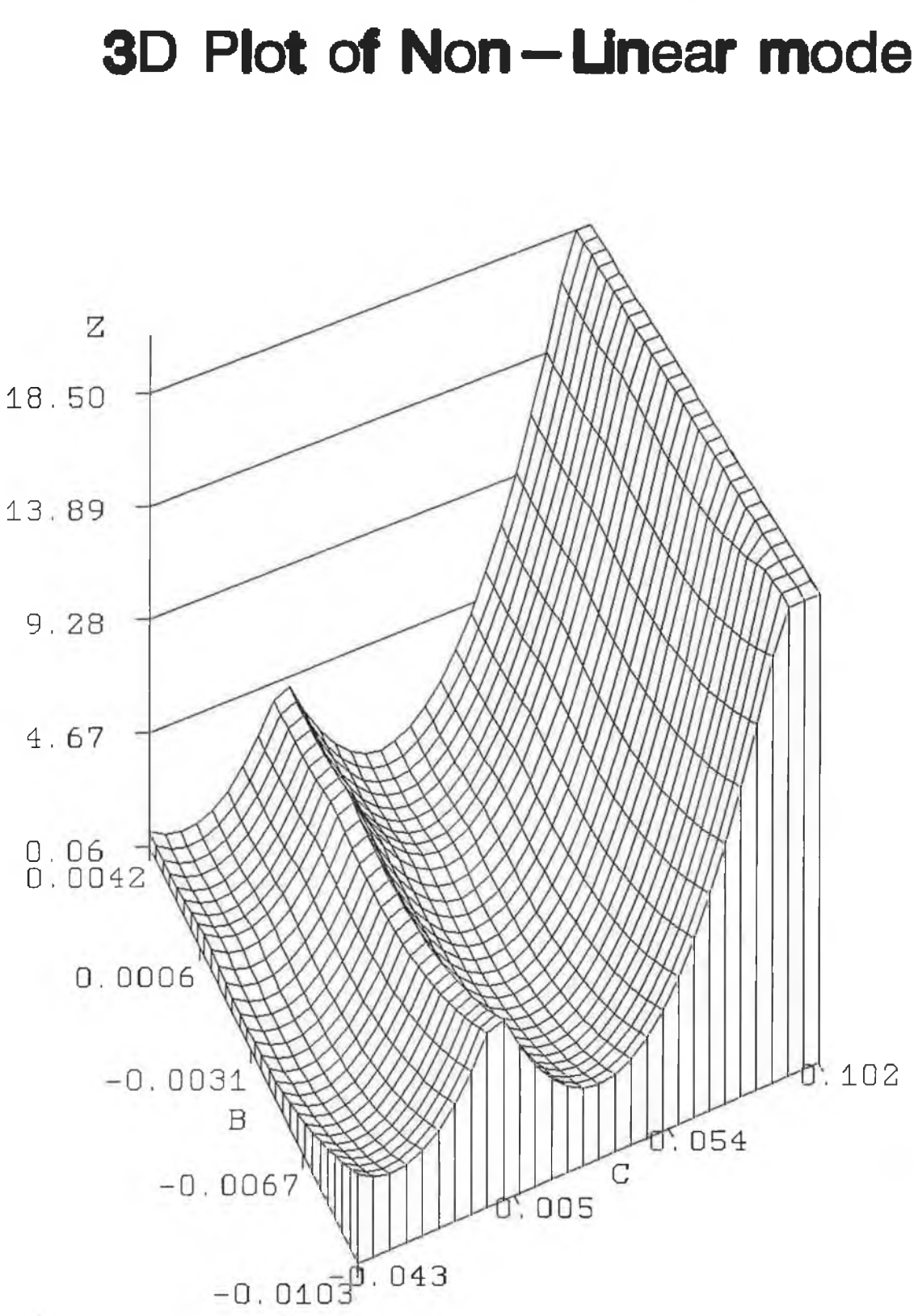


Figure 5.22 3-D surface plot of Beryllium Doped Silicon sum of squares, Campion et al.(1992), where B and C are parameters and Z is the sum of squares.

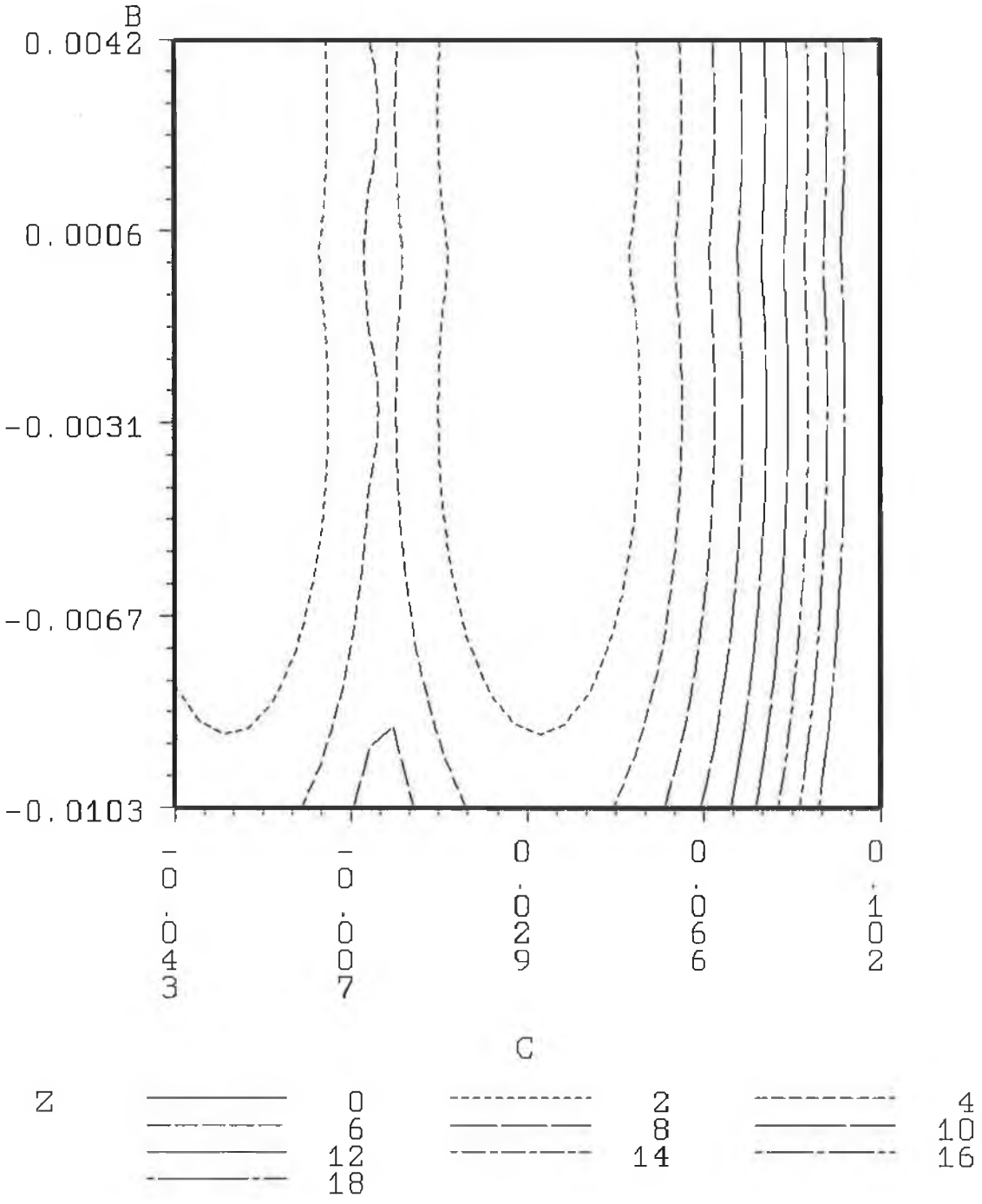


Figure 5.23 Contour plot of beryllium Doped silicon sum of squares, Campion et al(1992), where B and C are parameters and Z is the sum of squares.

5.7.1 Surface Plot and Contour Plots

The surface and contour plots of the sum of squares function for the beryllium doped silicon, Campion et al.(1992), illustrate the situation where the B and C parameters are included. Here we notice that the surface is highly indented and there seem to be at least four local minima. To get a clear image of the sum of squares function, additional plots of all combinations of the parameters must be taken, Appendix F. The value of such images are clearly to give approximate locations of local minima, allowing the experimentalist to assess the complexity and nonlinearity of the expectation surface. Additionally, when we inspect the surfaces of this problem, Appendix F, in more detail we find that close to the optimal parameter values the surfaces are quite well behaved. This implies that linear tests could well be applied in this case provided we have reasonable initial estimates. However, it should be noted that the linear tests apply to 4-dimensions in this case as opposed to just the two illustrated in the 2-dimensional plots. Further, looking at the surfaces plots for the do Carmo et al. case, also given in Appendix F, it again seems evident that the surfaces are close to linear for parameters E and L fixed, (section 4.6.2).

Given the amount of detail necessary to interpret several views, we also provide surface plots for a sample of the parameters for the Kehoe(unpublished) example in Appendix F, showing surfaces with highly non-linear characteristics. Given that this situation arises even for two parameters, then we would expect the sum of squares function to be far from planar with additional dimensions[4]. In order to make a decision on this situation we need a measure which takes account of the full dimensionality of a particular problem.

[4] *In all these examples the linear tests apply to n-dimensions where n=4 for Campion et al. and n=7 for Kehoe(unpublished).*

5.7.2 Acceleration Arrays

Generally, when modelling nonlinear data we need to determine the rate of change on the expectation surface and the parameters. If the expectation surface is planar or close to linear then we can say that the model is intrinsically linear, with the equivalent term for the parameter surface known as parameter effects linearity. When the model is intrinsically close to linear, then a reparameterisation of the model function can give improved parameter estimation and model fit, hence linear tests can more readily be applied. Bates and Watts (1980) describe a technique to determine the parameter effects and intrinsic nonlinearity. Measurement of these terms is based on the estimation of Velocity and Acceleration vectors, and can be described for a single response model as follows :

$$\{\dot{V}\}_{np} = \frac{\delta f(X_n, \theta)}{\delta \theta_p} \qquad \textbf{(5.34)}$$

$$\{\ddot{V}\}_{npq} = \frac{\delta^2 f(X_n, \theta)}{\delta \theta_p \delta \theta_q} \qquad \textbf{(5.35)}$$

where $f(\mathbf{X}_n, \theta)$ corresponds to the function response for independent vectors $\mathbf{x}_n$ and parameters θ).

The intrinsic and parameter effects measurements are then obtained by taking a QR decomposition of the acceleration and velocity vectors. If the intrinsic nonlinearity is acceptably low, there may nevertheless be significant parameter-effects non-linearity and the modeller may wish to seek an appropriate reparameterisation. The parameter-effects measure offers *no* guide however, to a suitable reparameterisation, and in a multi-parameter situation does not identify *which* parameters may be mostly responsible for the nonlinear behaviour, Ratkowsky(1983). On the other hand, if the estimator has small bias, a distribution close to normal and minimum variance, it seems reasonable to

assume that it is 'close to linear'. A problem with estimation of the velocity vectors for the Piezo-spectroscopic data, which does appear to exhibit some degree of non-linear behaviour is, however, the inclusion of multiple responses and allowance for discontinuities within the expectation surface of the model. Estimates of the parameter effects and intrinsic non-linearity can be completed for each response separately, but interpreting the results for the completed model involves additional complications and may lead to undesirable reparameterisation given that direct relation to the physical quantities of interest is our goal. An alternative way to investigate the nonlinearity of the parameters and the model is to investigate the properties of the model using various simulation studies, Ratkowsky (1989).

5.7.3 Lowry's Test

The plots of the sum of squares function do indicate the presence of nonlinearity of the estimates in some cases. An approach which specifically determines the amount of bias that is present in the parameter estimates is needed. We now examine a measure of bias and non-linearity due to Lowry, Ratkowsky(1983), (developed originally on unireponse models) and propose its application to the multi-response case. Consider the regression model

$$\mathbf{Y}_m = f(\text{Stress}_n, \underline{\theta}) + \varepsilon_m \qquad \textbf{(5.36)}$$

where $f(\text{Stress}_n,\theta)$ (n=1,..,N) may be linear or nonlinear in the parameter vector $\underline{\theta}$. Assuming that $\underline{\theta}$ is known, we can generate a series of independently identically distributed Normal (NID) errors ε_m (m=1,....,M) having mean zero and variance σ^2 corresponding to Stress_i , resulting in a set of random variables denoted Y_m^+, that is

$$Y_m^+ = f(\text{Stress}_n, \underline{\theta}) + \varepsilon_m \qquad \textbf{(5.37)}$$

The LS estimate of $\underline{\theta}$ corresponding to this generated model/data set combination is denoted $\underline{\theta}'^+$. By subtracting the error term from the individual

responses we can get the following equation

$$Y_m^- = f(Stress_n, \underline{\theta}) - \varepsilon_m \quad \textbf{(5.38)}$$

With the LS estimate of $\underline{\theta}$, equal $\underline{\theta'}^-$. For a linear model it is true that for each parameter θ_i (i=1,2,...,p)

$$(\theta'^+_i - \theta_i) = -(\theta'^-_i - \theta_i) \quad \textbf{(5.39)}$$

For a non-linear model this is not true, Ratkowsky(1981). It follows that the statistic can be used to measure the nonlinear behaviour of the LS estimator of

$$\Psi_i = \frac{(\theta'^+_i - \theta_i) + (\theta'^-_i - \theta_i)}{2} \quad \textbf{(5.40)}$$

θ_i. Since the distribution of θ'^+_i is identical with the distribution of θ'^-_i, each being an alternative LS estimator of θ_i, denoted θ'_i, it follows that

$$E(\Psi_i) = E(\theta'^+_i - \theta_i) = E(\theta'^-_i - \theta_i) = Bias(\theta'_i) \quad \textbf{(5.41)}$$

By obtaining the correlation between θ'^+_i and θ'^-_i, equation (5.42), we can determine the linearity of the model. For non-linear models, the correlation will vary between -1 and 0, with close-to-linear models producing values close to -1.

$$Corr(\theta'^+_i \theta'^+_i) = \frac{Cov(\theta'^+_i \theta'^-_i)}{\sqrt{Var(\theta'^+_i) Var(\theta'^-_i)}} \quad \textbf{(5.42)}$$

5.7.4 Results of Lowry's Test

Simulation runs for the examples described by do Carmo et al.(1989), Campion et al.(1992) and Kehoe(unpublished) are given in Table (5.10). The results show how the non-linearity changes according to problem type.

Table (5.10) : Lowry Correlation Coefficients (L.C.C.) for the problems of Campion et al.(1992), do Carmo et al.(1988) and Kehoe(unpublished).

Campion.	L.C.C	do Carmo	L.C.C	Kehoe	L.C.C.
A_1	-0.79	A_1	-0.68	A_1	-.926
A_2	-0.95	A_2	-0.72	A'_1	-0.317
B	-0.97	B_1	-0.60	A_2	-0.988
C	-0.78	B_3	-0.39	A'_2	-0.159
				B	-0.206
				D_1	-0.903
				E	0.3916

For the example described by Campion et al.(1992) (and similarly for that of do Carmo et al.(1988)) the L.C.C. estimates show high negative correlation indicating that the parameters are not far from linear. However, this cannot be said to be the case for the example of Kehoe(unpublished). Here, some of the correlation coefficients are extremely close to zero indicating that these parameters possess considerable nonlinearity. An interesting point with respect to the do Carmo et al. example arises when the E and L parameters are considered as estimated by the model as opposed to being separately obtained, section 4.6.2. Here the degree of non-linearity *markedly* increases, as illustrated by the surface plots, Appendix F. This suggests that when such parameters are estimated directly from the fit, as for example is also the case for the Kehoe data, then the model is markedly less linear. Experimentally, estimation of these where possible, prior to obtaining an overall fit, clearly greatly improves the stability of the set of parameter estimates as a whole. However, determining the unstable parameters or those most likely to adversely influence the fit is a difficulty. In many cases, they will be known to the experimenter because they relate to an aspect of the physics for which the expected behaviour is well known, which is supported by exact theory or for which independent

measurement is possible. However, this is not always the case and it is reassuring to note that our method *did* achieve identification of the correct transition and reasonably accurate parameter estimation in the do Carmo case, even when we allowed E and L to vary. Additionally, the exclusion of the autocorrelation terms increases the L.C.C. estimates, Table (5.11).

Table (5.11) : The L.C.C. estimates for the do Carmo example when the autoccorelation terms are not included.

Parameter	A_1	A_2	B_1	B_2
L.C.C	-0.90	-0.67	-0.96	-0.80

5.7.5 Other Simulation Results

The results from the L.C.C indicate that the parameters for the Campion et al. and do Carmo et al. models do not possess a significantly high degree of nonlinearity. This agrees with the surface plots discussed in section 5.7. However, for the Kehoe(unpublished) example we find that the L.C.C.s and surface plots indicate marked nonlinear tendencies. An alternative method to determine the appropriateness of the parameters for each of the models, is to investigate the properties of the distributions of each of the parameters through a simulation study as mentioned earlier, Ratowsky(1983),(1989). For a normal distribution we would expect the coefficient of skewness to be approximately normally distributed with mean zero and standard deviation $(24/N)^{1/2}$. The coefficient of kurtosis for large N is approximately normally distributed with mean 3 and standard deviation $(24/N)^{1/2}$. Table (5.12a) and (5.12b) give the mean, standard deviation and coefficients of skewness and kurtosis for the Campion et al.(1992) example and do Carmo et al. (1988) examples.

Table 5.12a : Distributional properties of the Campion et al. example taking 1000 simulations.

Campion	a1	a2	b	c
mean	0.01519	-0.0126	-0.00281	0.32
$\sigma(\theta)$	0.0003**	0.00015**	0.000189**	0.000299**
Skewness	-0.12496	0.0406	0.0458	0.837**
Kurtosis	5.0**	0.362**	-0.04921**	4.06**

Note *Probability< 0.05; **Probability < 0.01.

Table 5.12b : Distributional properties of the do Carmo et al.(1988) example taking 200 simulations.

do Carmo	a1	a2	b	c
mean	-0.00568	-0.01529	-0.008741	0.006729
$\sigma(\theta)$	0.000108**	0.000115**	0.000122**	0.000228**
Skewness	0.39332	-0.825**	-0.471	1.5533**
Kurtosis	0.438**	0.357**	3.636	1.5033**

Note *Probability< 0.05; **Probability < 0.01.

The results for the Campion and do Carmo examples show that the variance and kurtosis properties of the distributions of the parameters are significantly biased but that the mean and skewness properties are not, so that we would conclude that variance is over estimated. Such effects commonly occur in nonlinear models and might typically suggest a reparameterisation of both models. However, in the Campion et al. case it already relies on a reduced or minimal linear form for which expansion results in an expression having little physical meaning. Consequently, reparameterisation seems to be ruled out in terms of its likely value in explaining these distributional features. The conclusion then must be that the model possesses some level of intrinsic nonlinearity. For the do

Carmo example we see that a reparameterisation would be equally difficult for a problem of such high dimensions and equally unsuitable in terms of the physical interpretation. We conclude that the model parameters are not far from linear since the means are all unbiased and the surface plots show close to linear properties near the minimum, but that they again possess some intrinsic nonlinearity, which would generally be expected in any event in multiresponse models, Bates and Watts(1988).

The results for the Kehoe data (unpublished) are not so difficult to interpret since the covariance matrix of the estimates, equation (5.31), shows distinctly non-linear properties (i.e. the inverse of the hessian matrix gave some negative results on the diagonal elements). This may be due to the mesh routine which is used to approximate the hessian of the $S(\underline{\theta})$ or the fact that the function is highly nonlinear in at least one of the 7-dimensions. However, the results obtained here are in agreement with the surface plots and the L.C.C estimates. We can therefore conclude that, as with the other piezo-spectroscopic models, there is some intrinsic nonlinearity. Furthermore, in the absence of any independent estimates of subsets of the parameters, it is not possible to rule out the presence of a parameter effects component, since an alternative transition, implying a redefinition of the parameters, may be appropriate if unlikely.

5.8 Conclusions

The performance of the least-squares estimator and the Powell-Shell technique in modelling piezo-spectroscopic data is encouraging compared to those described by Stewart and Sorensen (1981), Bates and Watts(1988) and Seber and Wild(1989). Both ease of computation and time taken to achieve a good fit to the data are considerably improved. The parameter results show little difference when compared to estimates obtained by the Stewart and Sorenson method, although nonlinearity is not specifically accounted for. We can conclude therefore, that no correlation exists between responses and that the variance within responses is constant for piezo-spectroscopic models.

One of the predominant features of such data is auto-correlation which we have demonstrated can be handled satisfactorily using a two-stage process on a first-order autocorrelation model. Nonlinear features are less straightforward to assess and are highly data dependent. Estimating as many parameters as possible, independently and prior to overall fitting, clearly reduces the difficulties, but does not solve them. Bias in the parameter estimates may be assessed, and, and in the case of close-to-linear models, results of linear tests may be regarded as reasonably robust. This is confirmed by the results given by the surface plots and the Lowry coefficients. However the underlying characteristics of the simulation studies suggest that there is some intrinsic nonlinearity in piezo-spectroscopic models, which would bear further investigation.

The Stewart and Sorenson covariance estimator applies only asymptotically to close-to-linear models. Consequently the bias it produces is likely to be excessive when obtaining covariance estimates for piezo-spectroscopic models and should not therefore be considered to be particularly reliable in the majority of our examples. Inflation of the parameter variance estimates in such cases is

to be expected.

Chapter 6

The Tool

6.1 Introduction

In order to identify an appropriate defect for spectroscopic data, the approach summarised in Figure (4.1) was proposed. The fitting technique proposed in Chapter 4 seeks to facilitate and enhance the identification process. Some initial steps prior to optimisation must be carried out, which can be categorised as follows :

I) Reduction of the list of potential transition-type by comparison of the data set to known transitions, Fowler(1968).

II) Ranking the data in order to find a line matching pattern such that the minimum number of comparisons between the predicted eigenvalues and the experimental energy values will take place, section (4.5.2).

III) Seek good initial estimates for the Powell-Shell method, section (4.5.6).

We then proceed to obtain the estimates for the secular matrix using the Powell-shell method, section (4.5.4).

A system incorporating this approach has been implemented successfully for various data sets, described in Chapter 4. We have developed appropriate user interfaces to execute the initial data manipulation procedures, to supply parameters to the Powell-Shell optimisation and to select for the basis of these results the appropriate transition type. Additionally, we have incorporated a further matrix interface facility to enable the user to build his/her own secular matrix and enable a preliminary identification to be obtained.

In this chapter we describe the specification of TranId, demonstrate an

implementation of the system as an identification tool, analyze the speed of the method with respect to mainframe and PC machines, and finally discuss the software and hardware environments for the implemented tool.

6.2 System Design and Specification.

From previously, Figure (4.1), the precision of the fitting procedure has been recognised as an enhancement to the task of identification of spectroscopic data. This procedure requires data sorting and a database of possible transitions. Using this information we can then obtain a fit and hopefully identify the transition occurring.

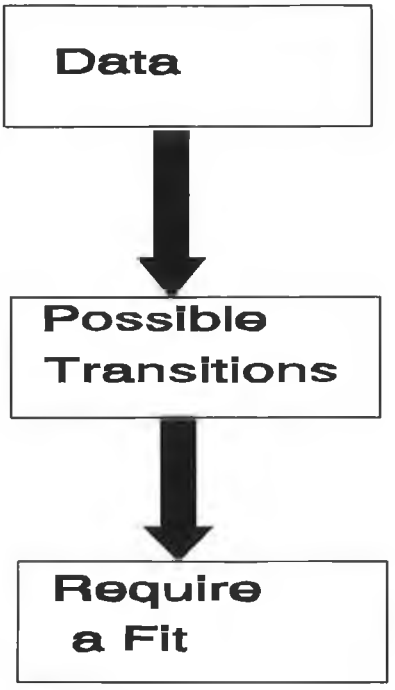


Figure 6.1 Basic system description

In the design of the system the user requirements are the initial starting point. Figure (6.1) describes the simplified system requirements as a subset of those shown in Figure (4.1). Elements of the various components have the following features ;

Data, for which there are two potential sources:

I) Manual entry of each record corresponding to the stress value applicable and the corresponding energy (experimental) value,

II) Importation of the previously set-up file,

Transition Box :

From the database of potential transitions we can focus on possibilities by:

I) identifying which points belong to individual lines,

II) comparing these lines to Table (3.2) to short-list the possible transitions.

Fitting:

Finally when a suspected transition type or types is short listed, estimates of its parameters can be obtained by fitting.

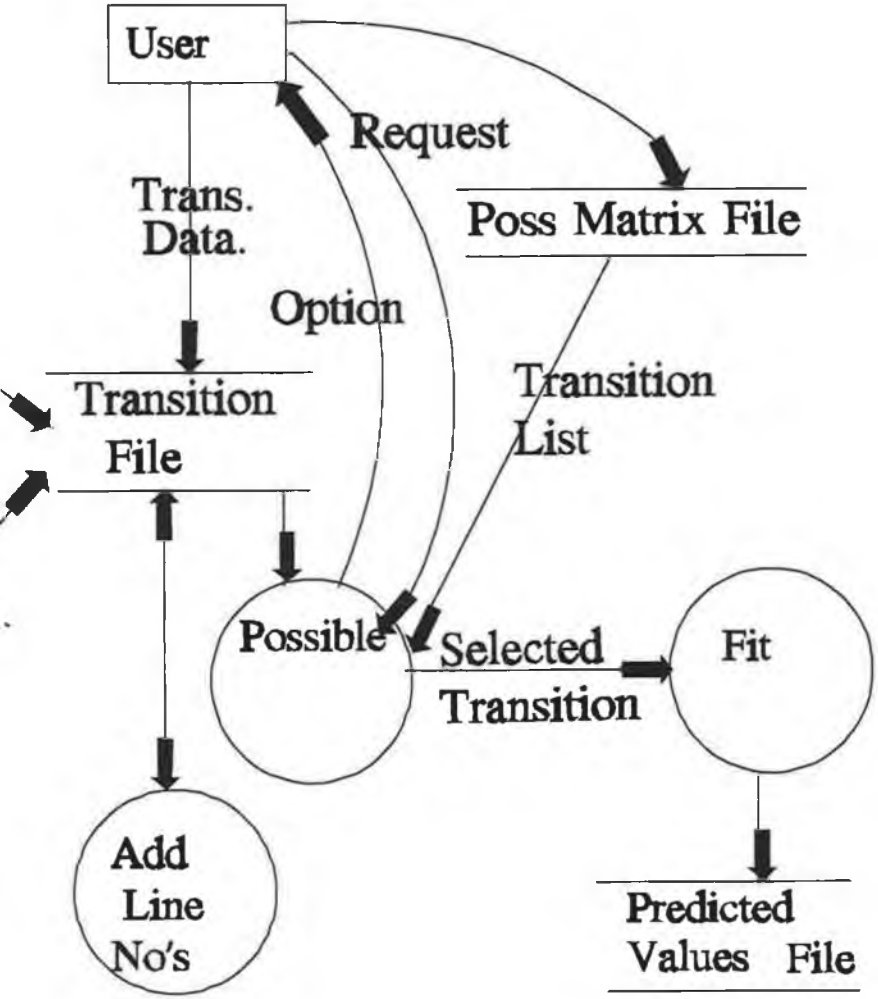
Figure (6.2) describes the data flow through the prospective system, corresponding to the above processes. This diagram corresponds to a Data Flow Diagram (DFD) which is a network representation of a system, De Marco (1979), and is made up of only four basic elements:

- data flows, represented by named vectors,
- processes, represented by circles or "bubbles",
- files, represented by straight lines,
- data sources, represented by boxes.

The DFD is useful in describing the overall specification of the system, and in particular explain the functionality of each of the above black boxes.

Raw
Import
Data
Import
Data
Spectrum
Spectrum
File
Extract
Data

Figure (6.2)


User
Request
Trans.
Data.
Poss Matrix File
Option
Transition
File
Transition
List
Possible
Selected
Transition
Fit
Add
Line
No's
Predicted
Values File

However in order to develop a suitable system it was clearly necessary to obtain detailed specifications from the proposed end user, Jones(1980). We discuss details of the specifications worked out with the D.C.U. solid state/optronics group in what follows.

6.3 Automation in the identification of transition data

In this section a broad overview of the most important points needed to automate identification of a transition are given. Initially we look at data entry, then move to data to line association and from there to initial parameter estimation. Final parameter estimation is then described together with a set of procedures to evaluate the statistical implications of the proposed parameter estimates.

6.3.1 Data Selection

For the first step, two methods for entering and selecting the data set from any particular experiment are permitted.

I) :- This allows the user to enter a complete data set for the experiment through a screen data entry facility. There are five columns in which the user will be able to enter the experimental data. Each column corresponds to Stress, Energy, Line number, polarisation intensity and line width respectively, as in Table (6.1).

Table (6.1)

Stress	Energy	Line No.	Polarisation Intensity	Line Width

II) :- A facility to import a file created with just stress and energy as its columns is also included. This file is checked to see that it conforms with the TranId file format, and if so, is assigned a signature name that uniquely identifies it.

6.3.2 Selecting a Transition

Having entered a data file into the system the next procedure to be undertaken is the association of data points to their corresponding line numbers. Two methods are available to the user.

I) The first approach enables the user to manually enter the line numbers associated with each data record through the procedure described in section (6.2.1).

II) The second method is a manual graphical procedure which allows the user to draw the lines on the screen with the aid of the mouse. The points are then associated to the nearest lines. The line numbers are then sent to the data files were they are associated to their corresponding stress - energy points.

Crossed lines will give rise to inaccurate fits since the eigenvalues generated from the selected matrix will all be ranked according to highest energy value. This implies that points should be assigned by rank as opposed to slope, McCarren and Ruskin (1994).

Here the user can start a new line by applying the <F4> key, Figure (6.3), and continue the line by moving the right or left arrow keys (← and →). When satisfied with the match of points with the actual lines the user can then initiate a new line by pressing the <F4> key again or leave the option by pressing the <ESC> key. When this process is fulfilled for the three families of directions a choice of possible transitions is given. From this list of transitions[1] the most likely candidate can be selected, with each choice corresponding to a pertinent file (TRD file), which describes the secular matrix of the transition in question. This file comprises the most recent initial estimates, data files applied to this transition, the applicable family of directions and the description of the transition. If a transition is selected then it's description will appear at the bottom of the screen.

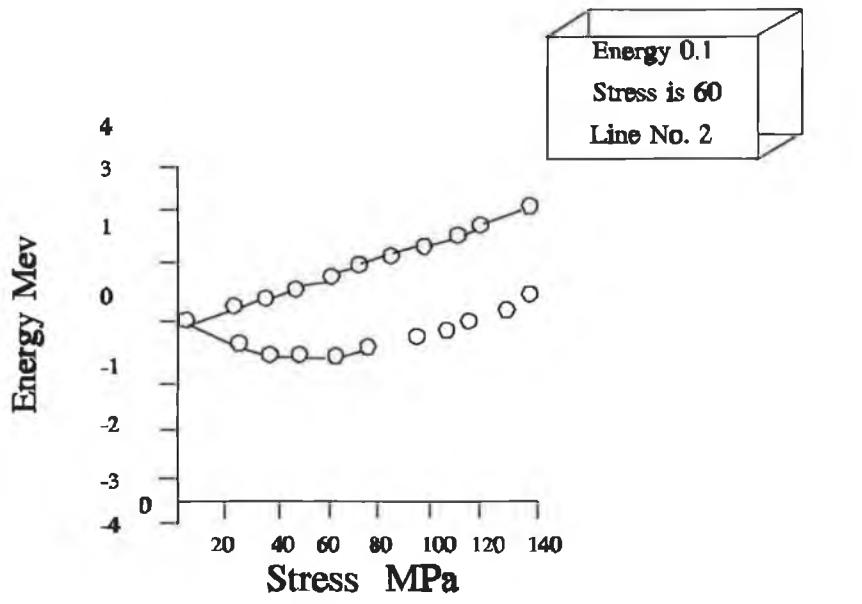


Figure 6.3 The line matching technique.

[1] *The list of transitions given is stored in a file denoted Base.dat.*

6.3.3 Obtaining a Fit

Having defined which points belong to which line the user must then obtain initial estimates for the parameter values and then obtain a fit based on the initial estimates. The reason for doing this is that even where non-linearity is present in these models, quite reasonable initial estimates can be obtained for the complete set of data by using the estimates derived from a fit on a subset drawn from the linear or close-to-linear part. A description of the processes and requirements needed to obtain a fit is given:

I) The user has the option of selecting the number of points that they perceive to be relevant to the initial estimation. This is done graphically by accepting certain points to the left of the line, Figure (6.4).

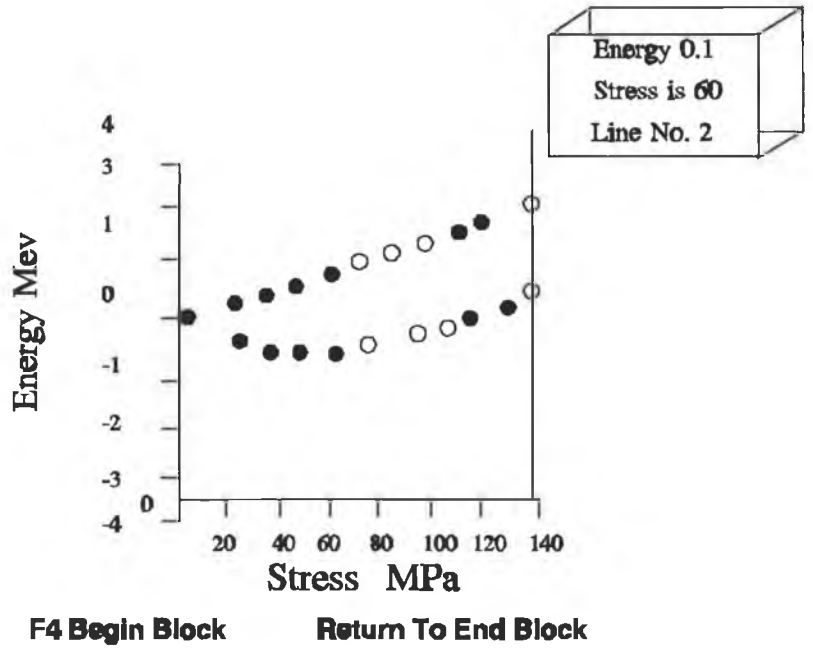


Figure 6.4 The data capture interface.

i.e by moving the line from left to right the user can select the relevant data by just clicking the mouse on the point where they wish the initial estimation of the parameters to be carried out.

II) The user can also select a subset of data by specifying in the secular matrix

file, n points from the N total stress points applicable. The first n points only will then be used for initial estimation.

In order to complete an estimation of the parameters the user has to choose a secular matrix or matrices which are likely to be appropriate.

III) The system automatically selects a matrix or matrices on the basis of the information known, but allows the user to confirm or reject the choice. The system's choice of transitions is performed by comparing the number of lines obtained in each data file to a database of transitions and their corresponding secular matrices similar to those given in Table (3.2).

IV) The user can also select and build alternative secular matrices by using the secular matrix generating routine.

Having completed the above tasks an estimation of the secular matrix parameters can now be carried out.

6.3.4 Machine Specification

In order to implement the user specification we need to take into account the environment within which the final product should be implemented. Typical requirements include:

- availability of computer hardware to the client,
- hardware cost,
- implementation speed,
- graphics capabilities,
- the number of users for the system.

The above points and the client's requirements can be satisfied by implementing the software on a PC system running on DOS 3.3 or above, since D.C.U. optronics have access to numerous PC's running on the DOS environment. Further if the group require the software to be run on higher performance

equipment then the cost of updating their current hardware will not be a major consideration, since the cost of high performance PC's has reduced considerably in recent years. Additionally the processing speed (section 6.5) and graphical capabilities of these machines are favourable with respect to a mainframe environment.

Setting up the data and secular matrices will require a finite period of time compared to the fitting technique which requires no user interaction once started, implying that a single user system is appropriate.

Additionally, the core C software has been implemented on PC,VMS and Unix enviroments, indicating that the system is not specific to any particular enviroment.

6.4 Tranid Requirements

6.4.1 Software Requirements

The Powell-Shell method was originally implemented on prototypes written in Borland's Turbo Pascal(5.5) and Vax Pascal(2.0), Appendix G. The final system software was however written in the Borland's Turbo C programming language, Barakati(1989). The implementation of this software followed as closely as possible the ANSI conventions, Kernighan and Ritchie (1988), deviating only when the Borland's graphics facilities were implemented. Appendix G gives the code written in C for the Powell-Shell fitting technique. It can be compiled on any operating system without major changes in the code and requires the data files and TRD files in to be in TranId format, Appendix G.

The software libraries used to implement the Powell and eigenvalue algorithms are described in the numerical recipes routines NRUTIL.C and NR.C, Press et al (1990) and the menu libraries included in the software were designed in conjunction with the Parker menu system, Parker (unpublished). Additionally an interface between TranId and Surfer(1989) was used to obtain the 3-dimensional

surface plots and the 2-dimensional contour plots and to implement the Surfer software two batch files *contour1.bat* and *contour2.bat* were used, Appendix G.

6.4.2 Hardware Requirements

Section (6.2.1) recommends the proposed software to be executed on a single user PC. In order to obtain fits in a reasonable time compared to that of a mainframe machine we propose that the PC possess a 486 microprocessor, with a minimum 640k of RAM. The disk space required to use the system is minimal but a minimum of 2 Megabytes is preferable due to inclusion of the graphic font files, the database of transitions file, the executable code and a comfortable amount of space for numerous TRD and data files.

6.5 Timing Analysis of Software

The implementation of the software of TranId on the PC environment requires an analysis of times to convergence of the Powell-Shell method on the various harware environments. Using column 4, Table (4.6), as initial estimates with a full data set, we obtained a final solution given in column 5, Table (4.6), in 1 hr 37 minutes 46 seconds on a Dell 486 PC running on the microsoft Dos operating system, and 1 hrs 3 minutes and 35.9 seconds on a VAX mainframe running on the VMS operating system, using the Powell minimisation algorithm. For the Powell versus Fletcher-Reeves algorithms, the convergence times and initial and final estimates of the Powell-shell are given in Tables (6.2a) and (6.2b) for both the do Carmo et al.(1989) and the Campion et al.(1992) problems.

Table (6.2a)

Parameter	Powell-Initial	Final-Estimates	F. Reeves Initial	Final-Estimates
A_1	-0.005	-0.00634	-0.005	-0.00625
A_2	-0.01	-0.01496	-0.01	-0.0148
B	0.01	0.009164	0.01	0.009466
C	-0.01	0.0057	-0.01	-0.00584
E	2.7	2.7145	2.7	2.6999
L	0.45	0.6136	0.45	0.45002
Time		3:12:59:62		2:21:8:19
NSSq Min		8.63932		9.5295

Table (6.2b)

Parameter	Powell-Initial	Final-Estimates	F.Reeves Initial	Final-Estimates
A_1	0.0175	0.0162	0.01	0.0162
A_2	-0.015	-0.0125	-0.01	-0.0125
B	-0.002	0.00311	-0.001	0.00311
C	0.01	0.0329	0.01	0.0329
Time		1:36:17		0:57:56
NSSq Min		0.18342		0.18348

A comparison of tables (6.2a) and (6.2b) show that Fletcher-Reeves algorithm comes to convergence quicker for both problems, since it uses fewer functional evaluations of NSSq(P). However for the do Carmo et al.(1988) example, table (6.2a), the final estimates found by the Powell algorithm gave a lower final NSSq(P) value. The initial estimates given in the Campion et al.(1993) example are different. This is because the sum of squares surface for this particular problem has several possible global minima, Figure (5.19). (diagram of Campion et al. sum of squares 3d surface) For example when the intial estimates used for both algorithms were alternated the final NSSq(P) value was approximately equal to 0.32, in both cases.

Using the Norton Utility "System Information " a comparison of the 486 to the lowest range of IBM compatible XT microprocessor is pertinent. The 486 gives a computing index of 55.1 compared to the XT processor. This implies that the 486 machine can compute and process information 55.1 times faster than the XT processor.

To compare different processing times on the PC platform, two problems are examined: The first deals with a relatively large matrix, (size= 8x8 including a spin orbital interaction, do Carmo et al(1988)) and the second with a relatively small matrix (size= 2x2, Campion et al (1992)), on 486 and 387sx[2] machines respectively. Table (6.3) contains the details of the comparison between the two machines.

[2] *The 387sx refers to a machine with a 386sx microprocessor with an 80387 maths coprocessor installed, running at approximately 15 times faster than the XT machine.*

Table (6.3)

Test	Machine	Problem	No. Parm	No of Points in 100 111 110.	Time to converge.
1	486	*	6	33 48 56	1:43:0.0
2	486	*	5 L	33 48 56	1:29:0.0
3	486	*	4 LE	33 48 56	0:44:28.23
4	486	*	5 E	33 48 56	0:52:26.7
5	486	**	4	Full Data Set	0:0:15.76
6	486	**	4	5	:0:14.78
7	486	**	4	10	0:0:30.0
8	387sx	*	6	33 48 56	7:17:32.0
9	387sx	*	5 L	33 48 56	5:57:0.0
10	387sx	*	4 LE	33 48 56	2:54:55.8
11	387sx	**	4	Full Data Set	0:1:12.72
12	387sx	**	4	10	0:2:26.87
13	387sx	**	4	5	0:1:9.20

Here the first two columns indicate number of test and machine type and column 3 indicates the problem tested (i.e. (*) do Carmo et al.(1988) and (**) Campion et al.(1992)). Column 4 describes the number of parameters included in the problem, and lists the parameters previously estimated by direct measurement. Finally columns 5 and 6 give respectively the number of stress points used in each stress direction and the time to convergence for the

parameter estimation.

From Table (6.3) the dimension of the matrix given by do Carmo et al.(1988) accounts for the bulk of the time taken in the parameter estimation. For example a comparison between tests 1 and 5 shows a difference of approximately 1 hour 42 minutes 45 seconds on the 486 machine, between a $16x16$[3] and 2x2 matrix. However the difference between tests 8 and 11 is approximately 7 hours 16 minutes and 20 seconds on the 387sx . Clearly the effect of machine type is also considerable. Nevertheless the times taken to converge decrease dramatically for the dimension of the matrix. Time consumption can be partitioned even further, since the number of parameters to be estimated also slows convergence. Tests 2 and 3 show a time variation of almost 45 minutes, yet only one parameter is held constant. The consequence of keeping a parameter constant will obviously depend on which parameter or parameters are involved (tests 2 and 4). However the cause of the reduction in time can in some instance (or "all" instances of exclusion from overall estimation) be explained in terms of the number of functional evaluations involved. This number is known to be directly related to the number of parameters in the functional form of the model for the Powell algorithm, Walsh (1975),Press et al.(1990).

An interesting apparent contradiction arising from the experiment summarized in Table (6.2) is that the time to convergence does not necessarily decrease as the number of points in the estimation decreases, (tests 12 and 13). In these cases this may be rationalised by the observation that the functional form of the sum of squares is changed owing to the truncation of the data set and hence initial estimates are not particulary good.

---

[3] *This matrix effectively doubles because of the inclusion of spin orbitals and hence a complex component is introduced. From section 3.7.3, a hermitian matrix is equivalent to a real symmetric matrix of double dimensions.*

6.6 Review

In this chapter a specification of appropriate user, machine and software specifications has been given and the tool developed to meet these requirements has been described, namely the software system known as TranId. From this a demonstration of times to convergence for different problems showed that size of matrix, number of parameters and capability of hardware had an adverse effect, whereas size of data set is fairly unimportant.

Conclusions and Further Work

In this thesis, the three general approaches to modelling spectroscopic data have been reviewed and assessed with respect to their relevance for piezo-spectroscopy. The approach which was found to be the most appropriate involved working with realisations of the spectrum, although no single method within this group was immediately applicable to piezo-spectroscopic data.

From the analysis of Chapter 4 it appears that while a more concise format for the characteristic polynomial or the eigenvalue equation may well prove satisfactory, the Powell-Shell approach of working directly with the secular matrix has been the most successful for the application described. Positive improvements with respect to time of convergence and parameter estimation have been obtained together with accurate transition identification in all cases. The use of two core algorithms, namely the Powell and Fletcher-Reeves has enabled us to independently confirm results for any particular problem. The Powell appears to be more robust with respect to initial estimates, but the Fletcher-Reeves is favourable with respect to time of convergence, since it uses considerably fewer functional evaluations. A further development should concentrate on reducing the number of functional evaluations required in locating a global optimum by imposing boundaries on parameter estimates with the optimisation algorithm.

The Powell-Shell primarily uses a least-squares estimator for multi-response models. The principal estimators for such data use the determinant criterion, Bates and Watts(1988) or the extension described by Stewart and Sorensen(1981) . The latter has been widely used for modelling multi-response data with *missing values*, Malcata et al.(1993), which is a feature of piezo-spectroscopic analyses. The Stewart and Sorensen estimator is, however, cumbersome to use with respect to the estimation of a large number of

parameters, while the least squares estimator is simpler, has other favourable characteristics and performs equally well in most cases. The interpretation of the parameters obtained in a physically meaningful way and the identification of the underlying transition has proved possible for all examples considered.

A major component of the errors in piezo-spectroscopic data analysis is the presence of autocorrelation and we have modelled this successfully using a two-stage procedure. The inclusion of extra autocorrelative parameters increases the complexity of modelling any data set and the use of multi-response estimators will in general have to allow for nonlinearity. Unfortunately, such estimators are frequently quite poor in close-to-linear models, which occur quite frequently in piezo-spectroscopic data and consequently, have a tendency to influence the parameter variance estimates. Some nonlinearity, which is probably intrinsic, does appear to be present in piezo-spectroscopic models and suggests that further investigation of the statistical properties of multi-response models with missing data would be desirable. In particular, diagnostic analysis has received scant attention to date and is predominantly confined to error plots and covariance estimates, Malcata et al.(1993). Additionally, it is not particulary clear if nonlinearity measures such as intrinsic and parameter effects, Bates and Watts(1988), are wholly applicable to multi-response problems, and these have yet to be seriously addressed in the current literature.

Finally, incorporating all the features discussed in Chapters 4 and 5, we have developed a complete software system. As part of this system a user interface was designed to allow the matching of experimental data points to proposed lines through the aid of a mouse. This procedure, however, is dependent on the user's interaction and experience in handling piezo-spectroscopic data and, during the course of our investigation, we also considered complete automation of this process. The matching was based on a function of a given point's horizontal distance, vertical distance and slope with respect to all other

previously matched data points. In general, however, this procedure was very unstable with respect to the ratio chosen between the three variables and proved inadequate as an all purpose matching procedure. The inclusion of additional explanatory variables in this function may improve this process.

References

Aries R., Lidiard D. and Spragg R. (1990), "Factor Analysis Applied to Sampling Methods for Infra-red Spectroscopy, Spectroscopy International, Vol.2, 3, 41-44.

Atkinson A.C. (1972), "Planning experiments to detect inadequate regression models", Biometrika, Vol. 59, 275-293

Atkinson A.C. (1981), "A comparison of two criteria for the design of experiments for discriminating between models", Technometrics Vol.23, 301-305

Barakati N. (1989), "Turbo C Bible", Howard W. Sams & Company.

Bates D.M. and Watts D.G. (1988), "Nonlinear regression analysis and its applications.", Wiley series in Probability and Mathematical Statistics
see also the following two articles :

Bates D.M. and Watts D.G. (1980), "Relative curvature measures of Nonlinearity", J. Royal Statistical Society B. Vol.42, 1-25

Bates D.M. and Watts D.G. (1987), "A generalised Gauss-Newton procedure for Multi-Response parameter estimation", SIAM J. SCi. and Stat. Comput. Vol.8, 49-55

Bates D.M. (1978), "Curvature Measures of Nonlinearity.", Ph.D thesis, Queens University at Kingston

Bates D.M., Hamiliton C. and Watts D.G. (1983), "Calculation of Intrinsic and parameter effects curvatures for non-linear models.", Communications in statistics- Simulation and computation, Vol.12, 469-477

Bates D.M. and Watts D.G. (1981)a), "Parameter transformation for improved approximate confidence regions in nonlinear least squares.", Annals of Statistics, Vol.9(6), 1152-1167

Bates D.M. and Watts D.G. (1981)b), "A Relative offset orthogonality convergence criterion for Non-linear Leasts squares. Technometrics, Vol.23(2), 179-183

Bates D.M. and Watts D.G. (1985), "Multiresponse Estimation With Special Application to Linear Systems of Differential Equations", Technometrics, Vol.24, 4, 329-339.

Beal E.M.L. (1960), "Confidence Regions in Nonlinear Estimation (with discussion)". Journal of Royal Statistical Society B Vol.22 41-88.

G. Beamson and D. Briggs (1991), "High Resolution Monochromated X-Ray Photoelectron Spectroscopy of Organic Polymers: A Comparison between Solid State Data for Organic Polymers and Gas Phase Data for Small Molecules, Molecular Physics, Vol. 76, No. 4, 919-936.

Beauchamp J.J. and Cornell R.G. (1966), "Simultaneous non-Linear Estimation.", Technometrics Vol.8, 319-326

BMDP (1985), "Statistical Software Manual", University of California Press.

Box M.J. and Draper N.R. (1972),"Estimation and Design Criteria for Multi-response Non-Linear Model With Non-homegeneous Variance" , App. Statist, Vol.21 , 13-14.

Box G.E.P., Hunter W.G. and Hunter J.S. (1988), "Statistics for experimenters.", New York: Wiley.

Brent R.P. (1973), "Algorithms for Minimization without Derivatives", Englewood Cliffs, N.J.: Prentice-Hall), Chapter 7.

Calderbank V.J. (1983), "A course on Programming in FORTRAN revised to incorporate FORTRAN 77", Chapman and Hall.

Campion J.D., Henry M.O., Nazare M.H. and Lightowlers E.C. (1992), "A New Photoluminescence Band in Beryllium Doped Silicon.", Proceedings of ICSIS, Kobe, Japan, to be published.

Carroll R.J. and Ruppert D. (1991), "Prediction and Tolerence Intervals With Transformation and/or Weighting", Technometrics Vol.33, No.2, 197-211.

Caspar J.V. and Meyer T.J. (1983), "Photochemistry of MLCT Excited States. Effect of Nonchromophoric Ligand Variations on Photophysical Properties in the Series *cis*-$Ru(bpy)_2L_2^{2+}$, Inorg Chemistry, Vol.22, 2444-2453.

Chesher A. (1991), "The Effect of Measurement Error", Biometrika Vol.78, No.3, 451-462.

Chojnicki T. (1978), "Estimation of Accuracy of Tidal Data Adjustment Results Based on Residual Spectrum", Inst. Geophys. Pol. Acad. Sc. F-4(129).

Croke J.F., Jenkins R. and Westberg R.G. (1978), "Comparison of Mathematical Models for Quantitative X-Ray Spectroscopy", Norelco Reporter, Vol 25, N0 1.

Dargahi-Noubary G.R. (1990), "A Method for Parameter Estimation of Non-linear Regression With Autocorrelation Errors", Commun. In Statist. Theory Meth, 19, Vol.8, 2907-2903.

Date C.J. (1985), "An Introduction to Database Systems Volume II", Addison-Wesley Publishing Company.

Davies G. and Lightowlers, E.C., "Optical Absorption by Platinum in Crystalline Silicon", Materials Science and Engineering, B4, 173-177.

Davies M.R., Shelley J.C. and LeRoy R.J., "Direct determination of long-range inverse-power potential coefficients from spectroscopic data", J.Chem.Phys. Vol.94, No.5.

Davies O.L. and Goldsmith P.L. (1984), "Statistical Methods In Research & Production", Longman.

Daly S., Unpublished data, to be presented as an Ph.D. Thesis, Dublin City University, (1994).

DEC (1982), "Digital - Vax-11 Pascal Users Guide, Version 2.0", (Digital Equipment Corporation)

DeMarco T. (1979), "Structured Analysis and System Specification", Yourdon Inc.

Diaz C.G., Fernandez F.M. and Eruardo A.C. (1991), "Construction of bound-state potential energy curves of diatomic molecules from spectroscopic data", Chemical Physics, Vol.157, 35-44

Do Carmo M.C., Calao M.I., Davies G. and Lightowlers E.C. (1988), "Photoluminescence from transition metals in silicon.", Proc.CONFS,ICDS-15, Budapeste.

Dzieciolowski K. (3-Qen), Ross W.H. (1990), "Assessing case influence on confidence intervals in non-linear regession"
(French Summary) Canad. J. Statist, Vol.18, No.2, 127-139.

Ehrenberg A.S.C. (1981), "The Problem of numeracy" American Statistican, Vol.350, Femenias J.L. and Cheval G.(1992), "Analysis of Residuals in fitting Spectroscopic Data", Can. J. of Phys., Vol.70, 152 - 163

Feynman R.P., Leighton R.B. and Sands M. (1966), "The Feynman lectures on Physics (Quantum mechanics)", (1966), Addison - Wesley.

Fomby T.B., Carter-Hill R. and Johnson S.R. (1988), "Advanced Econometric Methods", Springer-Verlag.

Fowler W. (1968), "Physics of colour centres", Academic Press.

Gelsema E.S., Beckers A.L.D, Sorber C.W.J. and De Bruijn W.C. (1992), "Correspondence Analysis for Quantification in Electron Energy Loss Spectroscopy and Imaging", Meth Inform Med, Vol.31, 29-35.

Gillis P.R. and Ratkowsky D.A. (1978), "The behaviour of estimators of parameters of various yield density relationships", Biometrics Vol.34, 191-8

Goldberg M.L., Bates D.M. and Watts D.G. (1983), "Simplified method of assessing Nonlinearity", American Statistcal Society, Proc. Bus Econ. Stats. Section 67-74.

Gordon S.H., Green R.V., Wheeler B.C. (1993), "Multivariate FTIR Analysis of Substrates for Protein, Polysaccharide, Lipid and Microbe Content: Potential For solid-State Fermentations", Biotechnology Advances, Vol.11, No.3.

Graftool (1990), "User's Guide", 3-D Visions Corporation.

Griffiths, J.S. (1961), "The Theory of Transition Metal Ions.", Cambridge University Press.

Hamilton, D.C. (1986), "Confidence regions for parameter subsets in non-linear regression", Biometrika Vol.73, 57-64.

Hatley D.J. and Pirbnai I.A. (1988), "Strategies for Real-Time System Specification", Dorset House Publishing.

Heine V. (1960) - Group Theory in Quantum Mechanics, Pergamon press New York.

Herbert, K.L. (1981), "An Evaluation of Mathematical Software that solves nonlinear Least Squares Problems", ACM Trans Math.

Honerkamp J., Maier D. and Weese J. (1993), "A Nonlinear Regularization Method For The Analysis of Photon Correlation Spectroscopy Data, J. Chem. Phys., Vol.98, No.2.

Huang M.N.L. and Huang M.K. (1991), "A Parameter-Elimination Method for Nonlinear Regression with Linear Parameters and Autocorrelated Errors", Biom. Journal, Vol.33, No.8, 937-950.

Hughes A.E. and Runciman W.A. (1967), "Uniaxial stress splitting of doubly degenerate states of tetragonal and trigonal centres in cubic crystals.", Proc Phys. Soc., Vol.90, 2279

Jellison(Jr.) G.E. (1991), "Use of the Biased Estimator In The Interpretation of Spectroscopic Ellipsometry Data", Applied Optics, Vol.30, No.23.

Johnston J. (1984), "Econometric methods", McGraw Hill International student edition.

Jones M.P. (1980), "The Pratical Guide to Structured Systems Design", Yourdon Press.

Joshi A.W. (1982), Elements of Group Theory For Phyicists, (1982), Wiley Eastern Limited.

Judge G.G, Griffiths W.E, Carter-Hill R., Lutkepohl H. and Tsoung-Chao

Lee,(1985), "The Theory and Practice of Econometrics.", Wiley and Sons.

Kang G. and Bates D.M. (1990), "Aproximate Inferences In Multiresponse Regression Analysis", Vol.77, No.2, 321-331.

Kaplyanskii A.A. (1964)a), "Noncubic centres in cubic crystals and their piezospectroscopic investigation.",OPT-SPECTROSC, Vol 16, 329

Kaplyanskii A.A. (1964)b), "Noncubic centres in cubic crystals and their piezospectroscopic investigation", OPT-SPECTROSC, Vol 16, 557.

Kauppinen J.K., Moffatt, D.J., Mantsch, H.H., and Cameron, D.G. (1981) Appl. Spec. 35, 271-276.

Kehoe T., Unpublished data, presented as an M.Sc. Thesis, Dublin City University, (1993)

Kernighan B.W. and Ritchie D.M. (1988), "The C programming Language", Prentice Hall Software Series.

Kobayashi M. (1991), "Testing For Autocorrelated Disturbances in Nonlinear Regression Analysis", Econometrica, Vol.59, No.4, 1153-1159.

Kober E.M. and Meyer T.J. (1983), Inorg Chem, Vol.22, 1614-6

Knuth D.E. (1973), "The Art of Computer Programming", Volume 3: Sorting and Searching", Addison- Wesley Publishing Co.

Kroonenberg P.M. (1992), Comment: "PARAFAC in Three-Way Land", Statisical Science, Vol.7 No.3, 312-319.

Leech D. (1975), "Testing the error specification in nonlinear regression", Econometrica, Vol.43, 719-725

Leurgans S. and Ross R.T. (1992), "Multilinear Models: Applications in Spectrscopy", Statisical Science, Vol 7 No. 3, 289-311.

Linssen H.N. (1975), "Non-Linearity measures a case study." Statistica Neerlandica, Vol.29, 93-99

Ludlow I.K. (1968), " Spin-orbit coupling and stress spectra of trigonal colour centres in magnesium oxide", J. Phys. C, Vol.1, 1194

MacDonald J.R. and Potter(Jr.) L.D. (1987), "A Flexible Procedure For Anlyzing Impedance Spectroscopy Results: Descripition and Illustrations", Solid State ionics, Vol.23, 61-79.

Macsyma (1988), "Users Guide", Symbolics.

Macsyma (1988), "Reference Manual Version 13",Symbolics.

Malcata F.X., Hill C.G. and Amundson C.H. (1993), "Hydrolysis of Butteroil By Immobilized Lipase using a Hollow-Fiber Reactor: Part VI. Multiresponse Analysis of Temperature and pH Effects.", Biocatalysis, Vol.8, 201-228.

Malinowski E.R. and Howery D.G. (1980), "Factor Analysis in Chemistry", Wiley, New York.

THE MATHWORKS INC. (1987), "PC-MATLAB Users Guide",Wiley.

McCarren A.L. and Ruskin H.J (1991), "The study of Methods used in Determining the Transition Type of Luminescene Data.", Technical Report, Ca-0891, School of Computer Applications, Dublin City University.

McCarren A.L., Ruskin H.J., McGuigan K.G. and Henry M.O. (1994), "Piezo-Spectroscopic Data Analysis: A PC Tool", IEE Proc-Sci Meas, Technol., Vol.141, No.3, 185-189.

McCaslin P.C. (1991), "Development of an Offline X-ray Photoelectron Spectroscopic Personal Computer-Based Data Analysis System.", J. Vac. Sci Technol., Vol.9, No.3.

McGuigan K. PhD Thesis D.C.U. (1989) ,"Luminescence studies of transition metal related defects in crystalline silicon."

McGuigan K., "Private Communication related to Indium Doped Silican Data".

Mendenhall W. and Reinmuth J.E. (1982), "Statistics for Management and Economics", Fourth Edition, (1982), Duxbury Press.

Mohammed K., Davis G. and Collins A.T. (1982), "Uniaxial stress splitting of photoluminescence transitions at optical centres in cubic crystals: theory and application to diamond.",J. Phys. C, Vol 15, 2779

Myers H.P. (1990), Introductory Solid State Physics, Taylor and Francis

NAG - (1988), "NAG Fortran Library Manual : Mark 13", Oxford, Unitied Kingdom; Downer's Gorve Ill : Numerical Algorithms Group.

Neter J. and Wasserman W. (1974), "Applied Linear Statistical Models",Irwin Inc.

Neter J., Wasserman W. and Kutner M.H. (1990), "Applied Linear Statistical Models", (Third Edition), Irwin.

Nicol R., Smith P. and Raggartt P. (1985), "A Robust Microcomputer Routine for the Identification of Outlying and Influential Points in Radioimmunoasssy Standard Curves", Computer and Biomedical Research, Vol.18, 334-346.

O'Brien S.K. (1989), "Turbo Pascal 5.5 The Complete Reference",

Ouyang X. and Varghese P.L. (1989), "Reliable and Efficient Program for Fitting Galatry and Voigt Profiles to spectral data on Multiple Lines, "Applied Optics", Vol.28, No.8.

Parker S., Unpublished document, "Menus.doc", (1989).

Piepponen S. and Lindstrom R. (1989), "Data Analysis of Heavy Metal Pollution in the Sea by Using Pricipal Component Analysis and Partial Least Squares Regression", Chemometrics and Intelligent Laboratory Systems, Vol.7, 163-170.

Pooi A.H. (1989), "Region estiamtes based on Likelihood ratio in a non linear model", Bull. Malaysian Math Soc.(2) Vol.12, No.2, 72-82

Press W.H., Flannery B.P., Teulblsky S.A. and Vetterling W.T. (1985), "Numerical Recipes", Cambridge University Press.

Press W.H., Flannery B.P., Teulblsky S.A. and Vetterling W.T. (1985), "Numerical Recipes Example Book [Pascal].", Cambridge University Press.

Press W.H., Flannery B.P, Teulblsky S.A. and Vetterling W.T. (1990), "Numerical Recipes Example, Book [C].", Cambridge University Press.

Press W.H., Flannery B.P., Teulblsky S.A. and Vetterling W.T. (1990), "The Art of Scientific Computing", Cambridge University Press.

Ratkowsky D.A. and Dolby G.R. (1975) "Taylor series linearization and scoring for parameters in nonlinear regression. Appl Statist. Vol.24, 109,111

Rudden M.N. and Wilson J. (1980), "Elements of Solid State Physics", Wiley.

Rusling J.F. and Kumosinki T.F. (1992), "New Advances in computer Modeling of Chemical and Biochemical Data", Elsevier Science Publishing Co. INC.

SAS (1988), "Sas/Graph user's guide, release 6.03 edition", SAS Institute Inc. Cary, NC, USA.

SAS (1988), "Sas/Stat user's guide, release 6.03 edition", SAS Institute Inc. Cary, NC, USA.

Seber G.A.F. and Wild C.J. (1989),"Nonlinear Regression", Wiley.

Sedgewick R. (1983), "Algorithms", Addison - Wesley Publishing Co.

Seraydarian R.P., Burrell K.H. and Groebner R.J. (1992), "A global fitting code for multichordal neutral beam spectroscopic data", General Atomics, San Diego, California, 92186-9784

Shenk J.S. and Westerhuas M.O. (1991), "Population Structuring of Near Infrared Spectra and Modified Partial Least Squares Regression", Crop Sci, Vol.31, 1548-1555.

Shmueli U. and Weiss G.H. (1990), "Probabilistic Methods in Crystal Structure Analysis", Journal of American Statistical Association, Vol.85, 409.

Sormann H., Kindl P. and Werner P. (1983), "Investigations on the Reliablity of a Multi-Component Anlaysis of Positron Lifetime Spectra, using a New Method of Producing Computer-Simulated Test Spectra.", Nuclear Instruments and Methods, Vol.206, 203-209.

Stewart W.E. and Sorensen, J.P. (1981), "Bayesian estimation of common parameters from multiresponse data with missing observations.", Technometrics, Vol.23, 131-141.

SURFER (1989), "Surfer Reference Manual ,Version 4.", Golden Software Incorporated.

Swokowski E.W. (1984), "Calculas with analytic geometry" , Prindle, Weber & Schmidt.

Taha H.A. (1987), "Operations Research An Introduction", Macmillan Publishing Company.

Tinkham M. (1964), "Group theory and Quantum Mechanics", New York, McGraw-Hill.

Vincient A. (1977), "Molecular symmetry and Group theory - A programmed Introduction to Chemical Apllications", Wiley J.

Walsh G.R. (1975), "Methods of optimization", Wiley J.

Wilks S.S. (1962), "Mathematical Statistics", New York, Wiley J.

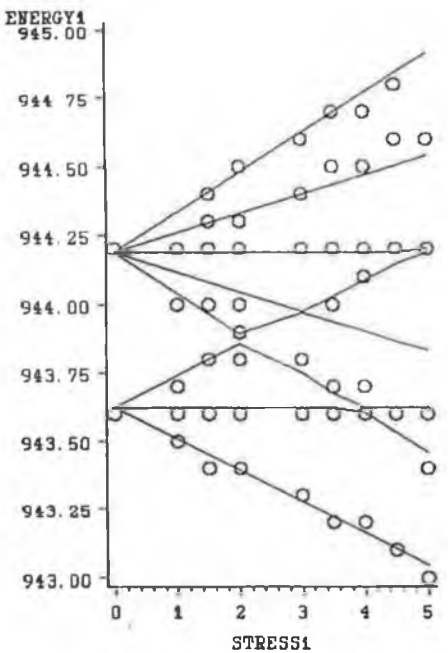
Windig W., Heckler C.E., Agblevor F.A. and Evans R.J. (1991), "Self-Modeling Mixture Analysis of Categorized Pyrolysis Mass Spectral Data with SIMPLISMA Approach", Chemometrics and Intelligent Laboratory Systems, Vol.14, 195-207.

Windig W., Lippert J.L., Robbins M.J., Kresinske K.R and Snyder A.P. (1990), "Interactive Self-Modeling Multivariate Analysis", Chemometrics and Intelligent Laboratory Systems, Vol.9, 7-30

Appendix A

The fits obtained using the Powell-shell for the problems outlined in Mcguigan(1989), Figures (A.1) to (A.3) and Daly(unpublished) Figures (A.4) to (A.6) are given below.

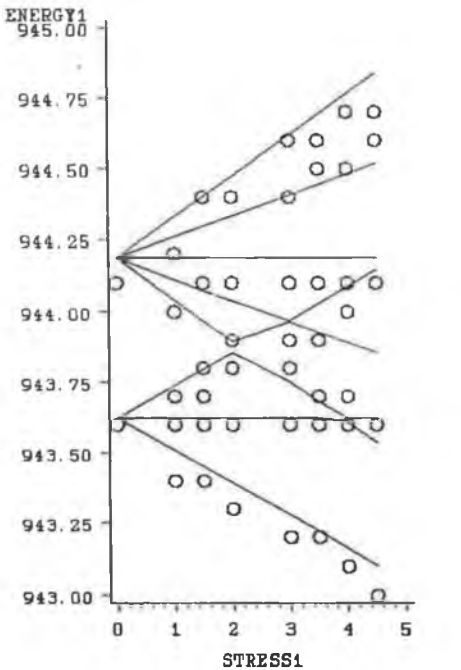
001 Direction



Circles denote actual points
Joint Lines denote estimated points

Figure A.1

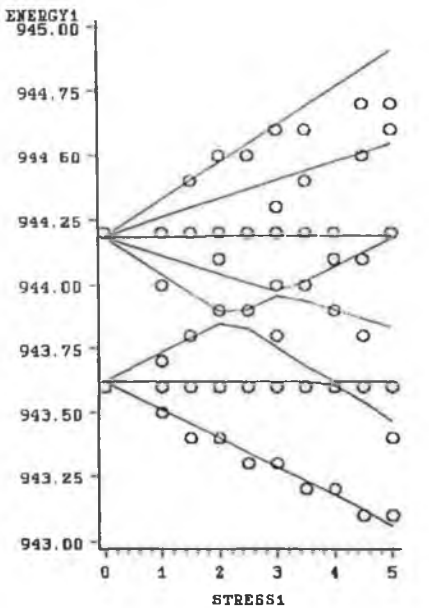
111 Direction



Circles denote actual data points
Joint Lines denote estimated data points

Figure A.2

110 Direction



Joint Lines denote the estimated points

Figure A.3

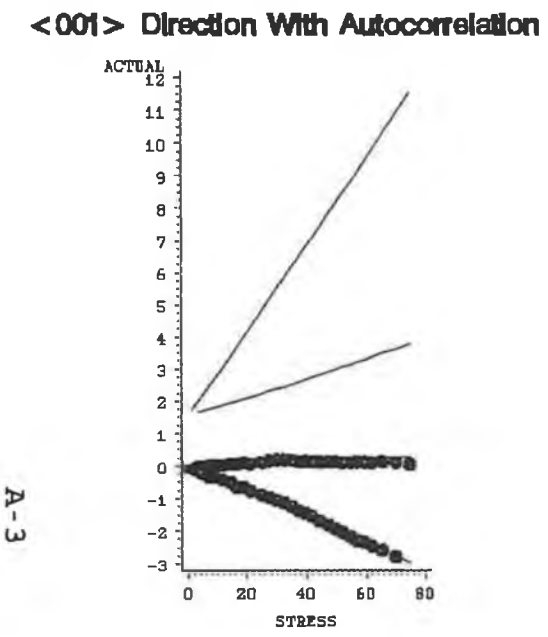


Figure A.4

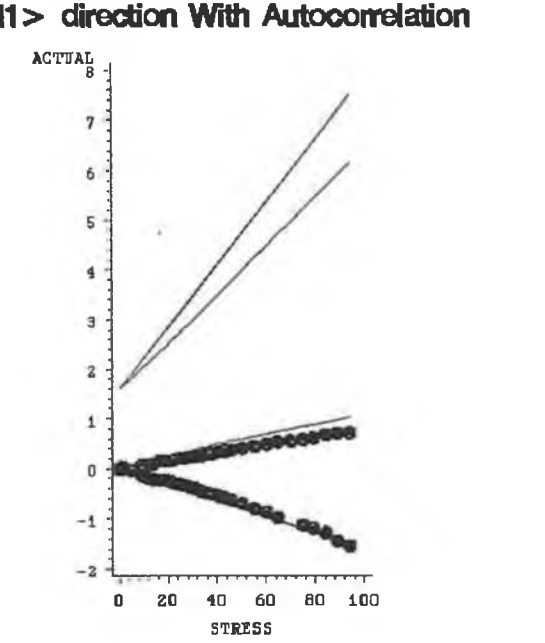


Figure A.5

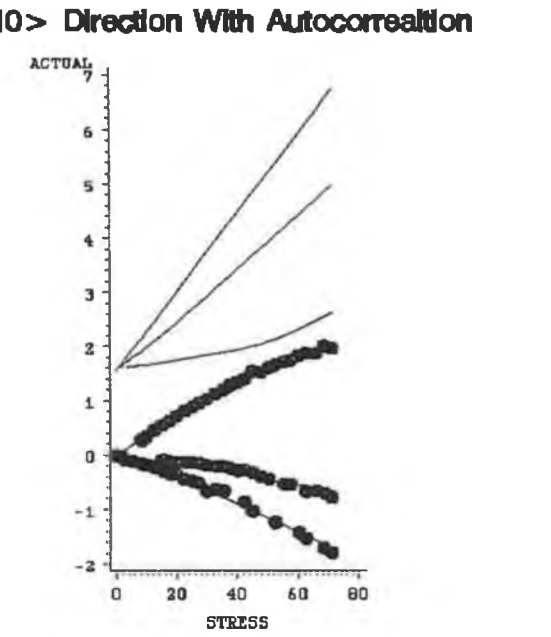
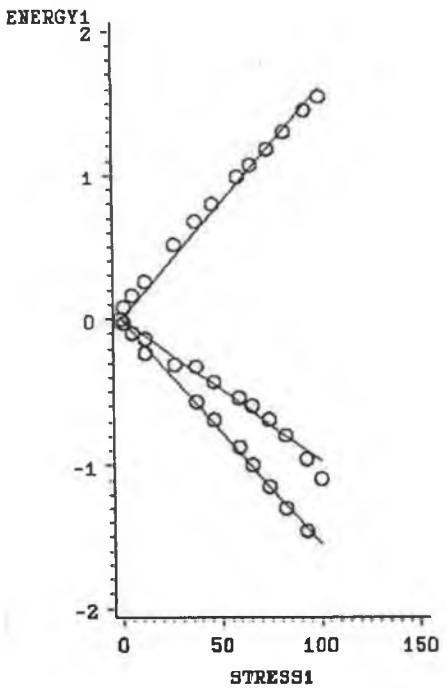


Figure A.6

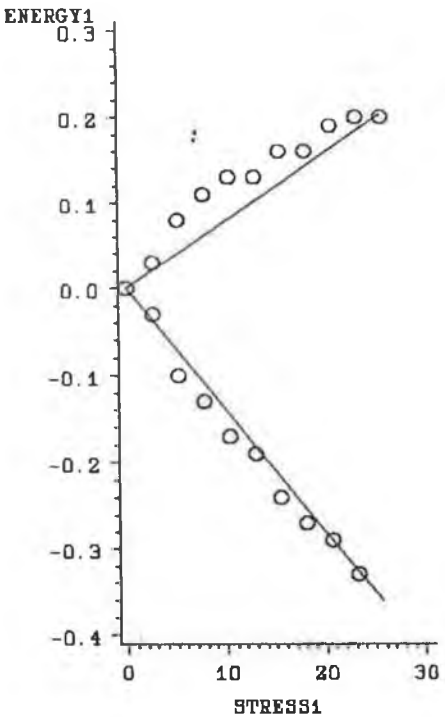
Campion et al.(1992) fits with no autocorrelation terms included.

<100> Direction



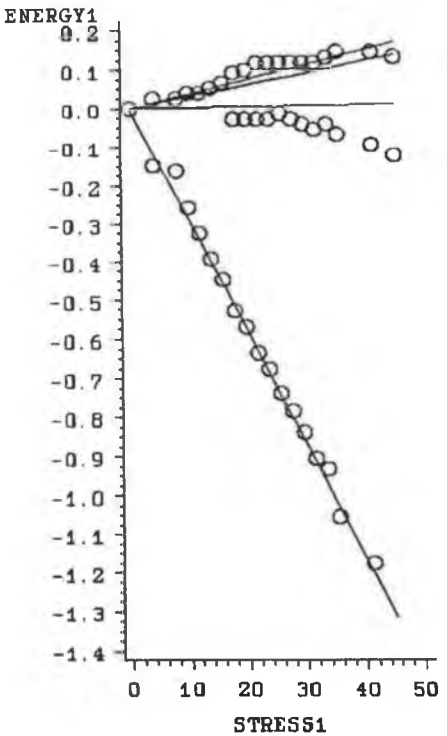
Source: Source Kevin McGuigans data
For Stress in <001> direction

<111> Direction



Source: Source Kevin McGuigans data
For Stress in <111> direction

<110> Direction



Source: Source Dr do Carmo
For Stress in <110> direction

Appendix B

The correlation matrices for the errors of the multi-responses in each of the three directions for the do Carmo et al(1988) example. Values marked ****** indicate that there were insufficient matching data points for the two vectors to obtain a correlation.

Correlation for Direction <001>

1.0000	1.0000	-0.4737	-0.4737	******	******	-0.6755	-0.6755
1.0000	1.0000	-0.4737	-0.4737	******	******	-0.6755	-0.6755
-0.4737	-0.4737	1.0000	1.0000	******	******	0.2508	0.2508
-0.4737	-0.4737	1.0000	1.0000	******	******	0.2508	0.2508
******	******	******	******	1.0000	1.0000	******	******
******	******	******	******	1.0000	1.0000	******	******
-0.6755	-0.6755	0.2508	0.2508	******	******	1.0000	1.0000
-0.6755	-0.6755	0.2508	0.2508	******	******	1.0000	1.0000

Correlation for Direction <111>

1.0000	1.0000	1.0000	1.0000	0.2640	0.2640	0.4245	0.4245	******	******	******	******	******	******	******	******
1.0000	1.0000	1.0000	1.0000	0.2640	0.2640	0.4245	0.4245	******	******	******	******	******	******	******	******
1.0000	1.0000	1.0000	1.0000	0.2640	0.2640	0.4245	0.4245	******	******	******	******	******	******	******	******
1.0000	1.0000	1.0000	1.0000	0.2640	0.2640	0.4245	0.4245	******	******	******	******	******	******	******	******
0.2640	0.2640	0.2640	0.2640	1.0000	1.0000	0.9418	0.9418	******	******	******	******	******	******	0.5142	0.5142
0.2640	0.2640	0.2640	0.2640	1.0000	1.0000	0.9418	0.9418	******	******	******	******	******	******	0.5142	0.5142
0.4245	0.4245	0.4245	0.4245	0.9418	0.9418	1.0000	1.0000	******	******	******	******	******	******	******	******
0.4245	0.4245	0.4245	0.4245	0.9418	0.9418	1.0000	1.0000	******	******	******	******	******	******	******	******
******	******	******	******	******	******	******	******	1.0000	1.0000	******	******	0.3223	0.3223	******	******
******	******	******	******	******	******	******	******	1.0000	1.0000	******	******	0.3223	0.3223	******	******
******	******	******	******	******	******	******	******	******	******	1.0000	1.0000	******	******	-0.3148	-0.3148
******	******	******	******	******	******	******	******	******	******	1.0000	1.0000	******	******	-0.3148	-0.3148
******	******	******	******	******	******	******	******	0.3223	0.3223	******	******	1.0000	1.0000	******	******
******	******	******	******	******	******	******	******	0.3223	0.3223	******	******	1.0000	1.0000	******	******
******	******	******	******	0.5142	0.5142	******	******	******	******	-0.3148	-0.3148	******	******	1.0000	1.0000
******	******	******	******	0.5142	0.5142	******	******	******	******	-0.3148	-0.3148	******	******	1.0000	1.0000

Correlation for Direction <110>

1.0000	1.000	0.9392	0.9392	******	******	******	******	******	******	******	******	******	******	******	******
1.000	1.0000	0.9392	0.9392	******	******	******	******	******	******	******	******	******	******	******	******
0.9392	0.9392	1.0000	1.0000	******	******	0.1450	0.1450	0.4554	0.4554	******	******	******	******	******	******
0.9392	0.9392	1.0000	1.0000	******	******	0.1450	0.1450	0.4554	0.4554	******	******	******	******	******	******
******	******	******	******	1.0000	1.0000	******	******	0.2837	0.2837	******	******	******	******	******	******
******	******	******	******	1.0000	1.0000	******	******	0.2837	0.2837	******	******	******	******	******	******
******	******	0.1450	0.1450	******	******	1.0000	1.000	-0.7172	-0.7172	-0.9506	-0.9506	******	******	0.9698	0.9698
******	******	0.1450	0.1450	******	******	1.000	1.0000	-0.7172	-0.7172	-0.9506	-0.9506	******	******	0.9697	0.9697
******	******	0.4554	0.4554	0.2837	0.2837	-0.7172	-0.7172	1.0000	1.0000	0.1185	0.1185	******	******	-0.3854	-0.3854
******	******	0.4554	0.4554	0.2837	0.2837	-0.7172	-0.7172	1.0000	1.0000	0.1185	0.1185	******	******	-0.3854	-0.3854
******	******	******	******	******	******	-0.9506	-0.9506	0.1185	0.1185	1.0000	1.0000	-0.0628	-0.0628	-0.5061	-0.5061
******	******	******	******	******	******	-0.9506	-0.9506	0.1185	0.1185	1.0000	1.0000	-0.0628	-0.0628	-0.5061	-0.5061
******	******	******	******	******	******	******	******	******	******	-0.0628	-0.0628	1.0000	1.0000	0.0190	0.0190
******	******	******	******	******	******	******	******	******	******	-0.0628	-0.0628	1.0000	1.0000	0.0190	0.0189
******	******	******	******	******	******	0.9698	0.9697	-0.3854	-0.3854	-0.5061	-0.5061	0.0190	0.0190	1.0000	1.0000
******	******	******	******	******	******	0.9698	0.9697	-0.3854	-0.3854	-0.5061	-0.5061	0.0190	0.0189	1.0000	1.0000

Appendix C

Plots of error versus stress with the autocorrelation terms excluded are shown in Figures (C.1) - (C.19) for do Carmo et al.(1988), in Figures (C.20)-(C.24) for Kehoe(unpublished) and in Figures (C.25)-(C.31) for Daly(unpublished).

Error versus Stress plots for each response in the <001> direction for the do Carmo et al.(1988) example with the autocorrelation terms excluded.

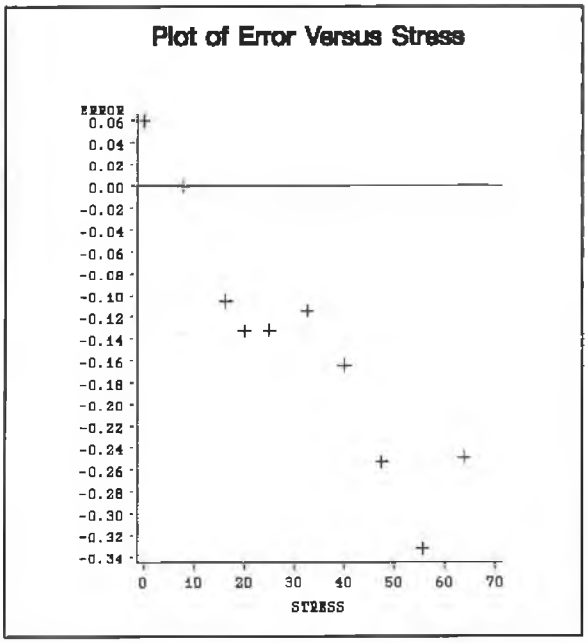


Figure C.1

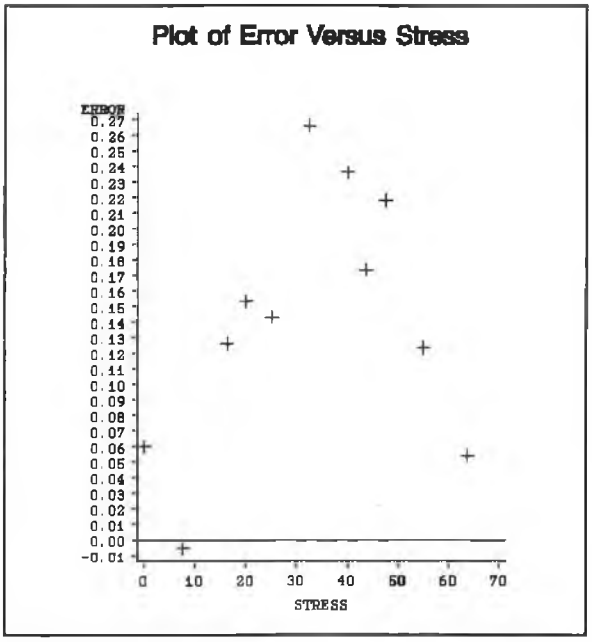


Figure C.2

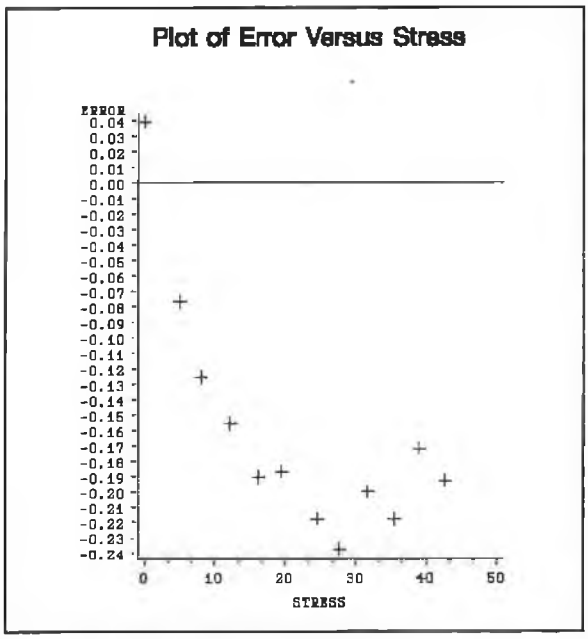


Figure C.3

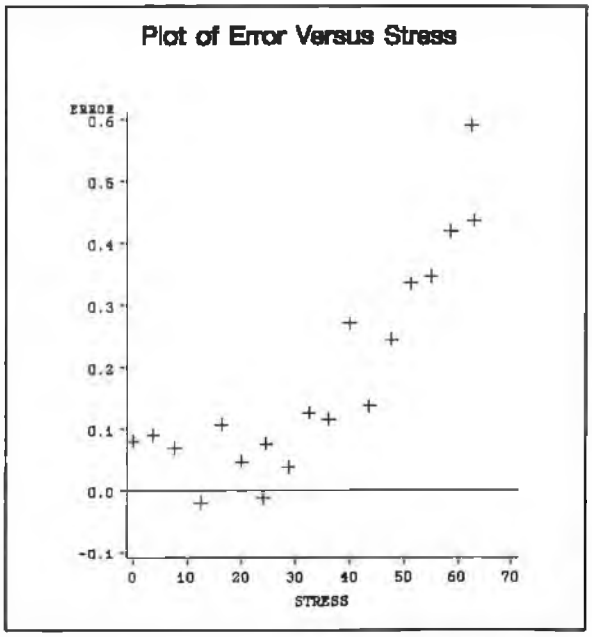


Figure C.4

Figures C.5-C11 give the Error versus Stress plots in the <111> direction, with the Autocorrelation terms excluded, for the do Carmo et al.(1988) example.

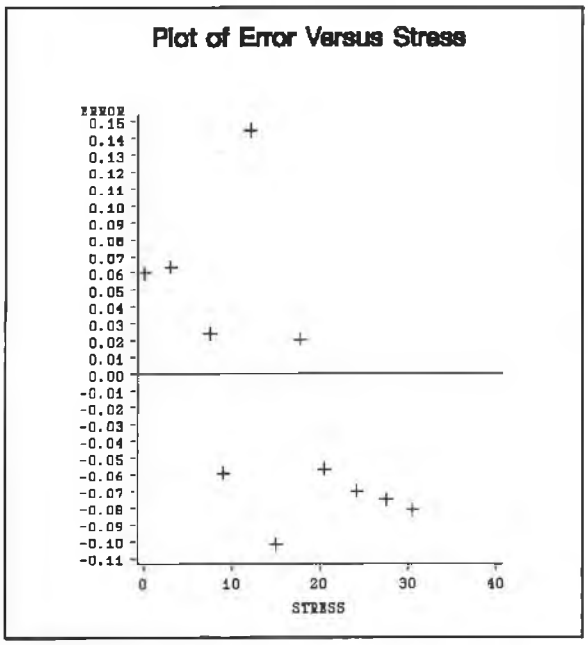


Figure C.5

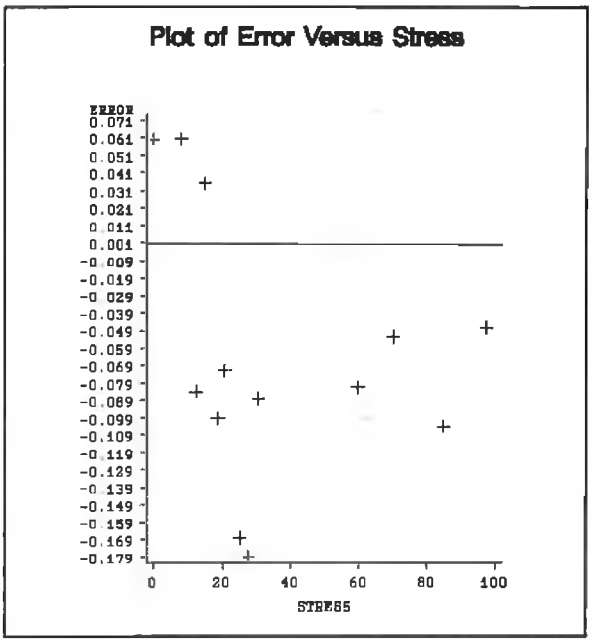


Figure C.6

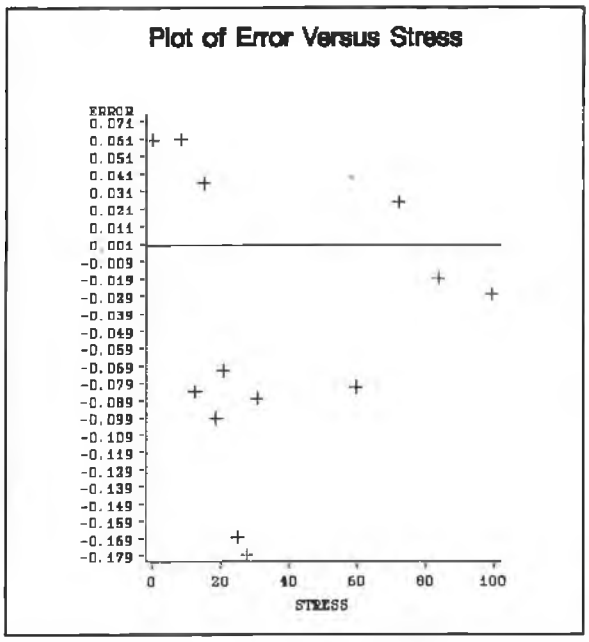


Figure C.7

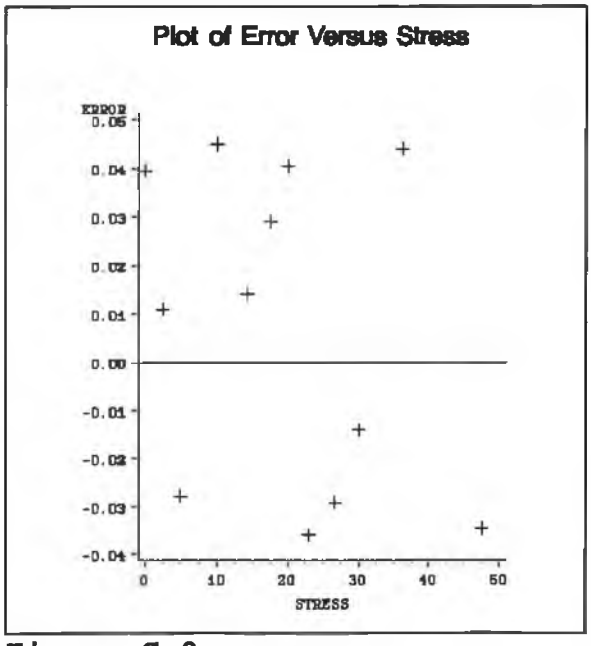


Figure C.8

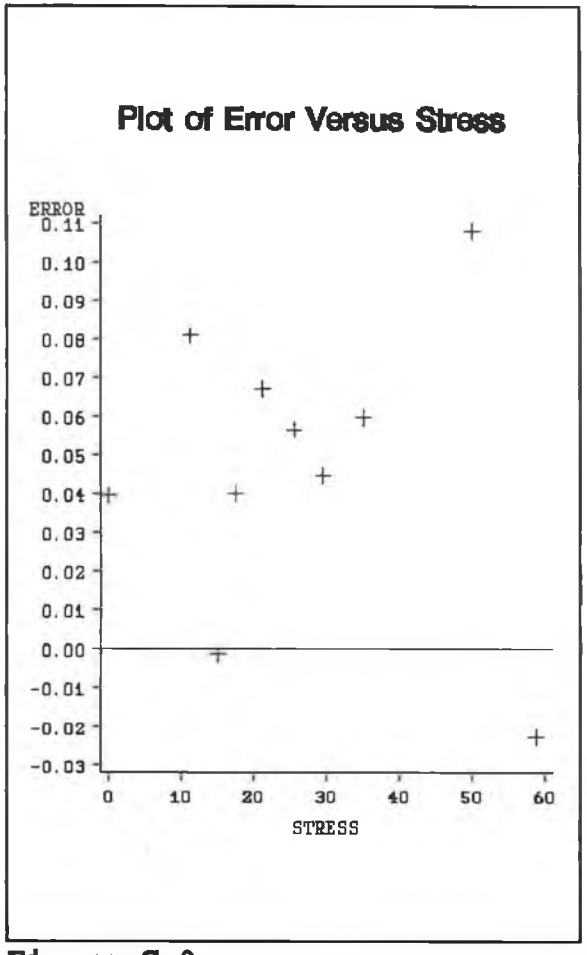


Figure C.9

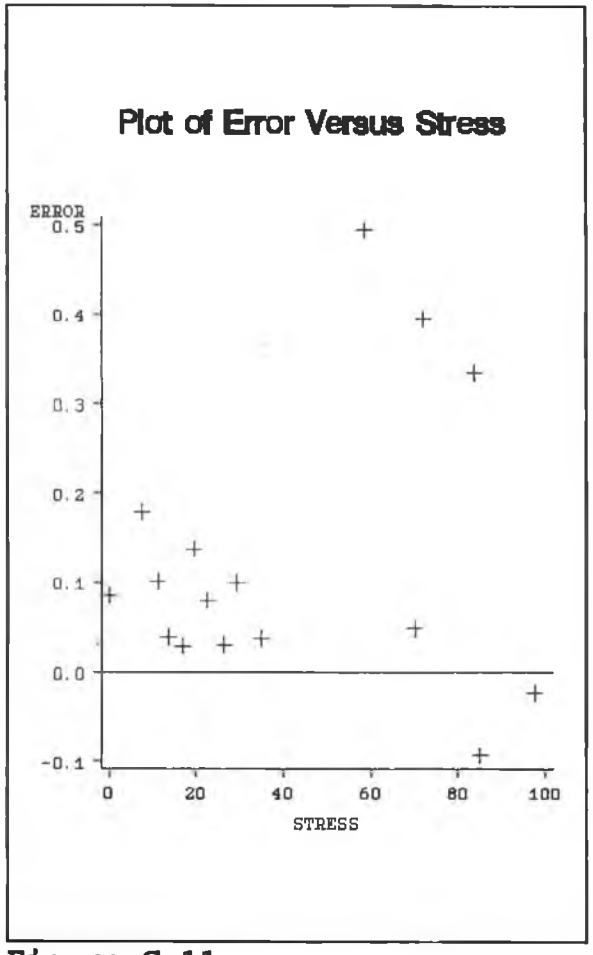


Figure C.11

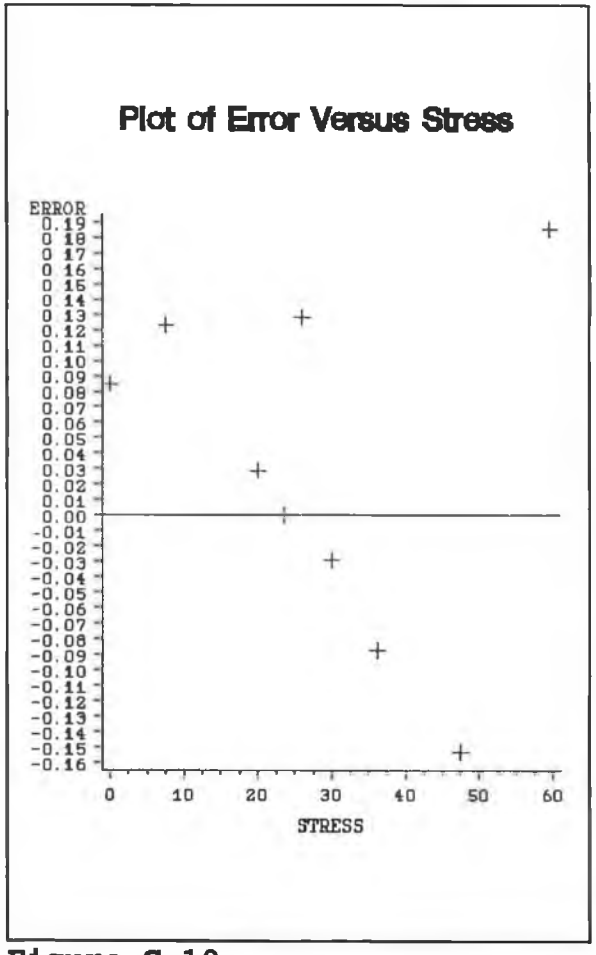


Figure C.10

Figure C.12 to C.19 give the Error versus Stress plots for the individual responses in the <110> direction for do Carmo et al.(1988) example, with the autocorrelation terms excluded.

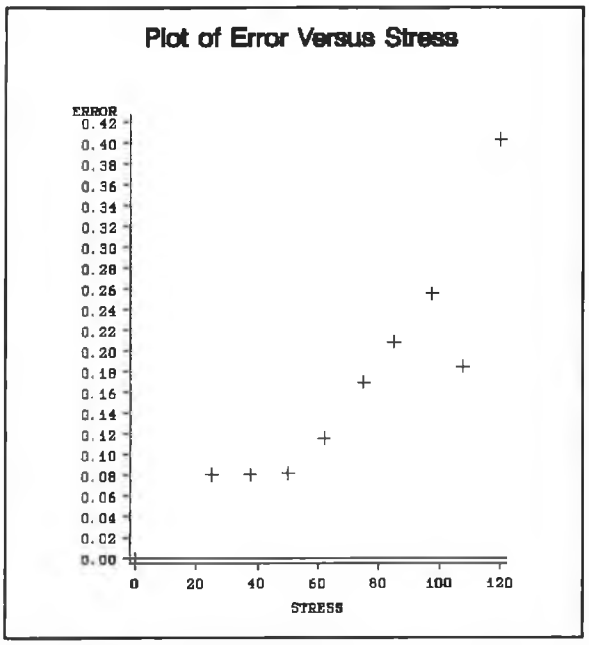


Figure C.12

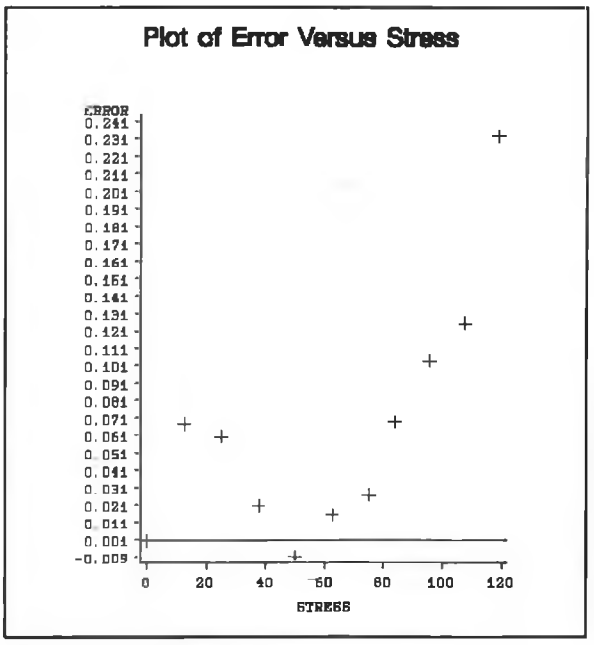


Figure C.13

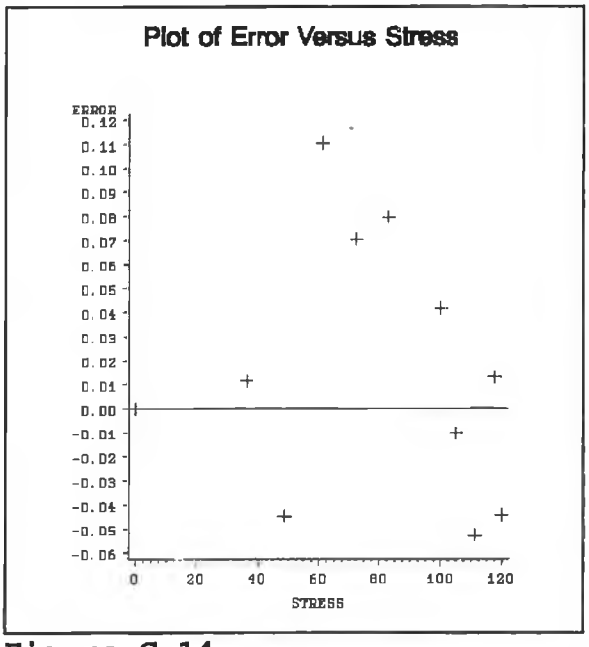


Figure C.14

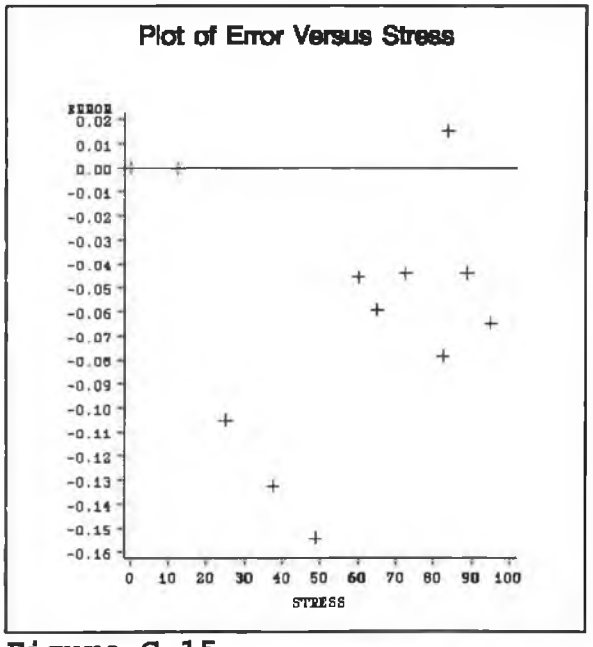


Figure C.15

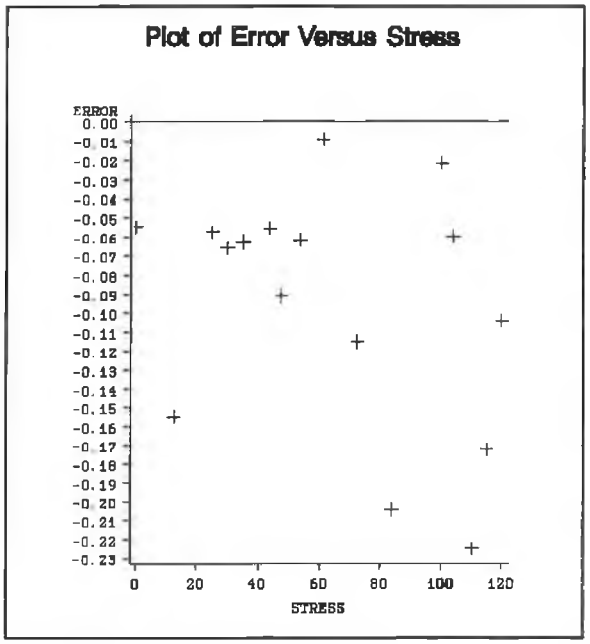


Figure C.16

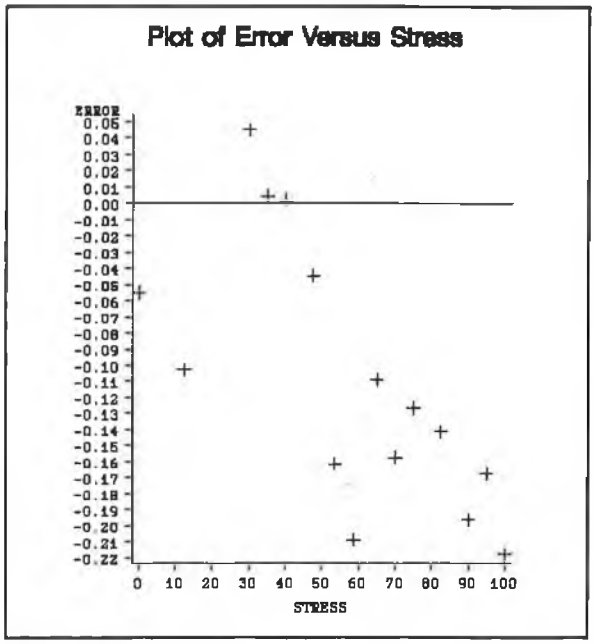


Figure C.17

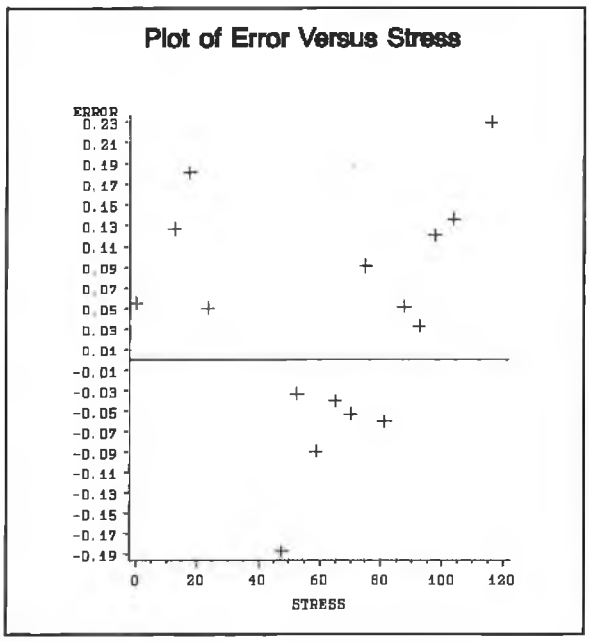


Figure C.17

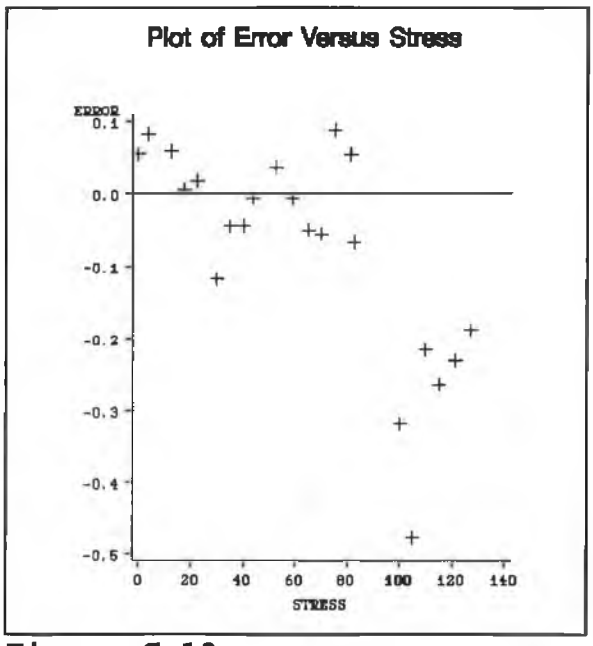


Figure C.19

Error versus Stress Plots for the <111> and <001> directions for Kehoe(unpublished), with the auto-correlation terms excluded.

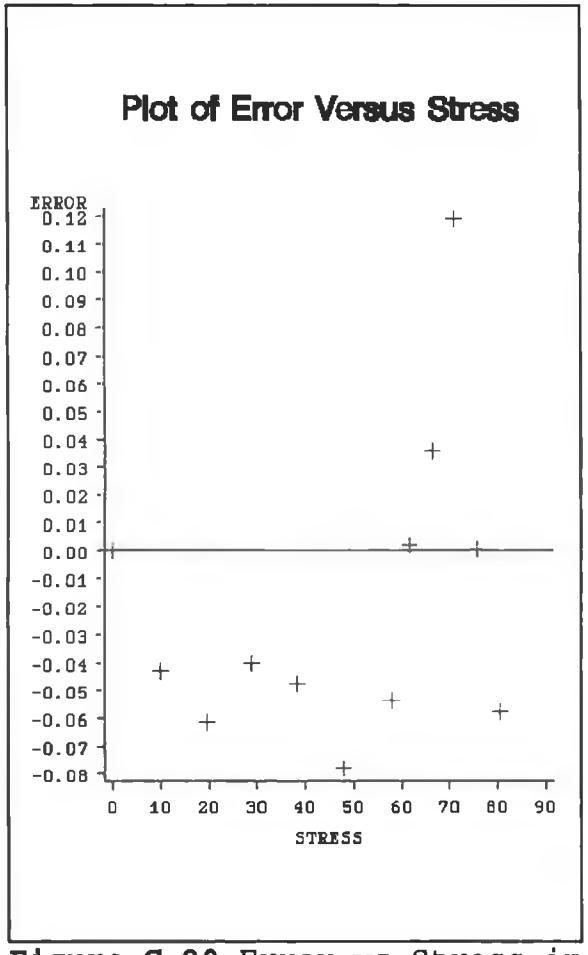


Figure C.20 Error vs Stress in <111>

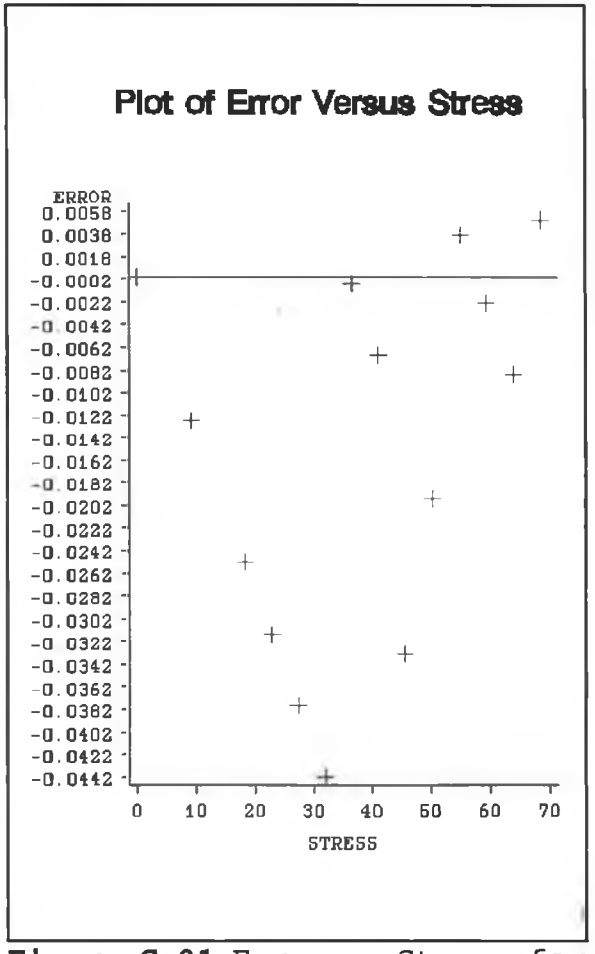


Figure C.21 Error vs Stress for <001> direction

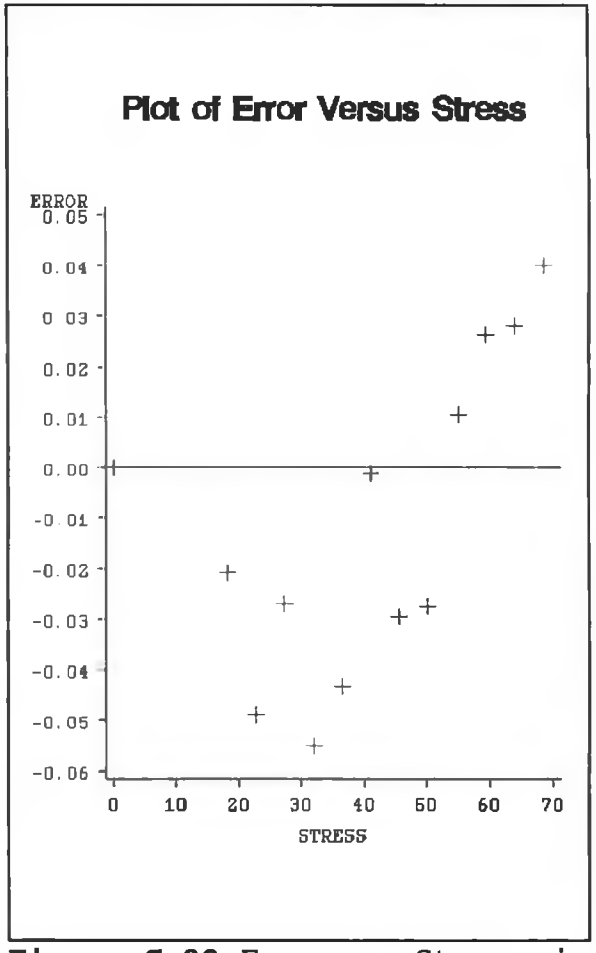


Figure C.22 Error vs Stress in <001>.

Error versus stress for the <110> direction in Kehoe(unpublished), with the autocorrelation terms excluded.

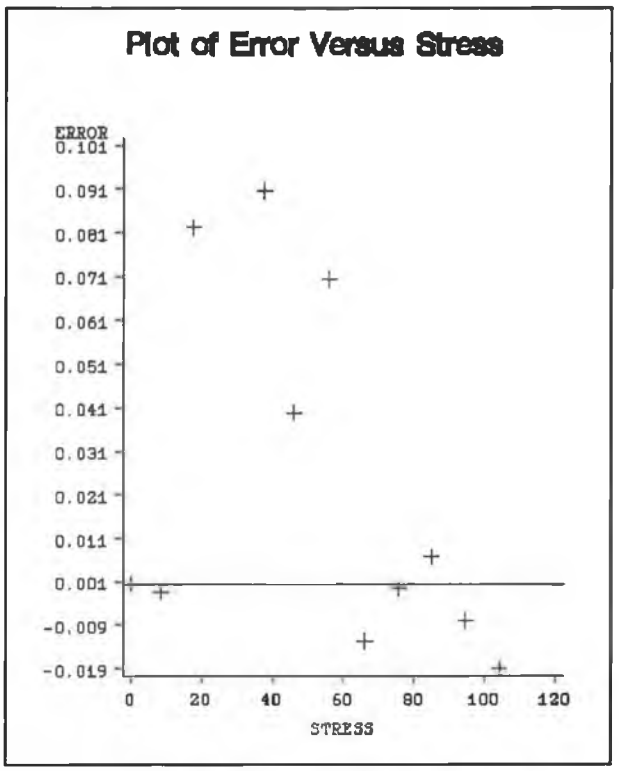


Figure C.23

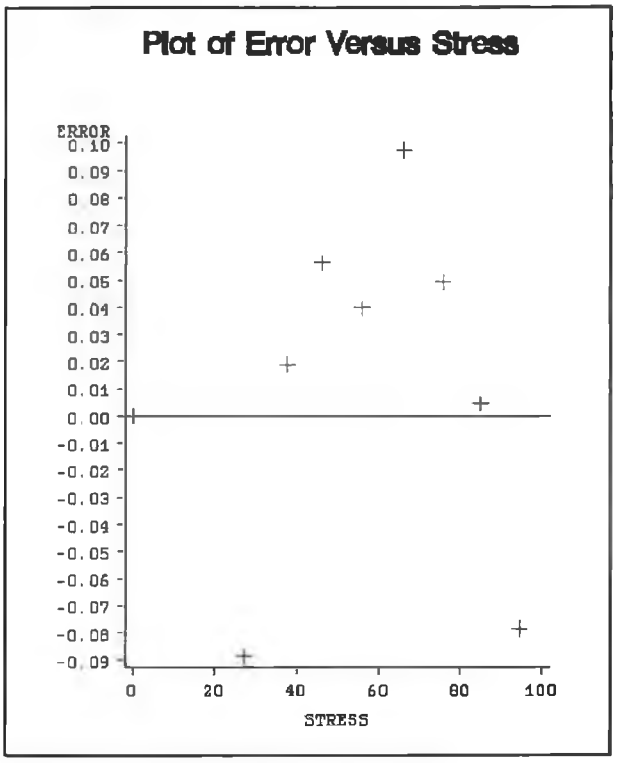


Figure C.24

Error versus Stress Plots for the individual responses in the <001> and <111> directions for the Daly(unpublished) example with the autocorrelation terms excluded.

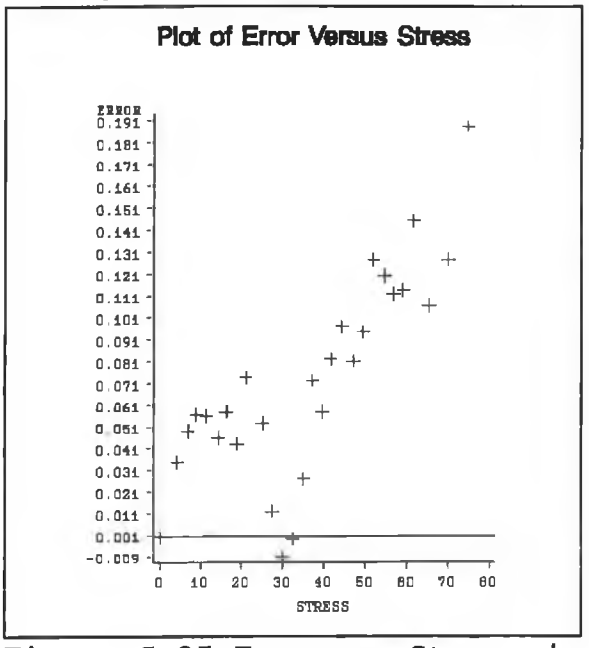


Figure C.25 Error vs Stress in <001>.

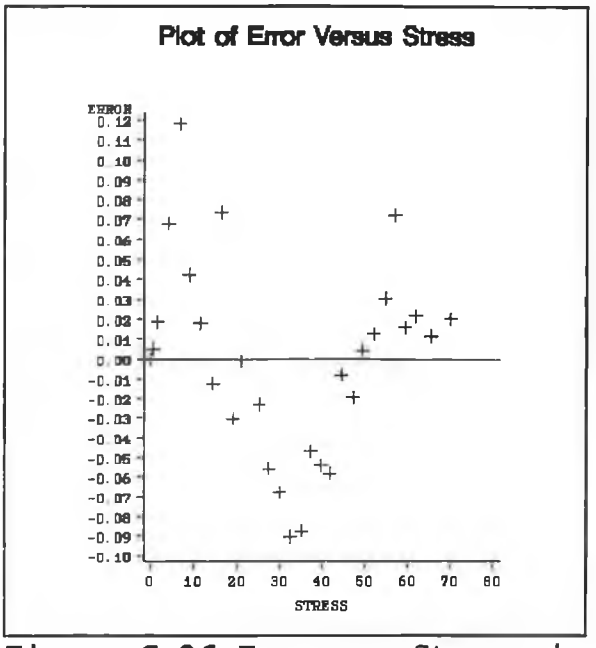


Figure C.26 Error vs Stress in <001>.

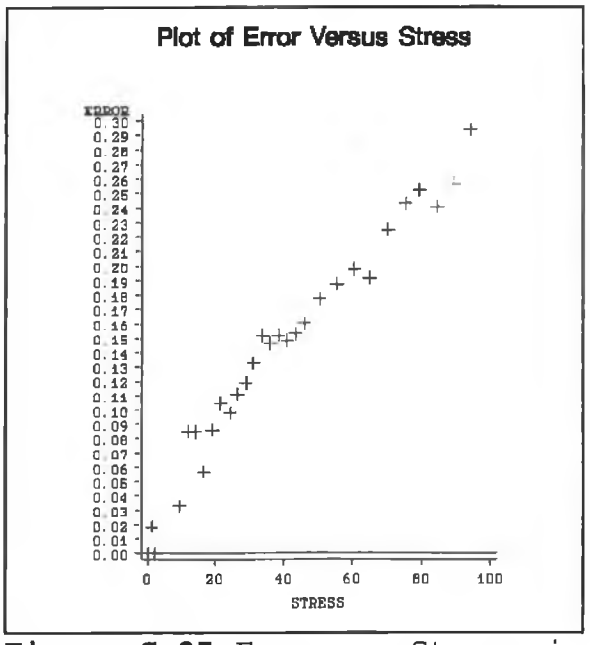


Figure C.27 Error vs Stress in <111>

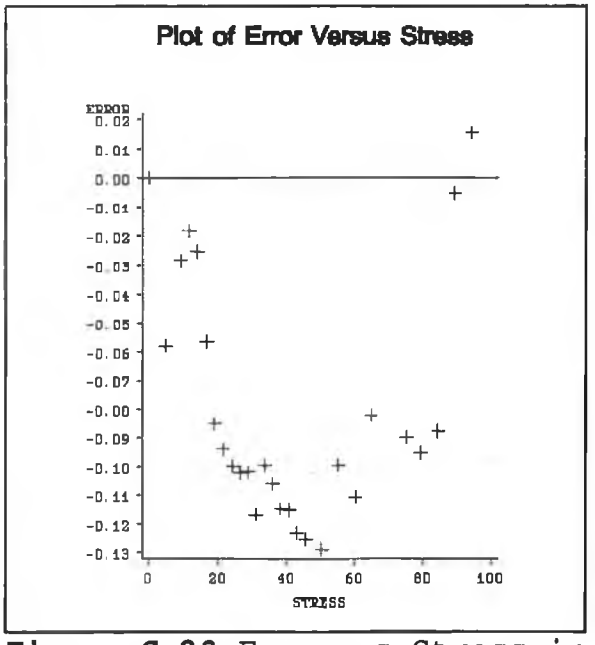


Figure C.28 Error vs Stress in <111>.

Error versus Stress plots for the individual responses in the <110> direction for the Daly(unpublished) example with the autocorrelation terms excluded.

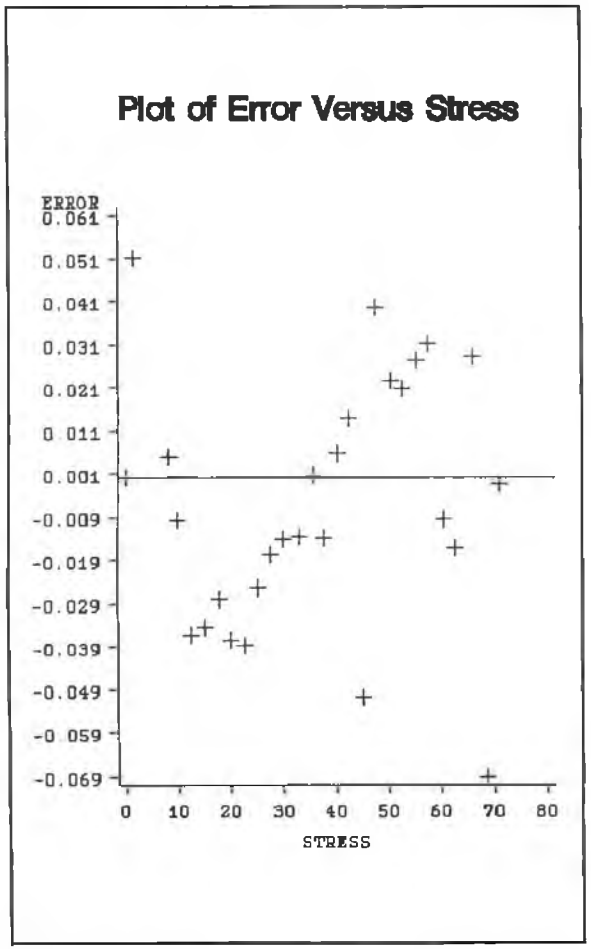


Figure C.29

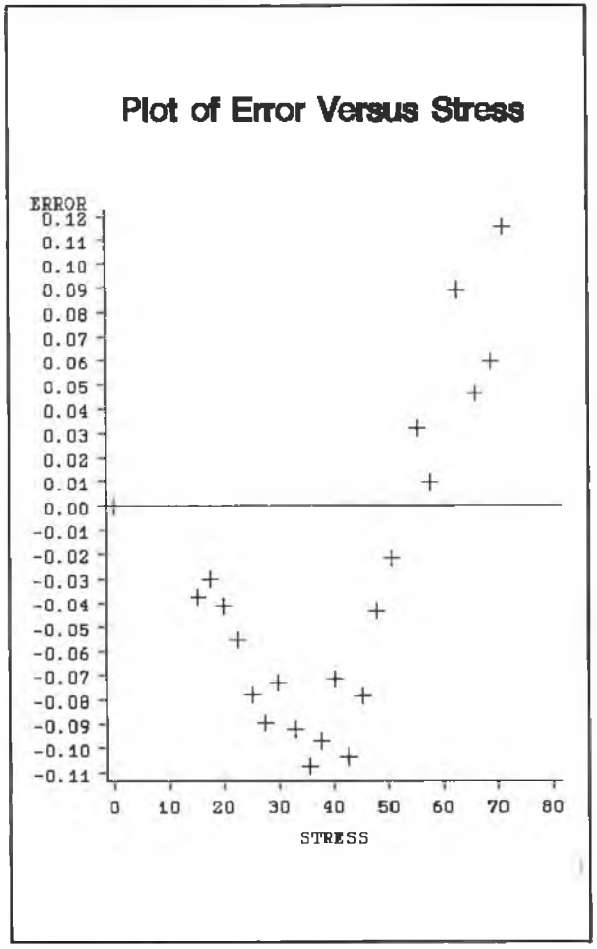


Figure C.30

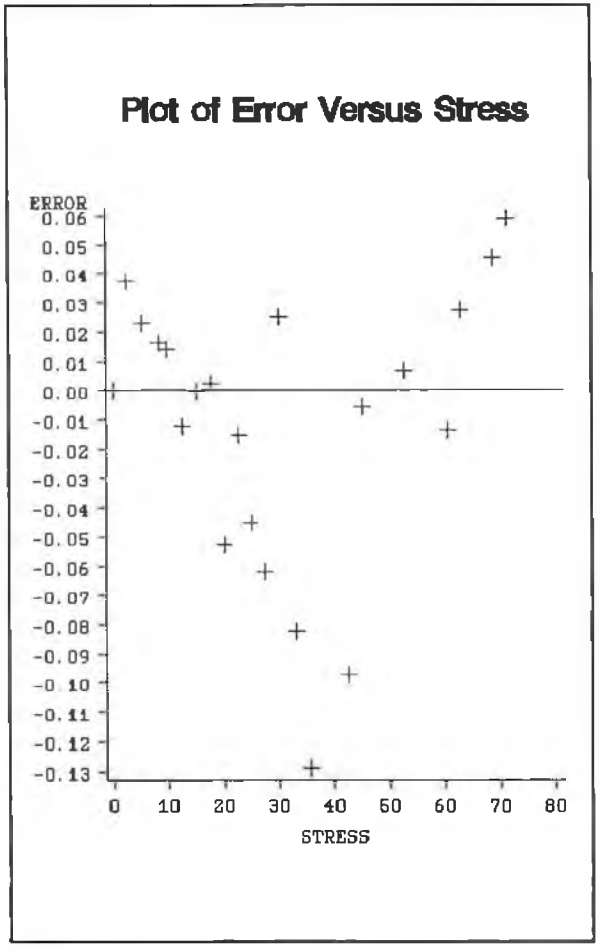


Figure C.31

Appendix D

Error versus stress plots with the autocorrelation terms included are shown in Figures(D.1)-(D.19) for do Carmo et al.(1988), in Figures (D.20)-(D.24) for Kehoe(unpublished), in Figures(D.25)-(D.31) for Daly(unpublished) and in Figures(D.32)-(D.36) for Campion et al.(1992).

Error versus Stress plots for each response in the <001> direction for the do Carmo et al.(1988) example with the autocorrelation terms included.

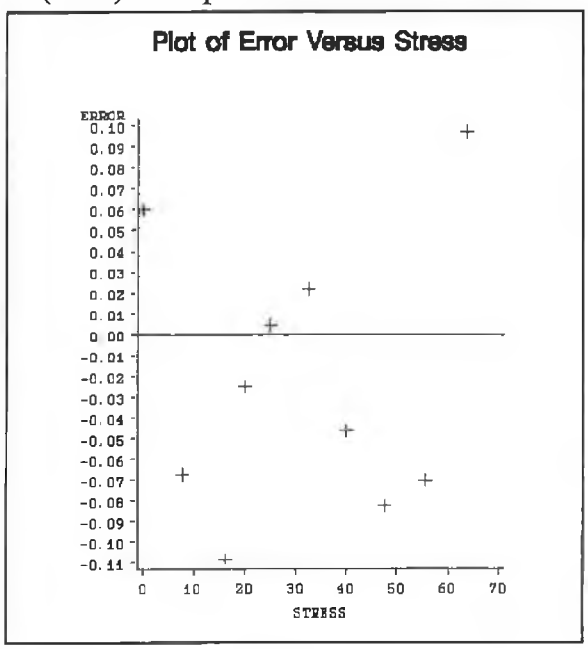


Figure D.1

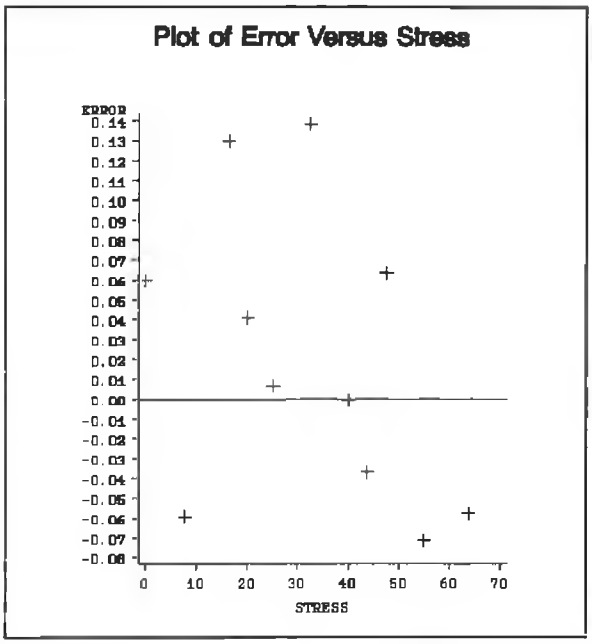


Figure D.2

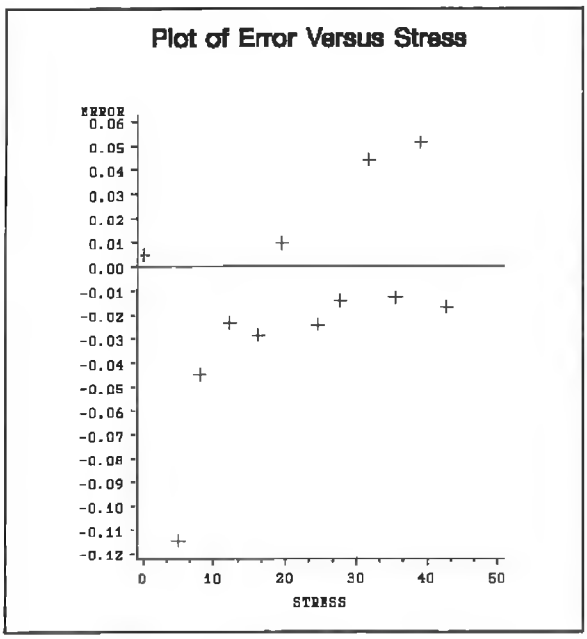


Figure D.3

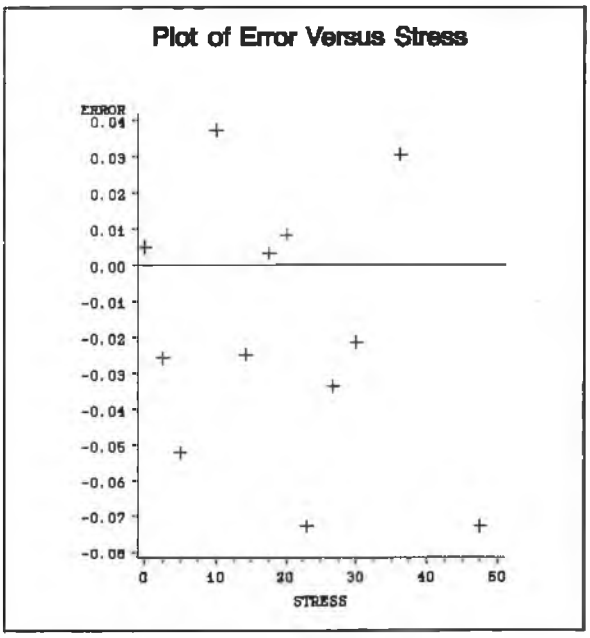


Figure D.4

Figures D.5-D.11 Error versus Stress plots in the <111> direction, with the autocorrelation terms included, for the do Carmo et al.(1988) example.

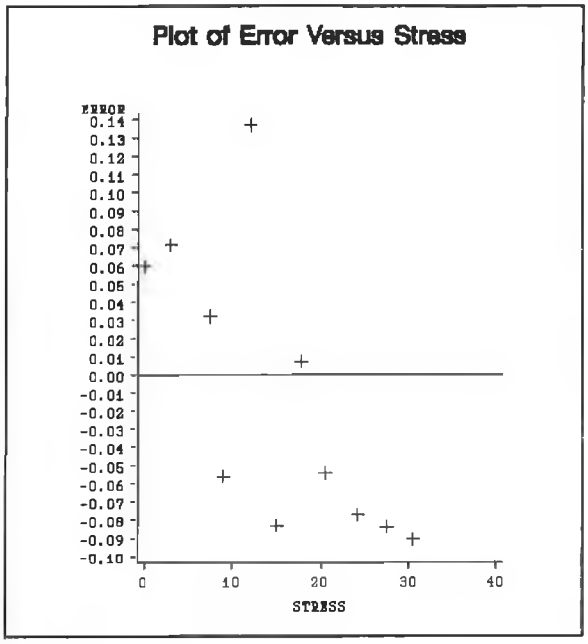


Figure D.5

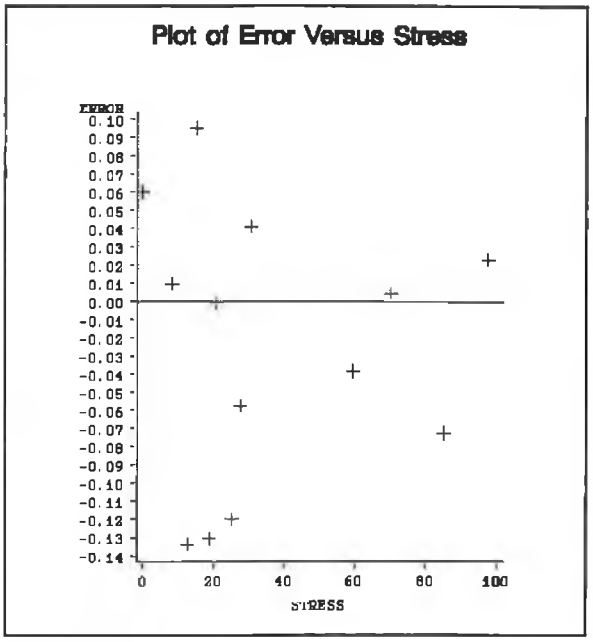


Figure D.6

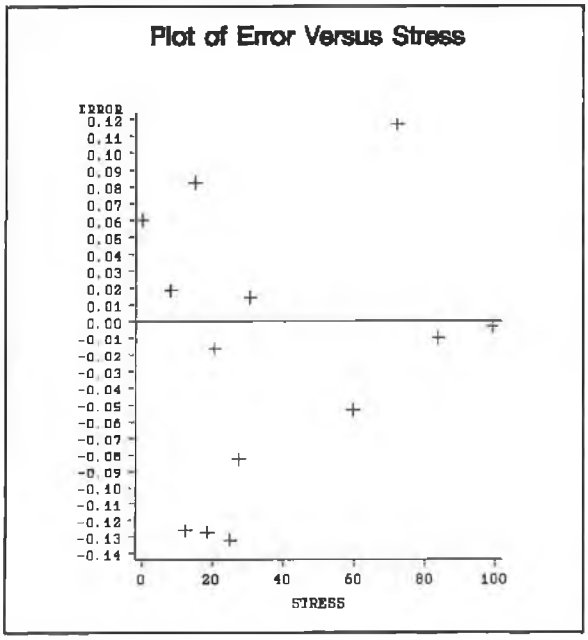


Figure D.7

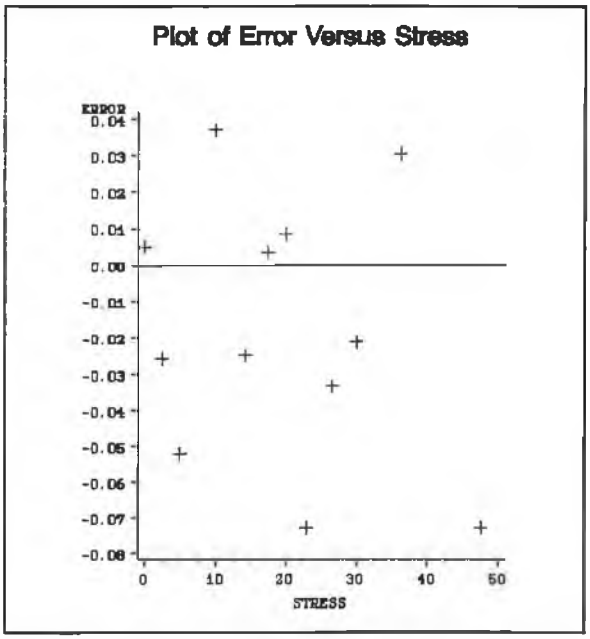


Figure D.8

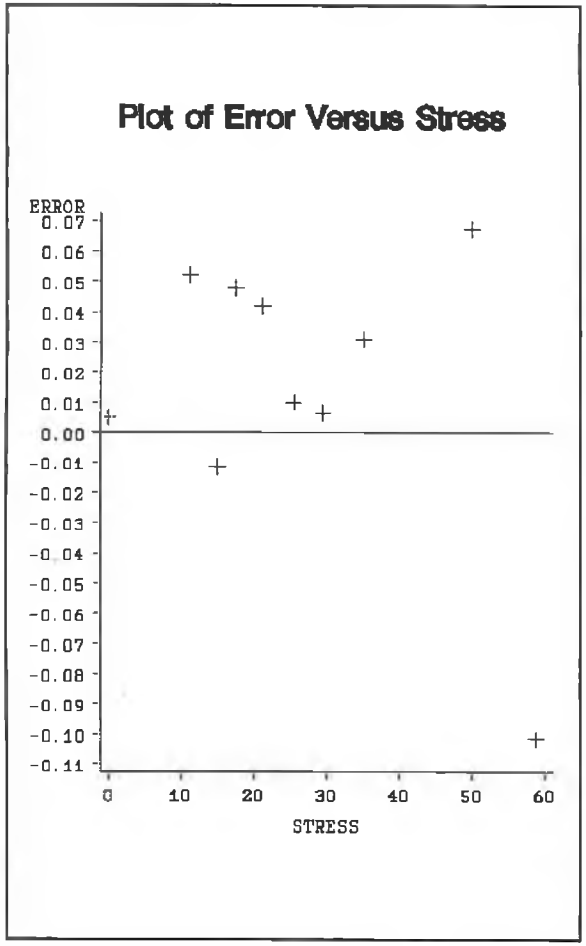


Figure D.9

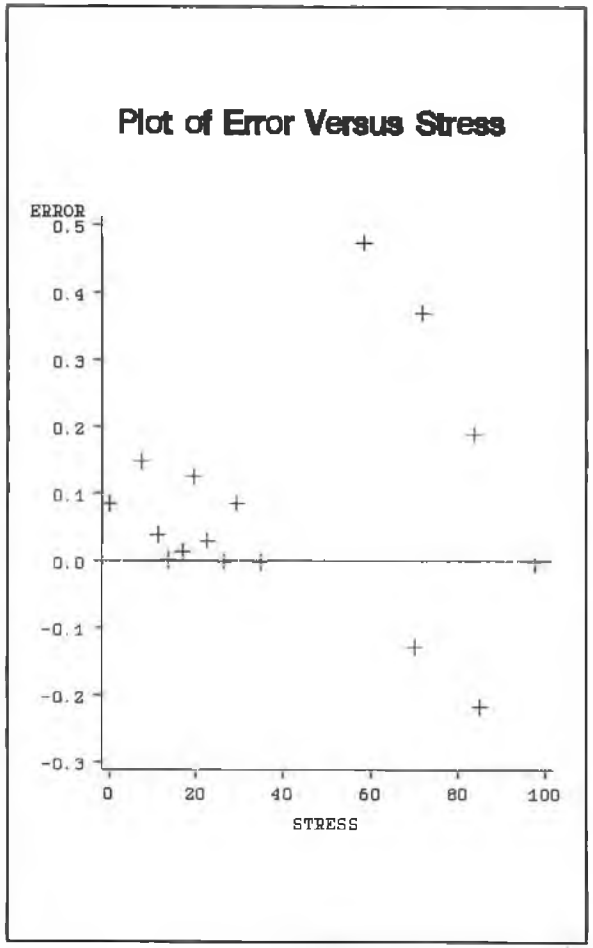


Figure D.11

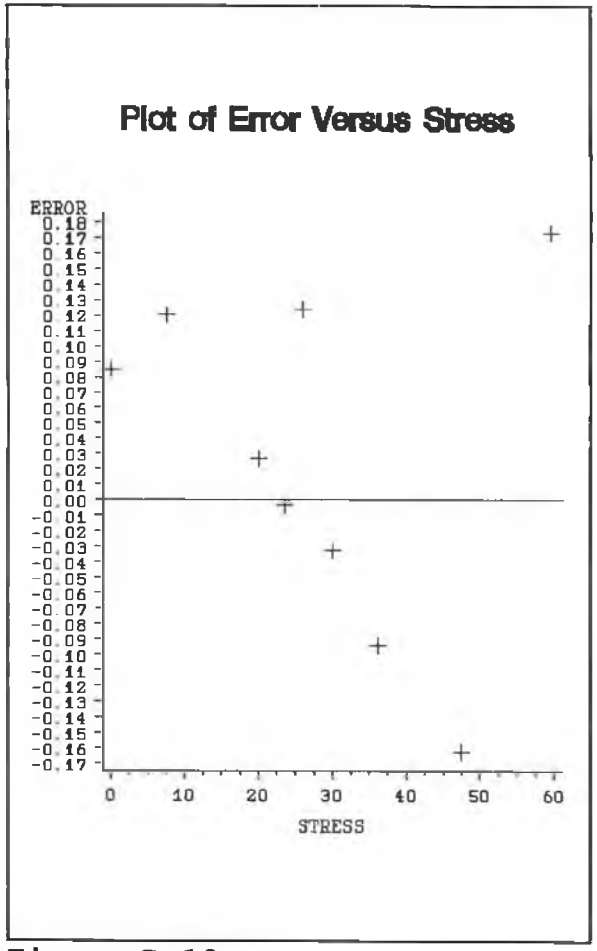


Figure D.10

Figures D.12 to D.19 : Error versus Stress plots for the individual responses in the <110> direction for the example given do Carmo et al.(1988), with the auto-correlation terms included.

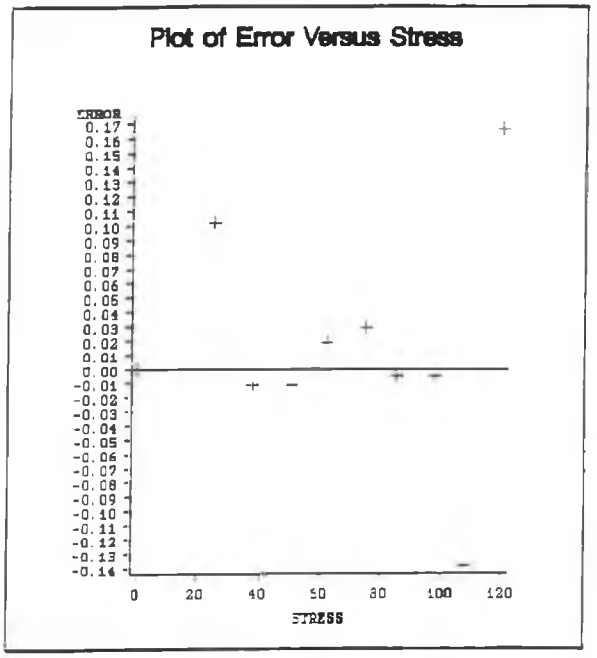


Figure D.12

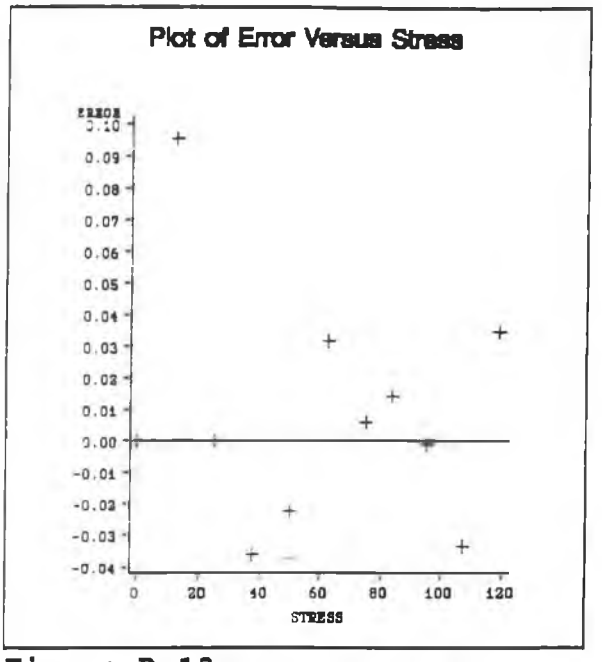


Figure D.13

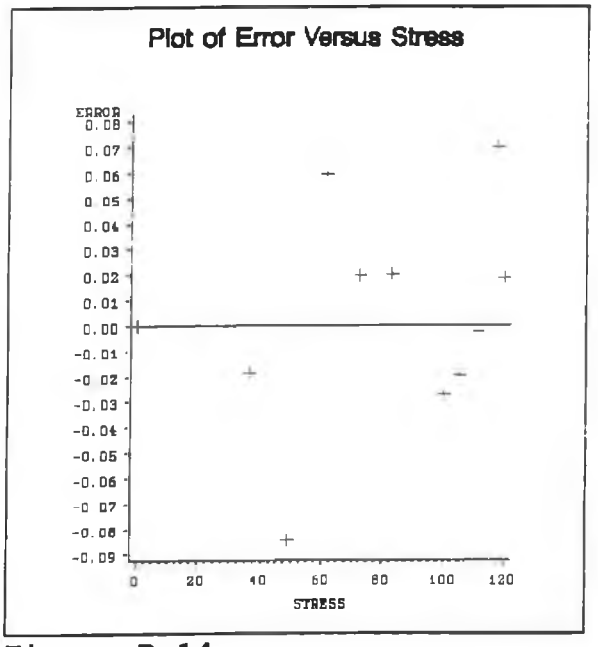


Figure D.14

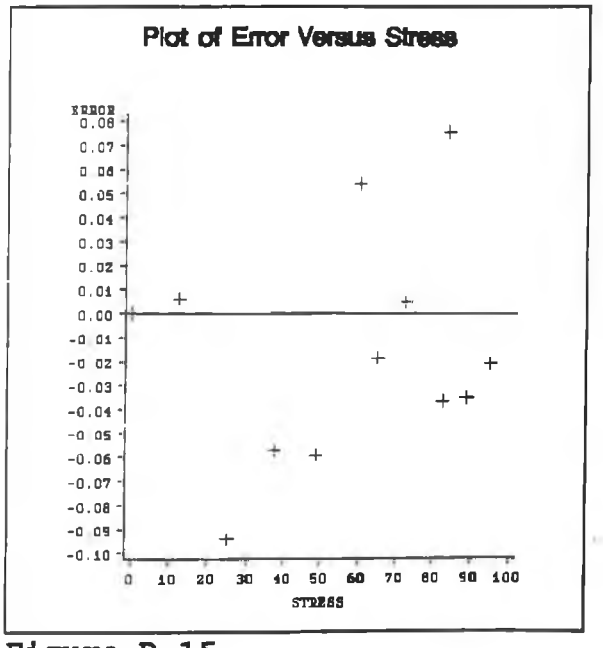


Figure D.15

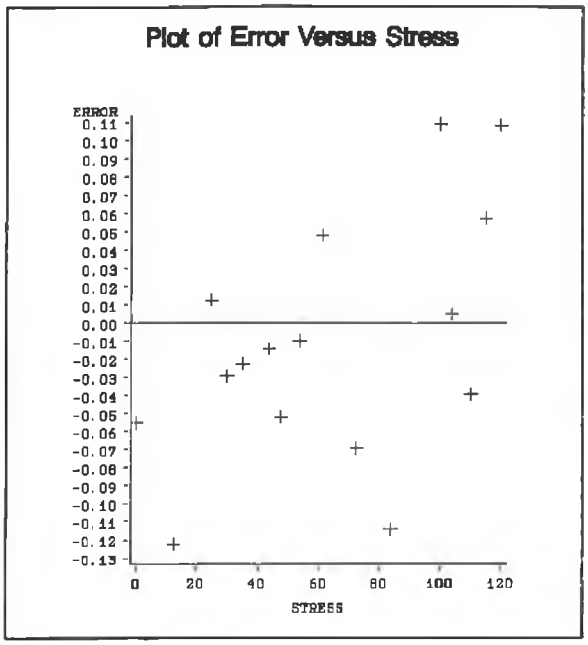


Figure D.16

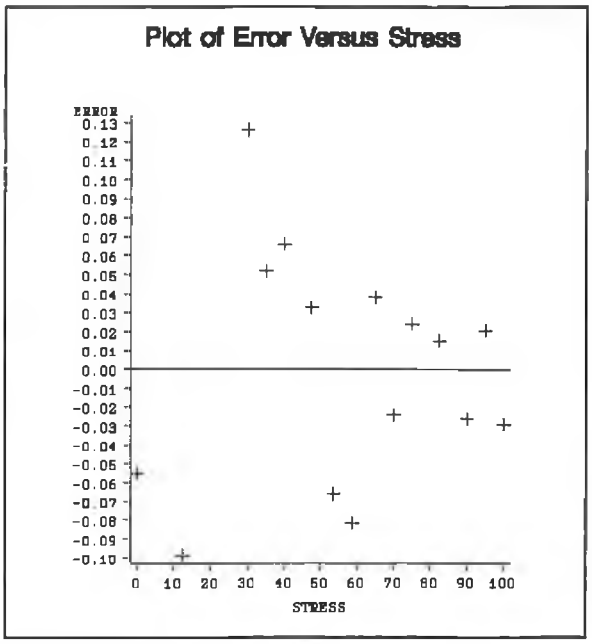


Figure D.17

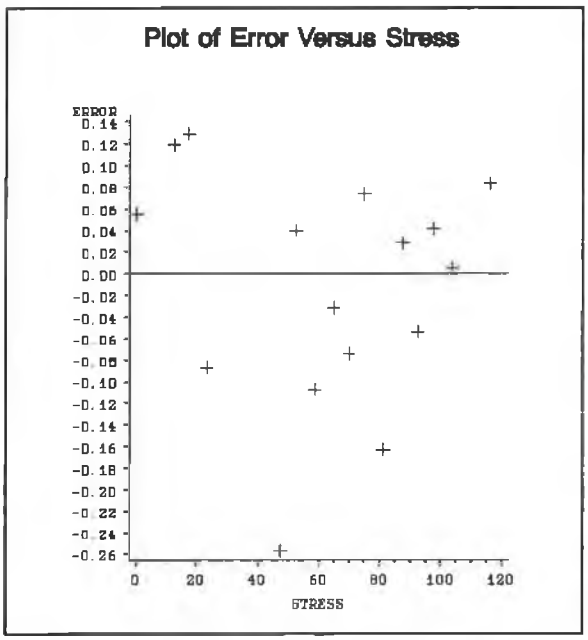


Figure D.18

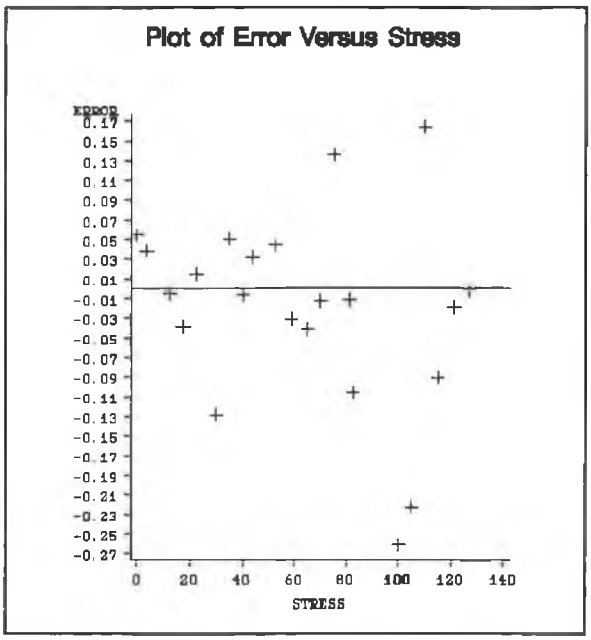


Figure D.19

Error versus Stress Plots for the <111> and <001> directions for Kehoe(unpublished), with the autocorrelation terms included.

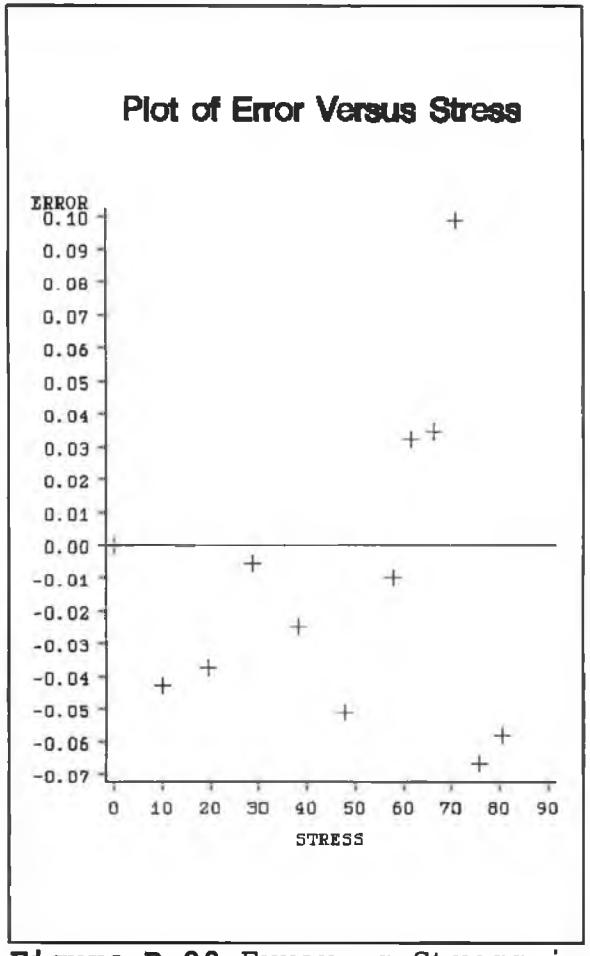


Figure D.20 Error vs Stress in <111>

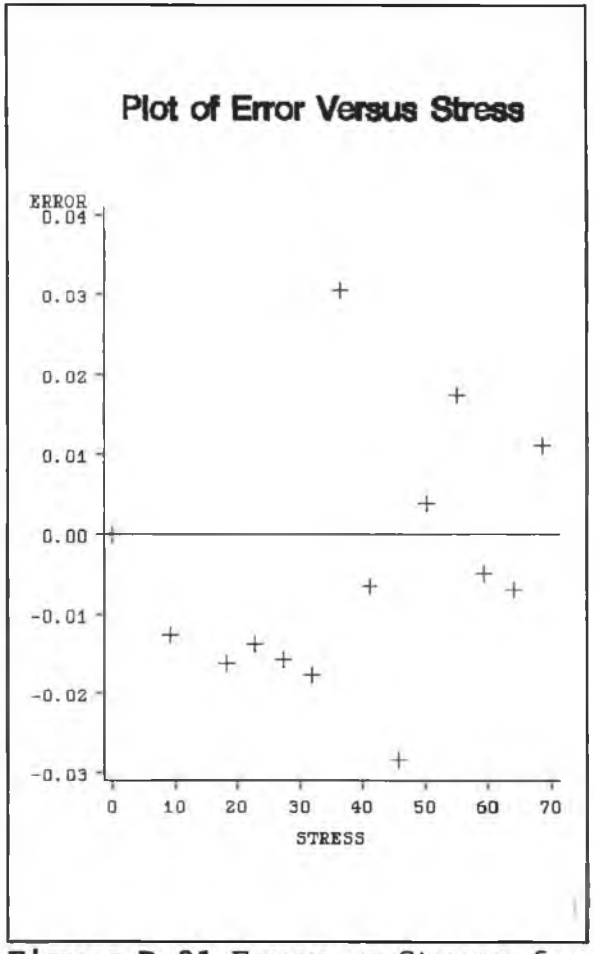


Figure D.21 Error vs Stress for <001> direction

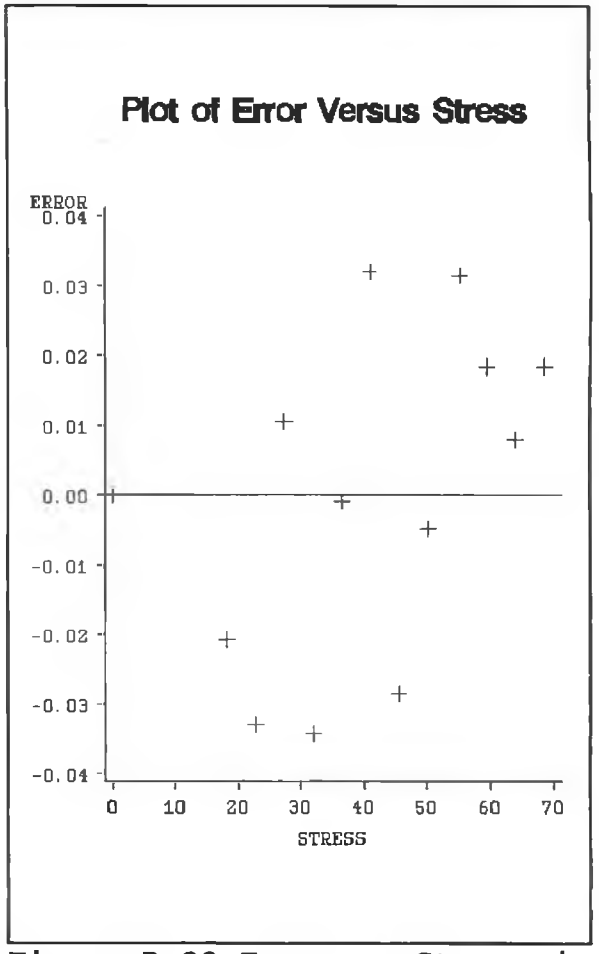


Figure D.22 Error vs Stress in <001>.

Error versus stress for the <110> direction in Kehoe(unpublished), with the autocorrelation terms included.

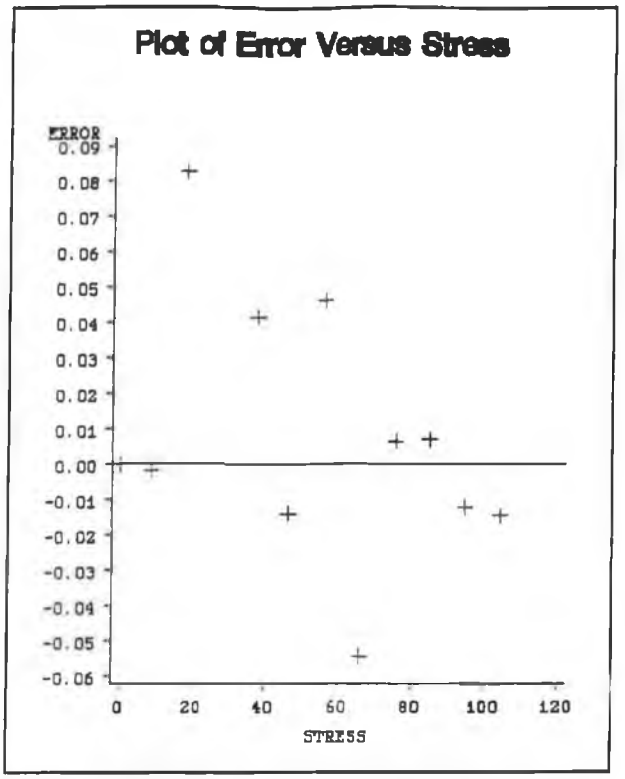


Figure D.23

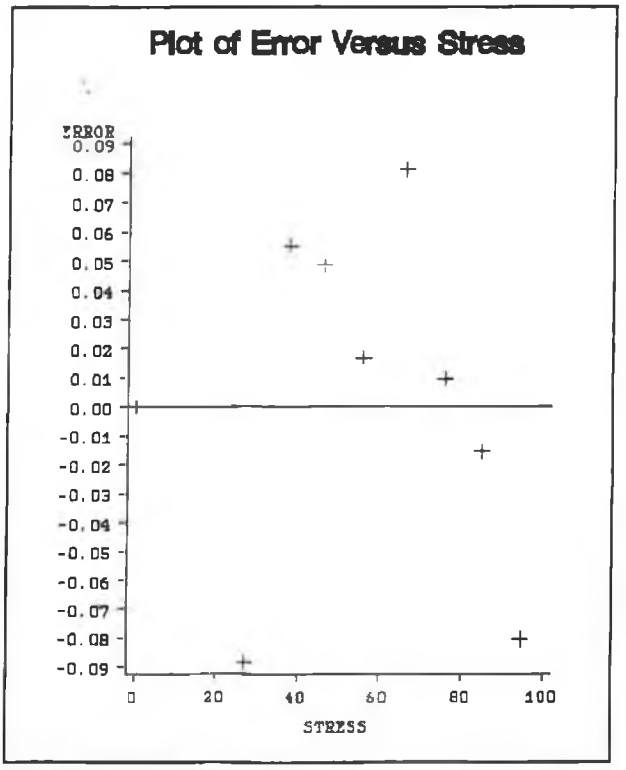


Figure D.24

Error versus Stress Plots for the individual responses in the <001> and <111> directions for the Daly(unpublished) example with the autocorrelation terms included.

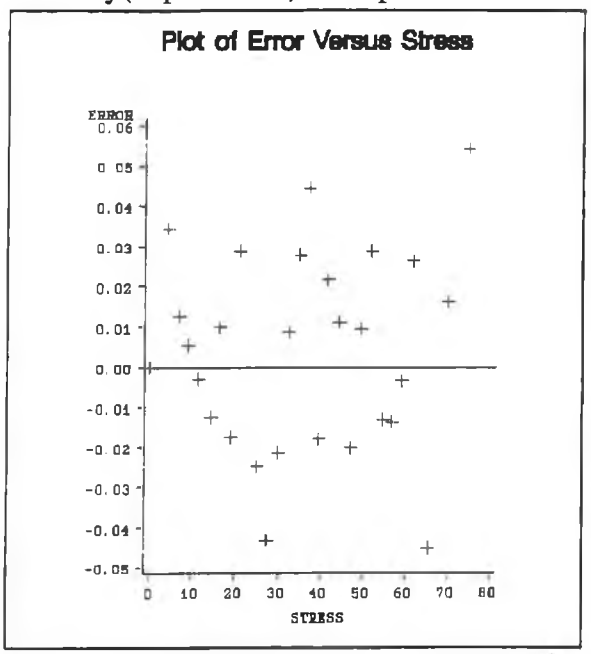


Figure D.25 Error vs Stress in <001>.

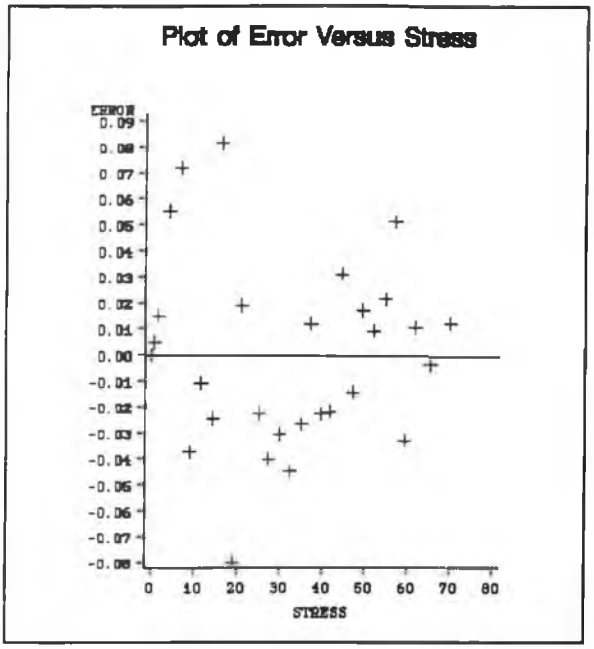


Figure D.26 Error vs Stress in <001>.

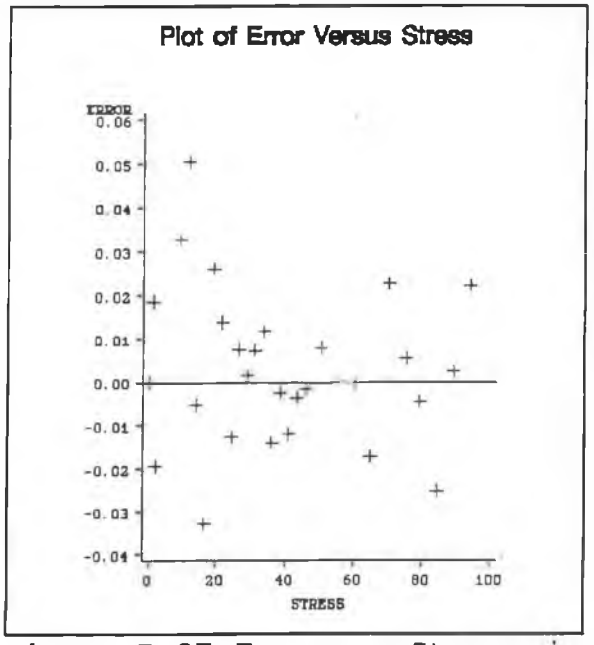


Figure D.27 Error vs Stress in <111>

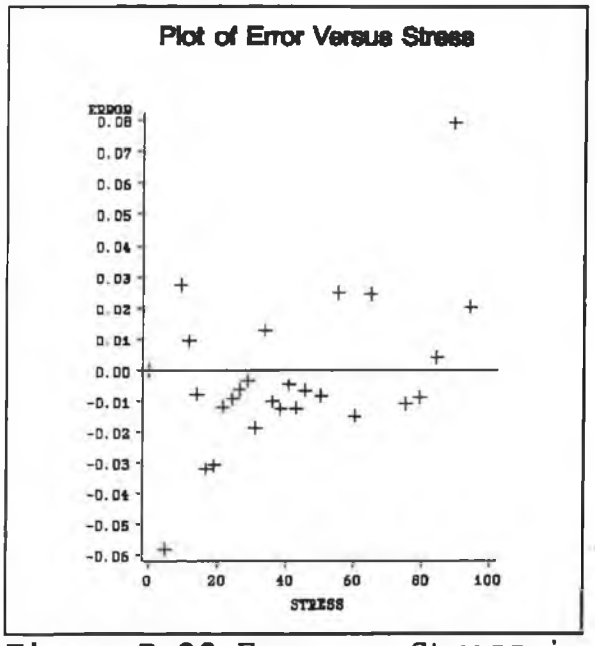


Figure D.28 Error vs Stress in <111>.

Error versus Stress plots for the individual responses in the <110> direction for the Daly(unpublished) example with the autocorrelation terms included.

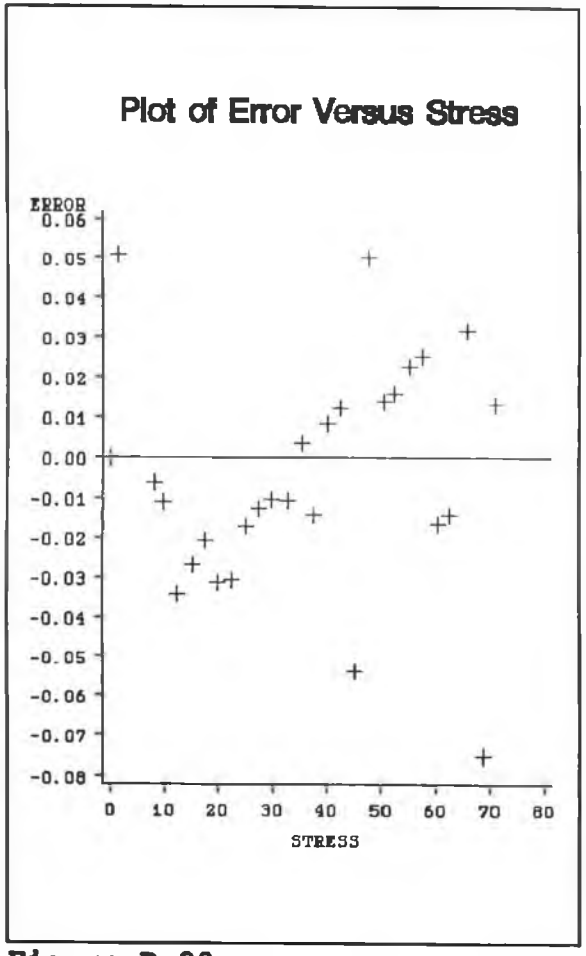


Figure D.29

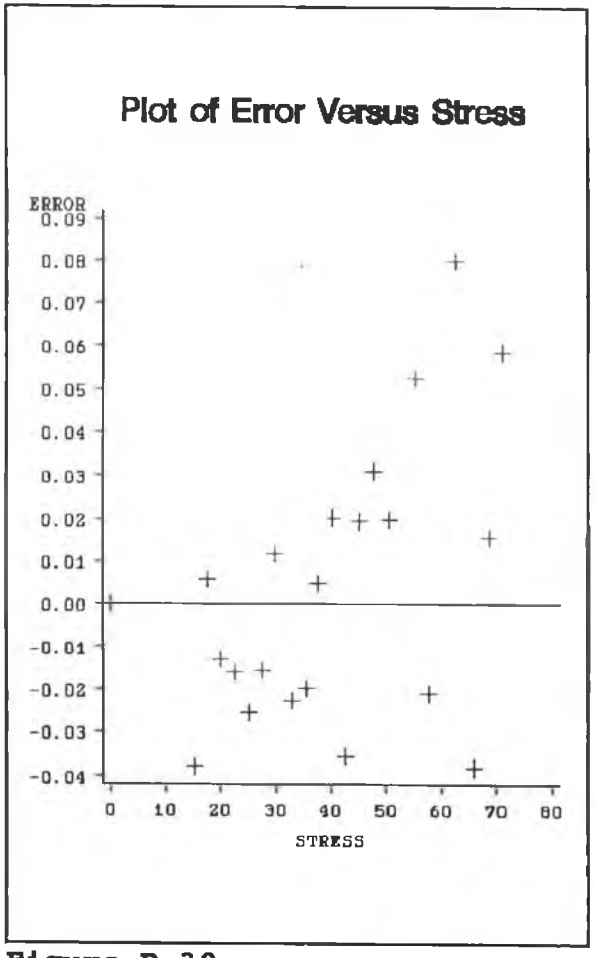


Figure D.30

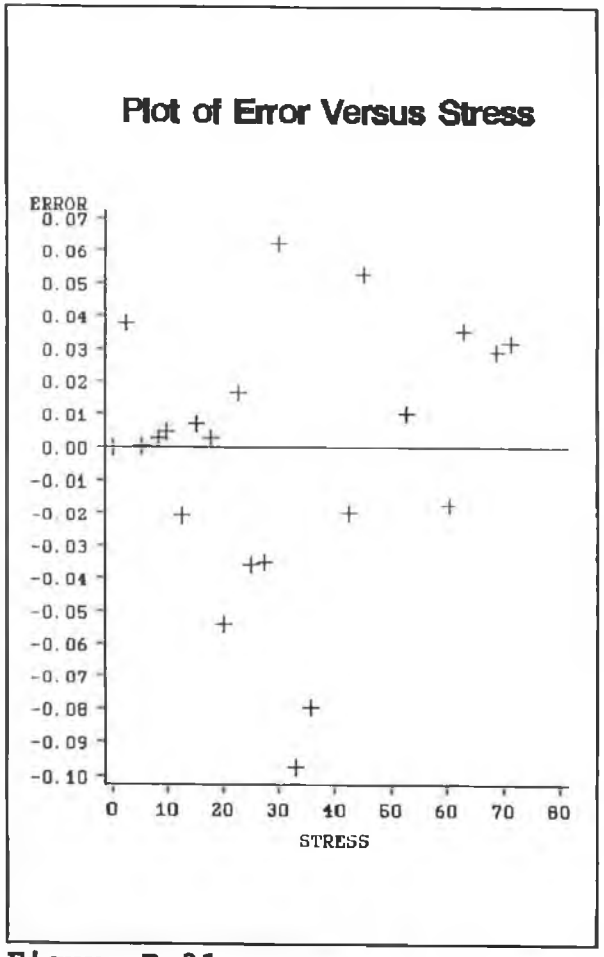


Figure D.31

Error versus Stress plots in Figures (D.32)-(D.33) for the <001> direction and in Figures (D.34)-(D.36) for the <111> direction for the Campion et al(1992) example with the autocorrelation terms included.

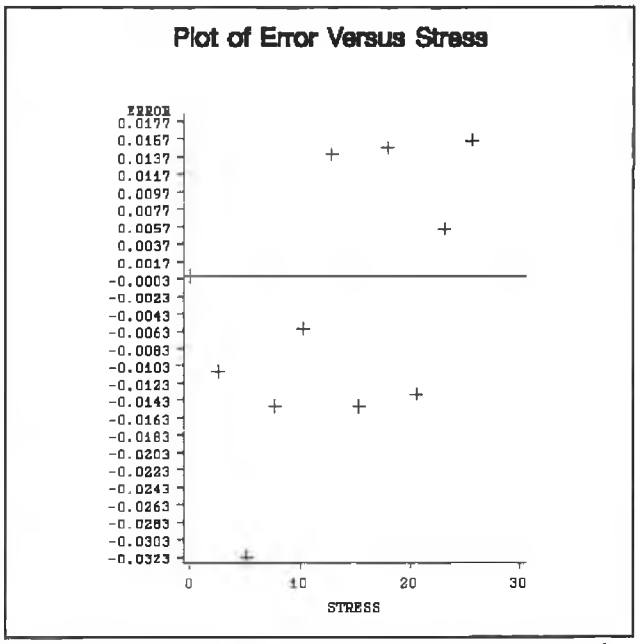


Figure D.32 Error vs Stress in <001>

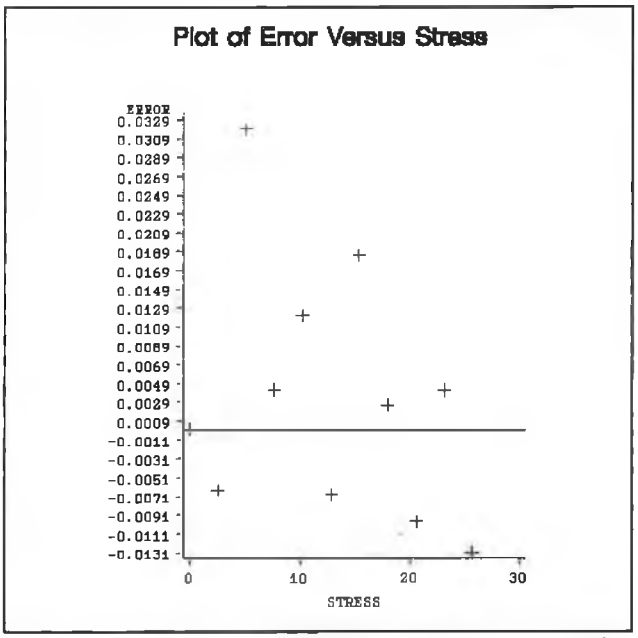


Figure D.33 Error vs Stress in <001>

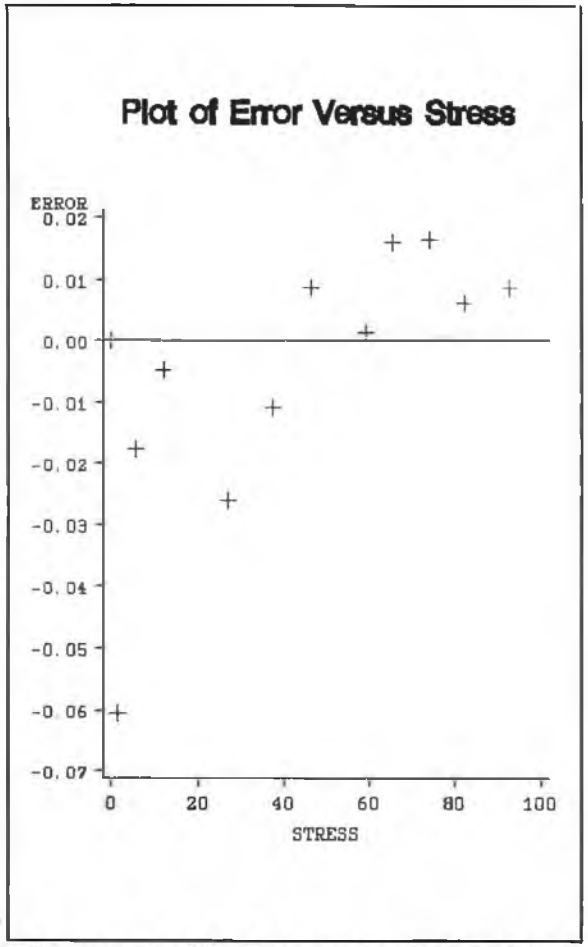


Figure D.34 Error vs Stress in <111>.

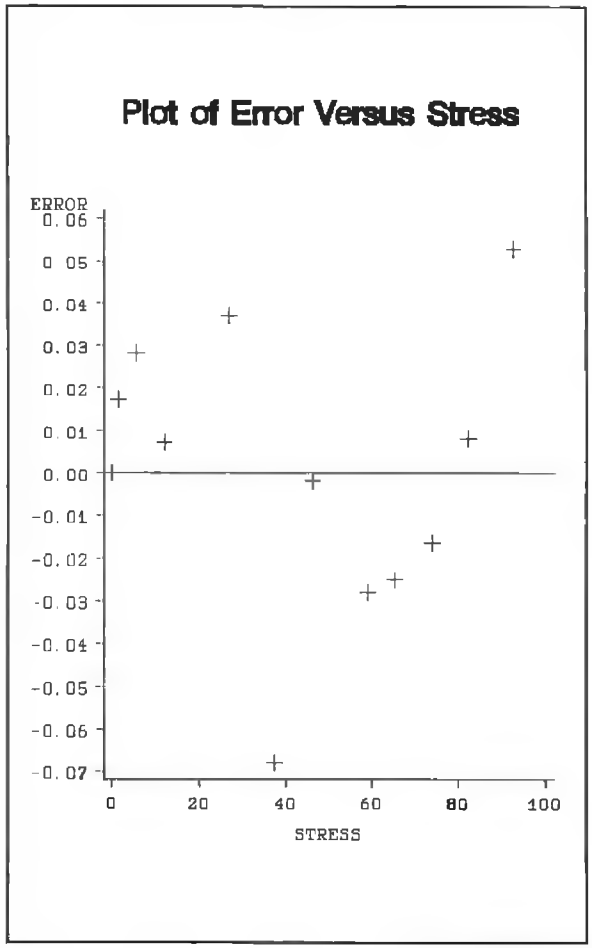


Figure D.35 Error vs Stress in <111>

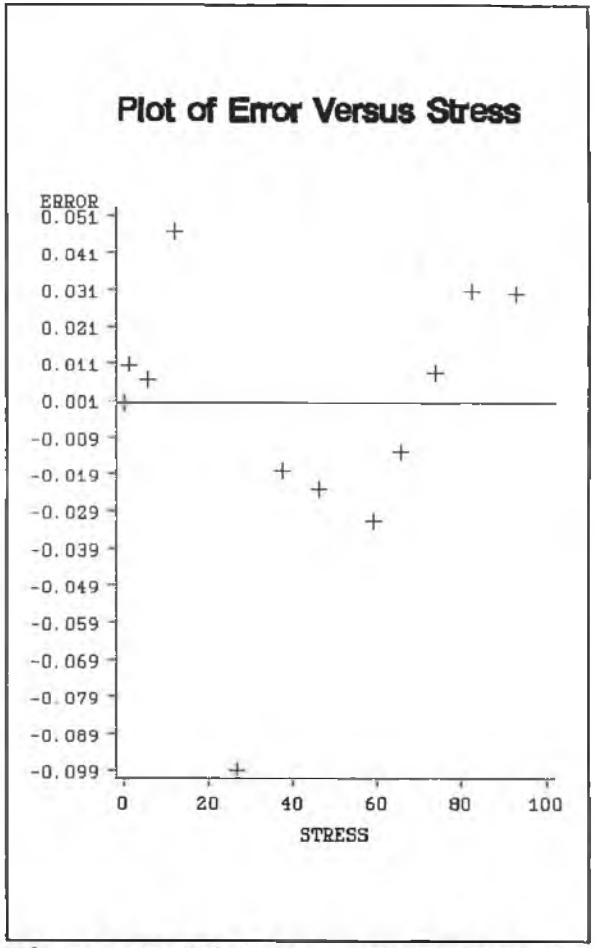
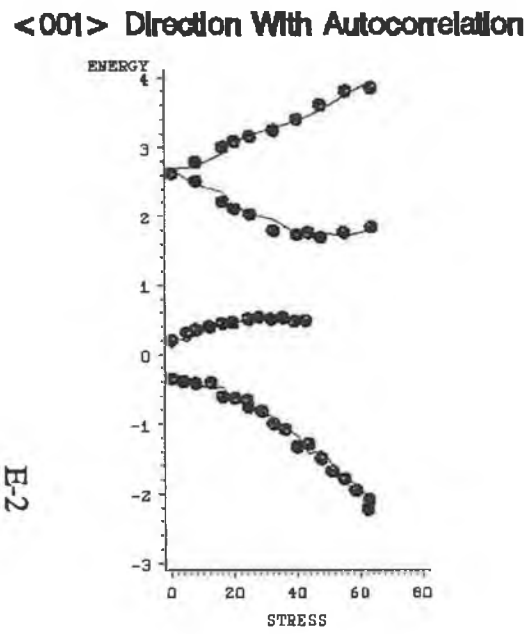


Figure D.36 Error vs Stress in <111>

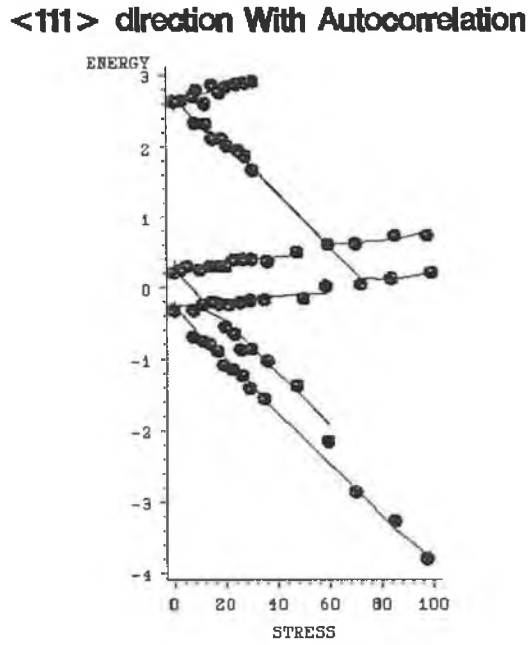
Appendix E

The fits obtained using the Powell-shell with the autocorrelation terms included for each response for the problems described by do Carmo et al(1988), Figures (E.1)-(E.3), Daly(unpublished), Figures (E.4)-(E.6) and Kehoe (unpublished), Figures (E.7)-(E.9). These should be compared with Figures (4.7)-(4.8) for do Carmo et al.(1988), Figures (4.10)-(4.12) for Kehoe(unpublished) and with Figures (A.4)-(A.6) for Daly(unpublished).



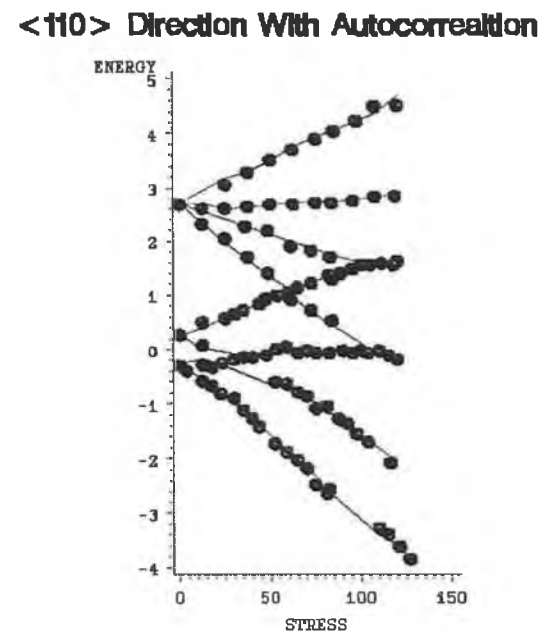
Circle denote energy data points
Joint Lines denote estimated points

Figure E.1



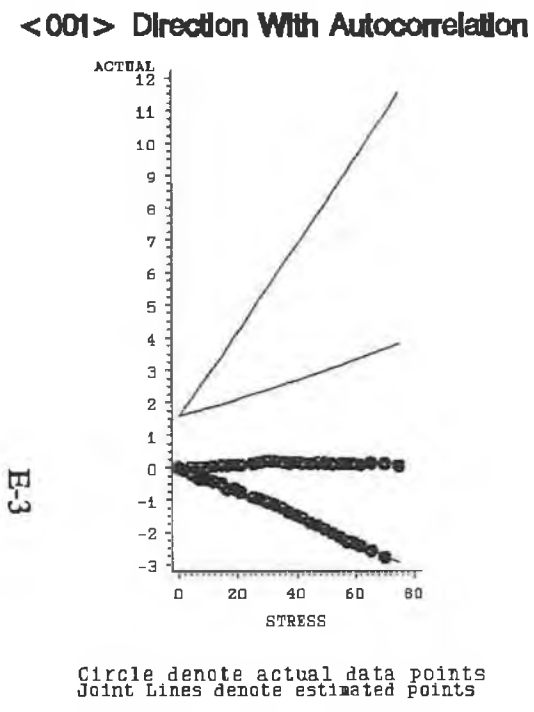
Circle denote energy data points
Joint Lines denote estimated points

Figure E.2



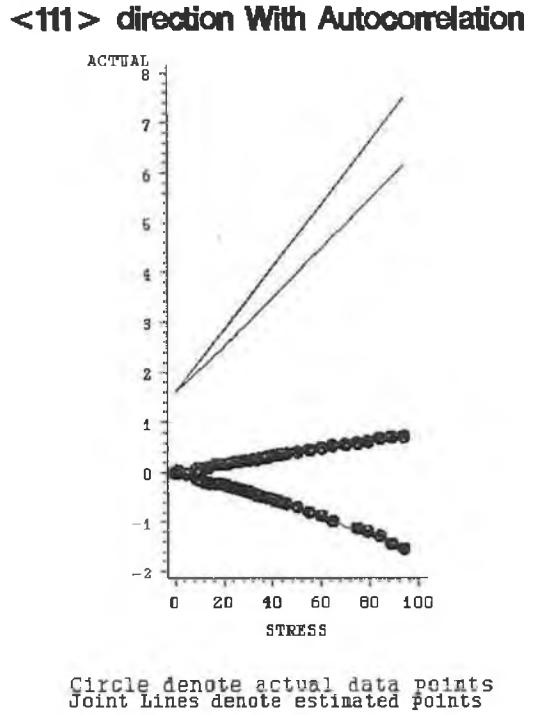
Circle denote energy data points
Joint Lines denote estimated points

Figure E.3



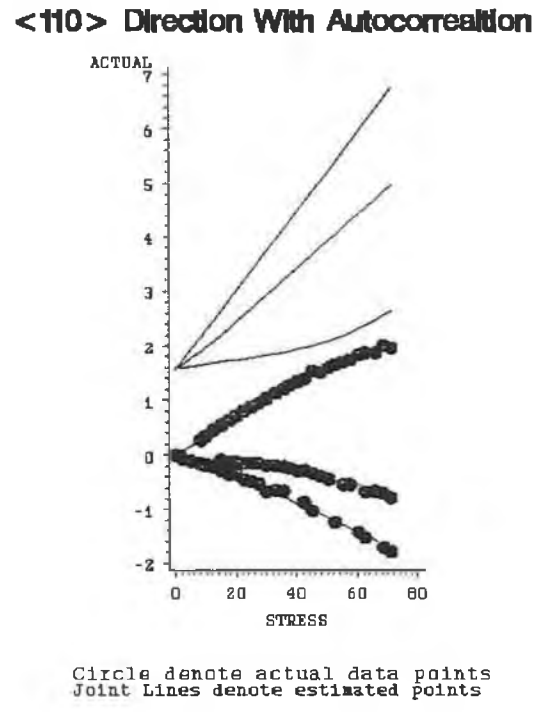
Circle denote actual data points
Joint Lines denote estimated points

Figure E.4



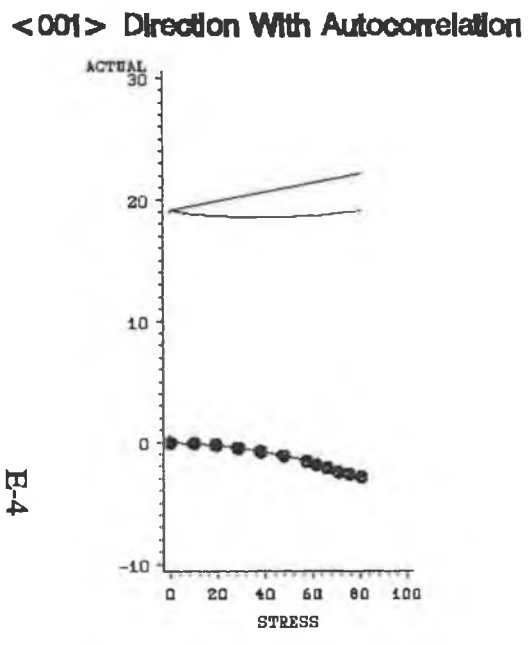
Circle denote actual data points
Joint Lines denote estimated points

Figure E.5



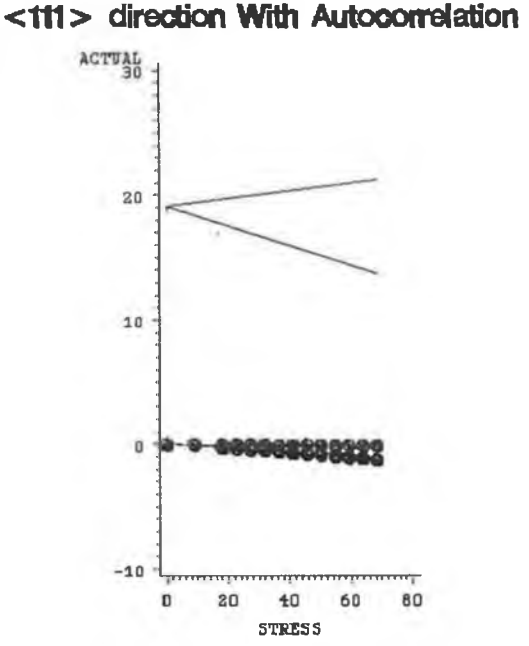
Circle denote actual data points
Joint Lines denote estimated points

Figure E.6



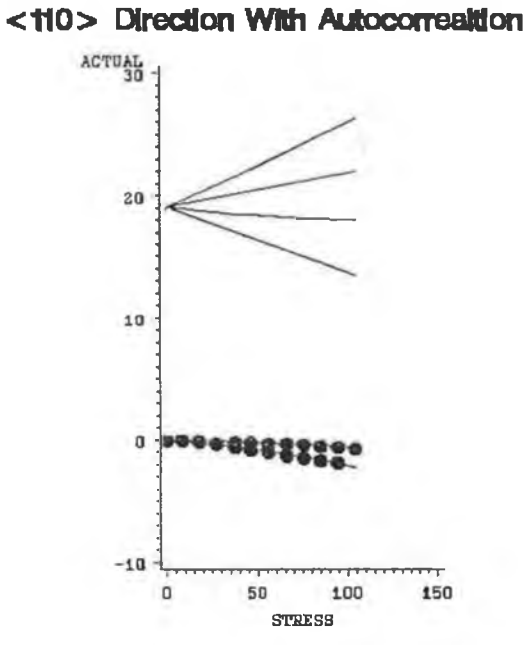
Circle denote actual data points
Joint Lines denote estimated points

Figure E.7



Circle denote actual data points
Joint Lines denote estimated points

Figure E.8



Circle denote actual data points
Joint Lines denote estimated points

Figure E.9

Appendix F

A selection of the surface plots described in section 5.7 are shown in Figures (F.1)-(F.4) for Campion et al.(1992), in Figures (F.5)-(F.8) for do Carmo et al.(1988) with E and L excluded, in Figures (F.9)-(F.12) for Kehoe(unpublised) and in Figures(F.13)-(F.16) for do Carmo et al.(1988) with E and L included.

A selection of surface plots for Campion et al.(1992), where SSq corresponds to the Sum of Squares.

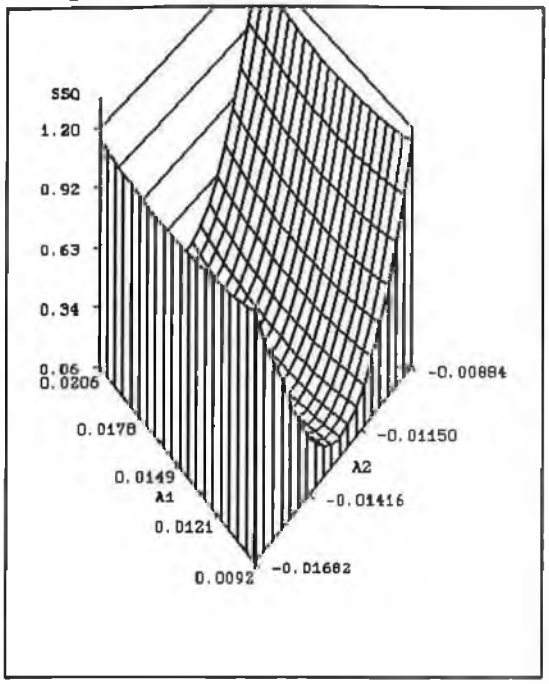


Figure (F.1)

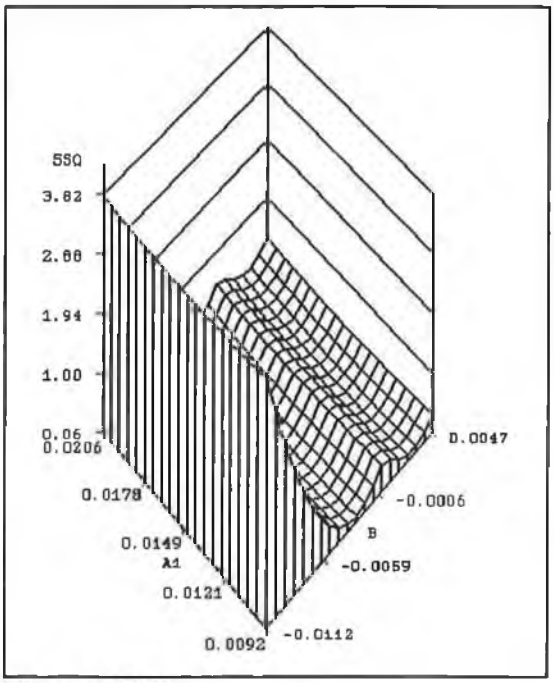


Figure (F.2)

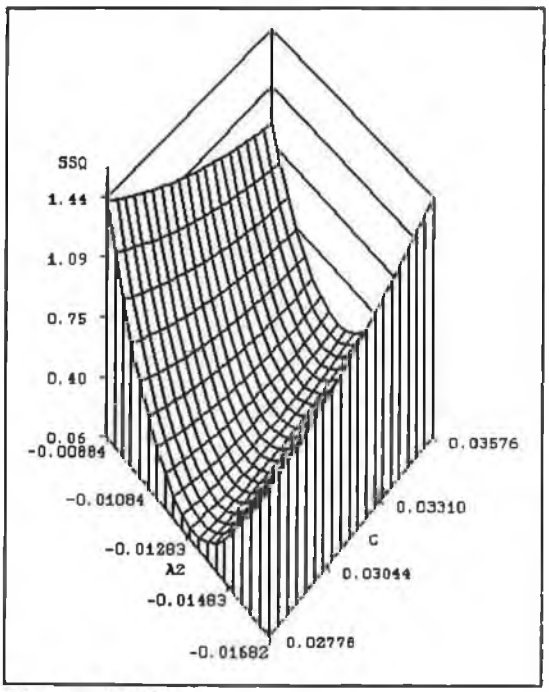


Figure (F.3)

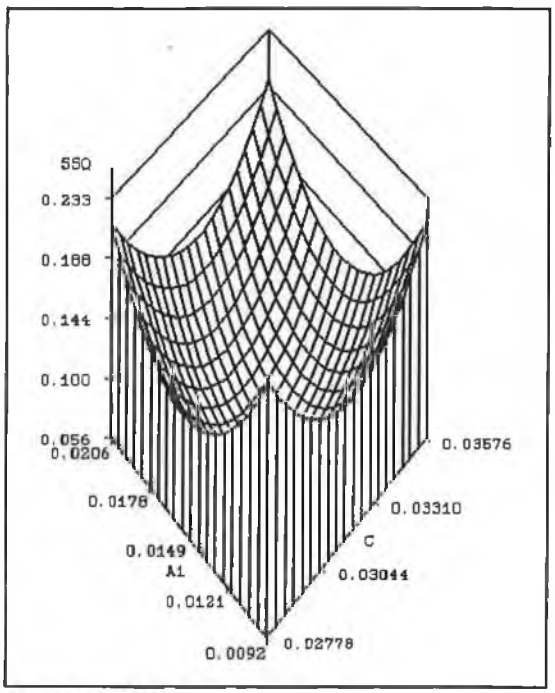


Figure (F.4)

A selection of surface plots for do Carmo et al.(1988), where SSq corresponds to the Sum of Squares, and E and L are excluded.

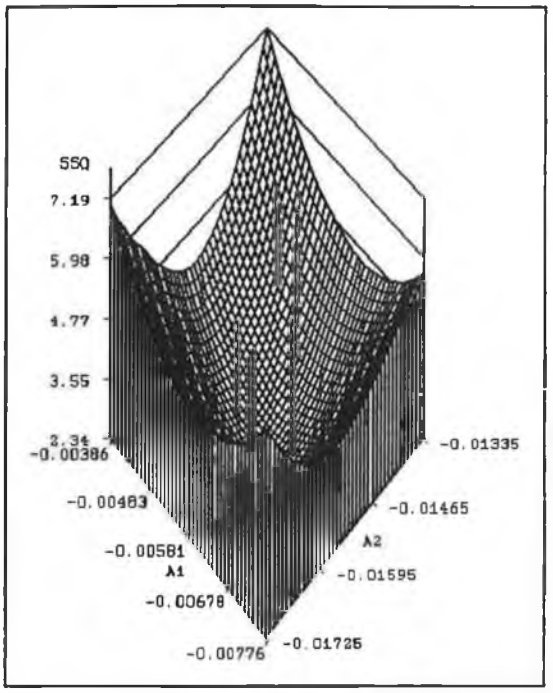


Figure (F.5)

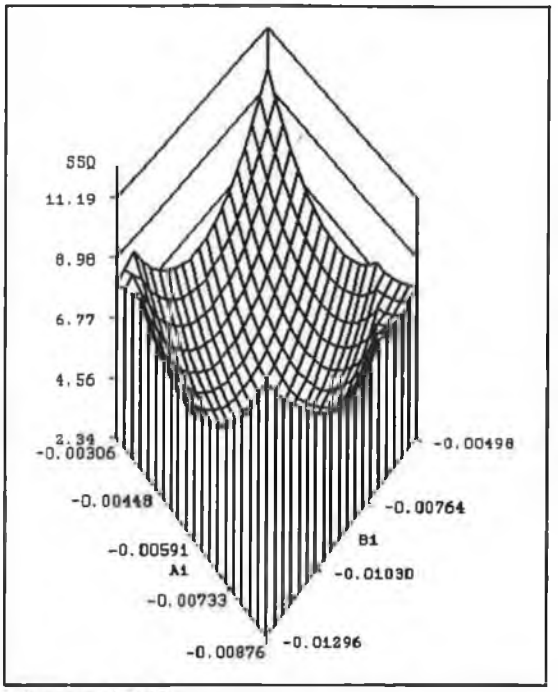


Figure (F.6)

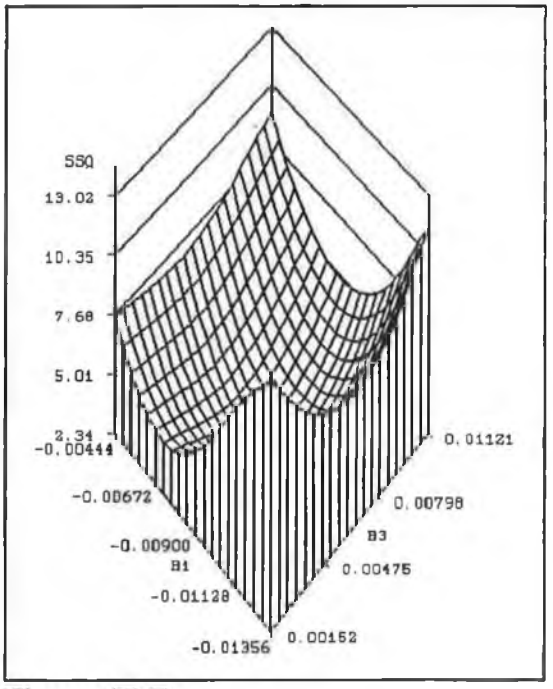


Figure (F.7)

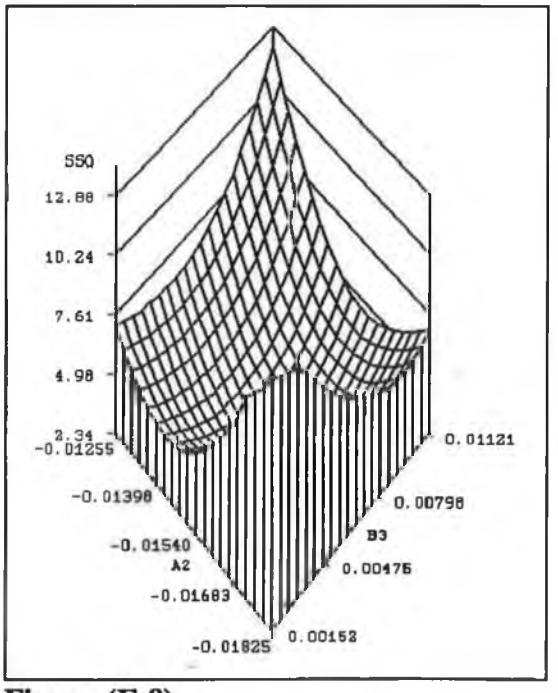


Figure (F.8)

A selection of surface plots for Kehoe(unpublished), where SSq corresponds to the Sum of Squares.

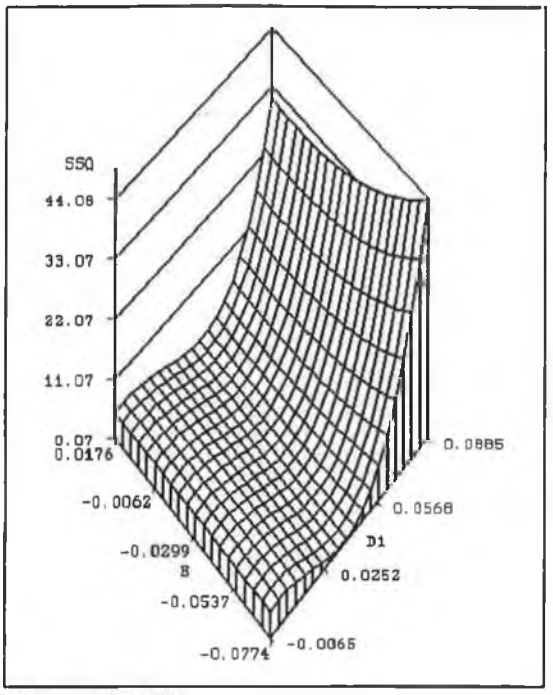


Figure (F.9)

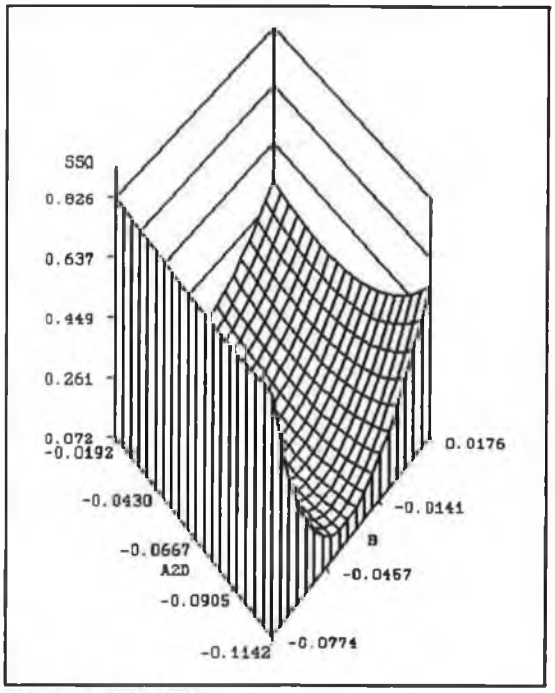


Figure (F.10)

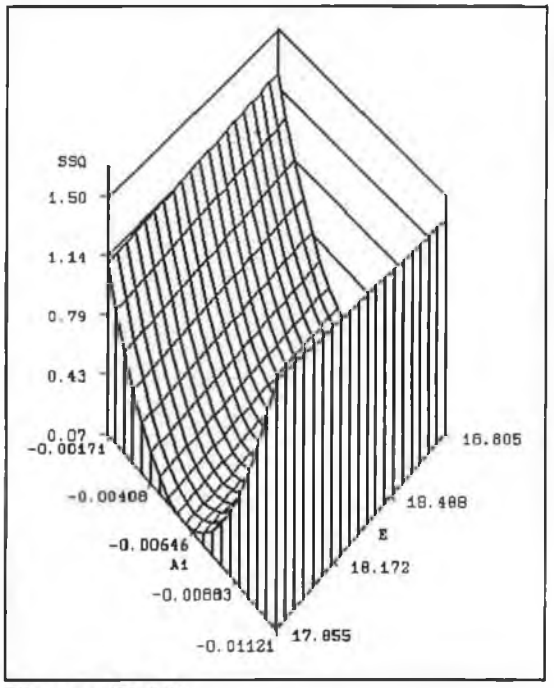


Figure (F.11)

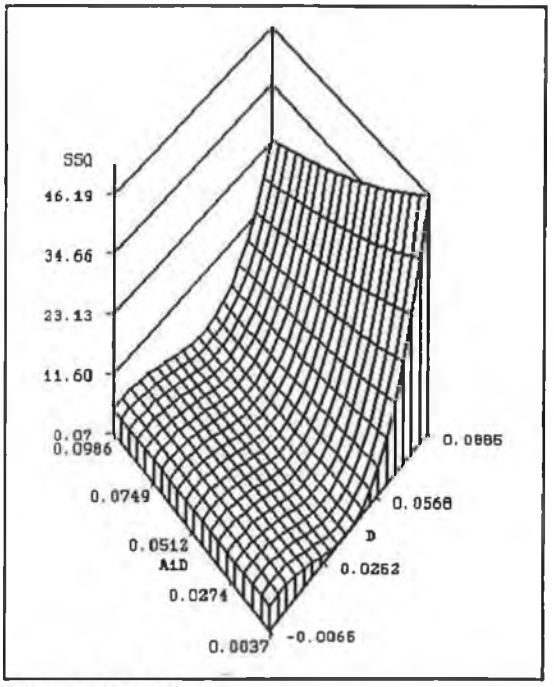


Figure (F.12)

A selection of surface plots for do Carmo et al.(1988), where SSq corresponds to the Sum of Squares, and E and L are included.

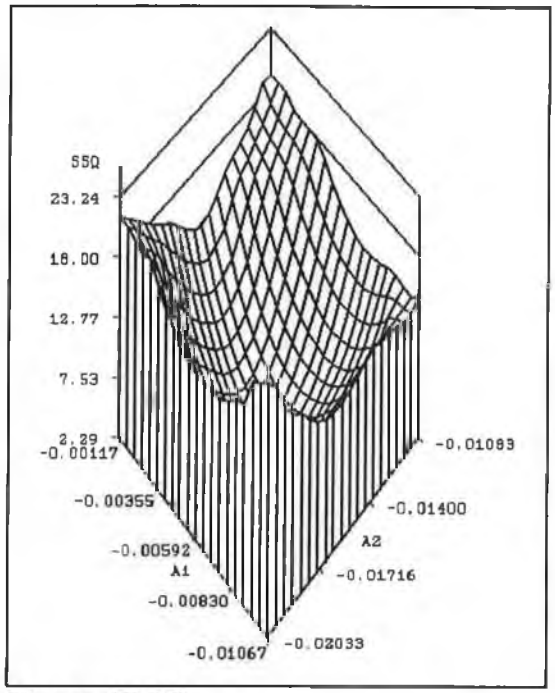


Figure (F.13)

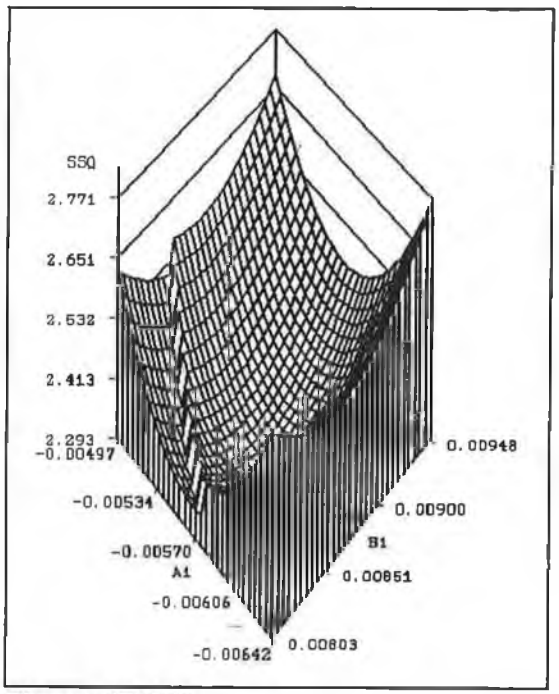


Figure (F.14)

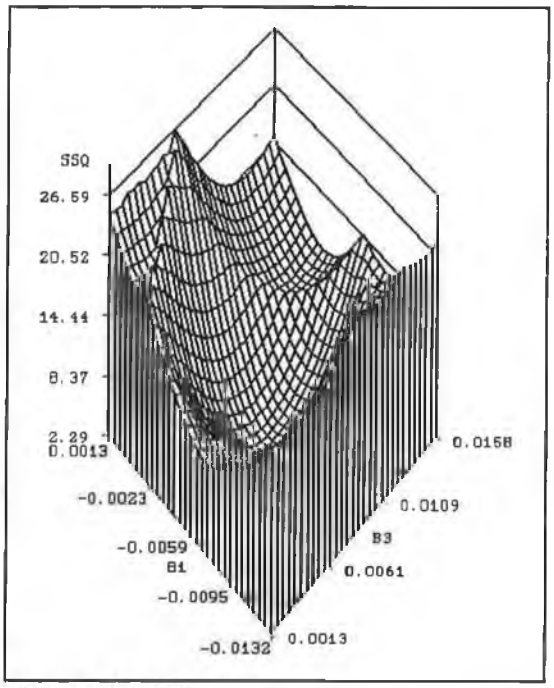


Figure (F.15)

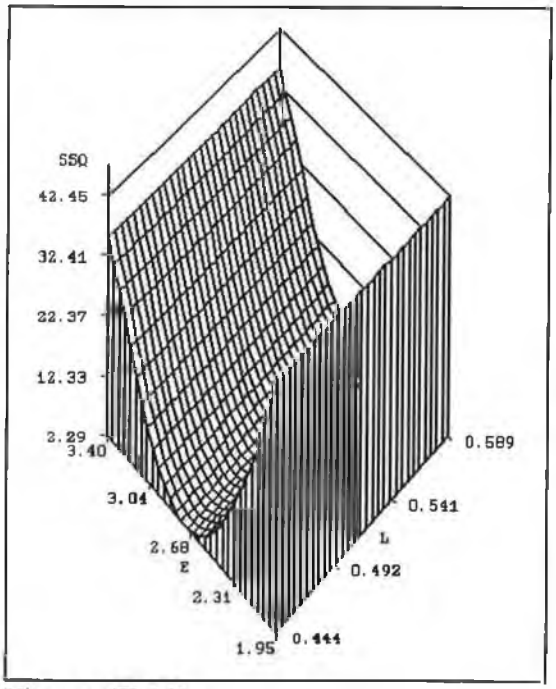


Figure (F.16)

Appendix G

The following are the list of files located on disk A.

Proto.c :- C prototype.

Proto.pas :- Pascal prototype.

Tran.txt :- TranId Format.

Coutour1.bat :- Surfer batch program for surfaces plots.

Coutour2.bat :- Surfer batch program for contour plots.

Man.txt : A brief user manual.

Piezo-spectroscopic data analysis: a PC tool

A.L. McCarren
H.J. Ruskin
K.G. McGuigan
M.O. Henry

Indexing terms: Piezo-spectroscopy, Powell algorithm, Spectroscopic data analysis

Abstract: Information from the many kinds of spectroscopy used by chemists and physicists is fundamental to our understanding of the structure of materials. Numerical techniques have an important role to play in the augmentation of the instrumentation and technology available in the laboratory, but are frequently viewed as separate from the laboratory procedures. We describe an integrated PC-based approach for obtaining *directly* the parameter estimates of transition types in piezo-spectroscopic measurements of crystalline materials. Typically, the analyses in question are required to handle complex secular matrices, to distinguish between components in the experimental results, and to identify the transition types as rapidly and as efficiently as possible. The method described, based on providing a discrete shell to the Powell algorithm, is shown to give both accurate identification of the transition type in the case of new data and improved fits (i.e. reduction in residual variation) when compared with results obtained via standard procedures. In addition it is flexible with respect to the language used and possesses a high degree of portability. We illustrate the success of the approach using (i) data previously reported on the solution of a trigonal defect which includes both mixing of states and spin orbit interactions and (ii) new data obtained for a defect related to beryllium impurities in silicon.

1 Introduction

The fitting of spectroscopic data is a topic which is characterised by a lack of coherent reporting in the literature. This is due both to the diversity of current experimental techniques and applications and the obvious corollary of the lack of a common forum in terms of reporting on the analyses performed. Such considerations have ensured that the effort of reconciling common features in the underlying linear models is nontrivial, even where such models have long been employed by mathematicians, statisticians and computer scientists. Similarly, an informed discussion of the scope of the optimisation technique used is usually lacking, since the primary aim is to interpret the values obtained for the parameters defining the problem rather than to assess the efficiency of the estimation process. The wider aspects of achieving a good fit to spectroscopic data have not, of course, been entirely ignored and notable examples include References 1–3 and references cited therein.

In this paper we describe a PC-based approach to the fitting of data obtained in piezo-spectroscopic measurements of crystalline materials. The central issues are those of obtaining the best fit, i.e. the most precise estimates of the problem's parameters, and identifying the transition which produces the spectrum. The requirements for the tool include flexibility, good interactive facilities and comparable performance to that achieved on a mainframe, (VAX/VMS 6230 clustered with Vax Stations), for a reasonable choice of PC. Procedures for analysis of data of this type typically involve obtaining initial estimates by preliminary manipulation of the data and the application of an optimisation technique which finds the roots of the characteristic polynomial (i.e. the eigenvalues of the energy shift matrix) [4]. In the present instance, however, these methods do not yield directly the quantities of interest, namely the parameters defining the optical transition. The alternative approach which we describe involves the provision of a discrete shell to the Powell algorithm [5, 6]. The method is used to re-examine published data for a defect related to iron impurity atoms in silicon [7] and to analyse new data for a defect in beryllium doped silicon [8].

2 Background physics

The study of electronic and optoelectronic materials has been intensive for decades. This reflects the central role of materials analysis and characterisation in the improvement and development of device technology. For these materials, much of the research has been directed at obtaining a full understanding of the properties of defects. This applies equally to defects deliberately produced for device optimisation and to unwanted defects which degrade device performance.

Spectroscopic techniques, of which electron paramagnetic resonance and photoluminescence are good examples, possess high sensitivity and accuracy and are widely used in the study of defects in materials. The power of

Paper 9936A (E1), received 6th July 1993

A.L. McCarren, H.J. Ruskin and M.O. Henry are with the School of Computer Applications and School of Physical Sciences, Dublin City University, Dublin 9, Eire

K.G. McGuigan is with the Department of Physics, Royal College of Surgeons in Ireland, Dublin 2, Eire

We thank J. Campion for making his data available to us prior to publication. We are also indebted to Dr. C. Do Carmo for several useful discussions.

spectroscopic measurements can be considerably enhanced by applying known perturbations to the system under study. Magnetic fields and both hydrostatic and uniaxial stresses are widely used perturbations in conjunction with optical spectroscopy. From the effects produced by the perturbation, much detailed information on the physical structure and electronic state of the defect may be obtained. The effects of a uniaxial stress on a crystal containing randomly distributed defects and possessing an axis of symmetry may be as simple as a shift of the spectral line position; this occurs when all defects have the same projection on the stress axis and the initial and final states of the transition producing the spectral line are both nondegenerate. Where the stress axis is not oriented equally to the symmetry axis of all defects, the orientational degeneracy of the defects is lifted under stress and a line splitting will be observed. Further complications may occur when the states are degenerate and/or adjacent states interact.

Since only a finite number of possible point defect configurations exist for crystals, general equations, which describe the effects of an applied uniaxial stress on cubic crystals, have been obtained for all possible configurations [9–11]. These equations, expressed in terms of several parameters (in units of energy per unit stress or shift rates), are known for stresses along ⟨0 0 1⟩, ⟨1 1 1⟩ and ⟨1 1 0⟩ crystallographic directions. In practice, data are recorded for stresses along these three directions and a best fit is sought between the data and the appropriate equations. In a successful fit, the symmetry of the defect, the symmetry properties of the electron energy states, the dipole nature of the transition and (in favourable cases) the nature of the defect vibrational nodes may be obtained. The difficulty, of course, lies in reconciling 'less than perfect' data with a set of equations describing a particular transition. Common occurrences include data missing for high stress values and overlapping spectral lines, leading to additional complexity in terms of both fitting and identification.

The basis of the mathematical method used to model the data is perturbation theory. It is assumed that the effect of the stress is a small perturbation to the zero-stress Hamiltonian for the defect. First order effects only are considered in general. In cases where the effect of the stress is to produce line shifts which are linear in the stress, the eigen-value equations may be derived analytically and the parameters estimated from the data in a straightforward manner. It is not uncommon, however, for a spectrum to consist of several closely spaced lines involving adjacent electron energy states, which may interact under stress [7, 12, 13]. Although these interactions can be formally deduced in an analytical solution, the parameters for each state will become imbedded in the solution. Consequently, recovery of the original parameters, (the physically meaningful results being sought), becomes difficult and is frequently compounded in situations, (see previously), where fewer experimental values than expected can actually be observed. The secular matrix containing the information on the energy shifts is then of order $n \times n$, whereas data may be available only for $m < n$ lines. The problem of matching a particular data point to a particular eigenvalue clearly becomes more complicated with increasing n in this situation.

The Hamiltonian with stress included may be expressed as

$$\hat{H} = \hat{H}_0 + \hat{H}' \tag{1}$$

where $\hat{H}'$ is the perturbation produced by the stress. The form of $\hat{H}'$ depends on the crystal structure and defect type, and suitable forms have been derived for commonly occurring situations. For the case considered below, a trigonal defect in a diamond type lattice, the perturbation may be expressed, from Reference 4 as

$$\begin{aligned}\hat{H}' = {} & \hat{a}_1(S_{xx} + S_{yy} + S_{zz}) + \hat{a}'_1(S_{yz} + S_{zx} + S_{xy}) \\ & + \hat{E}_x(S_{xx} + S_{yy} - 2S_{zz}) + \hat{E}'_x(S_{xx} + S_{yy} - 2S_{zz}) \\ & + \sqrt{(3)}\hat{E}_y(S_{xx} - S_{yy}) + \sqrt{(3)}\hat{E}'_y(S_{yz} - S_{zx})\end{aligned} \tag{2}$$

where $\hat{a}_1$, $\hat{a}'_1$, $\hat{E}_x$, $\hat{E}_y$, $\hat{E}'_x$, $\hat{E}'_y$ are operators, and $S_{ij} = |\boldsymbol{P}| \cos(\boldsymbol{P}, \boldsymbol{i}) \cos(\boldsymbol{P}, \boldsymbol{j})$ with $|P|$ the magnitude of the applied stress and $\cos(\boldsymbol{P}, \boldsymbol{i})$ the cosine of the angle between the direction of $\boldsymbol{P}$ and the crystal axis $\boldsymbol{i}$. In the absence of $\hat{H}'$ a transition between two states Ψ_a, Ψ_b of energy W_a and W_b, respectively, involves a photon of energy $W = W_a - W_b$. For the case where Ψ_a and Ψ_b are nondegenerate and when $\hat{H}'$ is included the energy of the photon becomes

$$W' = W_a - W_b + \langle \Psi_a | \hat{H}' | \Psi_a \rangle - \langle \Psi_b | \hat{H}' | \Psi_b \rangle \tag{3}$$

Accordingly, the energy shift observed in the measurement under $\hat{H}'$ is the difference of the shifts of the initial and final states and is given by the last two terms. The analysis of the data in general involves calculating matrix elements of the form $\langle \Psi_m | \hat{H}' | \Psi_m \rangle$, where $\hat{H}'$ typically takes the form of eqn. 2 above and the Ψ_m state vectors may be one, two or three-dimensional. Fortunately, group theory arguments simplify the calculations considerably by enabling factors which do not contribute to be identified and ignored. For the first example considered here (transitions between two-dimensional (E) states and one-dimensional (A) states at a trigonal defect in silicon), the effect of $\hat{H}'$ on the E states [4] is given by

$\hat{H}$	Ψ_x	Ψ_y
Ψ_x	$\alpha - \beta$	γ
Ψ_y	γ	$\alpha + \beta$

where

$$\begin{aligned}\alpha &= A_1(S_{xx} + S_{yy} + S_{zz}) + 2A_2(S_{yz} + S_{zx} + S_{xy}) \\ \beta &= B(S_{xx} + S_{yy} - 2S_{zz}) + C(S_{yz} + S_{zx} - 2S_{xy}) \\ \gamma &= \sqrt{(3)}B(S_{xx} - S_{yy}) + \sqrt{(3)}C(S_{yz} - S_{zx})\end{aligned} \tag{4}$$

Ψ_x and Ψ_y are the basis functions for the two-dimensional E state. A_1, A_2, B and C are derived from the operators in eqn. 2 [14], and are parameters which describe the line shifts with units of energy/stress.

The shifts observed for components of a spectral line due to transitions between an E and an A state are linear combinations of α, β and γ given by solutions λ_1 and λ_2 to the characteristic polynomial

$$\begin{vmatrix} \alpha - \beta - \lambda_1 & \gamma \\ \gamma & \alpha + \beta - \lambda_2 \end{vmatrix} = 0 \tag{5}$$

for the different stress directions. For the case of E to A transitions at a trigonal defect, there is a total of five characteristic polynomials, two each for stress along ⟨1 1 0⟩ and ⟨1 1 1⟩ directions and one for ⟨0 0 1⟩ stress. The measured values of the shifts may be used to fit simultaneously for each component and stress direction and to obtain best estimates of the parameters A_1, A_2, B, C. A simple system of this type is readily solved but seldom arises in practice. Factors which add to the com-

plexity were noted previously and include stress-induced interactions between adjacent energy states, overlap and crossover of components of lines when there are several closely spaced energy states. As a consequence, it is frequently difficult to relate experimental results to components of known transitions.

3 Solving for complex systems: Powell shell

Obtaining the characteristic polynomials or eigenvalue equations for each individual stress direction in a system can be facilitated by using models of the form

$$\hat{W}_t = \pi \hat{P}_t + v_t \quad \text{for } t = 1, \ldots, n \tag{6}$$

where $\hat{P}_t$ denotes the stress component, (the predetermined variable), $\hat{W}_t$ corresponds to the current variable (energy) and v_t is a stochastic error between the experimental energy values and predicted eigen values.

The matrix π takes the following form

$$\pi^T = [a \quad b \quad c] \tag{7}$$

where a is given by

$$a^T = [A_1 + 2B \quad A_1 - 2B \quad A_1 + 2A_2] \tag{8}$$

and correspondingly for b and c.

Parameter estimates for such models can be readily obtained by using simultaneous regression methods but it is evident from the form of eqns. 4 and 5 and following remarks that neither the eigenvalue nor characteristic polynomial approach will yield directly the physical parameters of interest, but will imbed these within the overall solution. Unfortunately, both are also highly spendthrift of computer time, since models of the type described by eqn. 6 require evaluation of the determinant of high order matrices.

A computer-based solution, while not essential for simple systems of the type described in eqn. 2, is clearly highly desirable for problems of increased complexity. It seems obvious therefore that an efficient method, which does not imbed the parameters but permits direct association with the experimental results, should be sought. We have found that the 'Powell–Shell' approach meets these requirements. The ranking involved in the matching procedure, (described below), means that we do not have an analytical expression for the overall residual sum of squares, where numerical approximation of the derivatives would involve a large number of extra evaluations. However, knowledge of the partial derivatives is not required for the class of optimisation procedures which include the Powell algorithm and its robustness and general applicability make it a good choice [5, 6, 15].

The matching of experimental to theoretical values is a lengthy procedure which can be minimised by ranking both sets of data from the highest to the lowest at each stress value. Effectively we define which experimental energy values belong to which 'line'. Having done this we associate experimental data points with eigenvalues, and avoid the problem of matching lines as such and hence the problems of 'cross-over lines' since associations now depend on ranked position. The number of comparisons is clearly sensibly diminished, since the highest ranked theoretical points need only be compared to the highest ranked actual data observations and similarly for the lowest ranked points, since otherwise we would be left with unassociated values. Thus for, say, eight expected values W_j of the given secular matrix and four experimental measurements M_j, the scheme of comparisons includes W_1 associated with M_1; W_2 with M_1, M_2; W_3 with M_1, M_2, M_3; W_4 with M_1, M_2, M_3, M_4; W_5 with M_1, M_2, M_3, M_4; W_6 with M_2, M_3, M_4; W_7 with M_3, M_4 and W_8 with M_4. The method generalises readily to matrices of any order.

This process is then carried out for all appropriate stress directions. The least squares fit is then obtained by checking each estimated line against each of the actual lines proposed by the matching technique described and the pairing with the minimum residual sum of squares is accepted and added to the overall residual sum of squares, ($S(\theta)$), dependent on θ, the parameter vector. The search for the optimum vector θ is accomplished in a number of iterations, each one of which normally involves many evaluations of $S(\theta)$ and consists of p (or $p + 1$) one dimensional searches in the p-dimensional space defined by the parameter set. Values of the initial points are user specified, but are updated together with step sizes during the optimisation process. Three consecutive points are used to fit a parabola and the abscissa of the minimum is used as a predictor for the minimum point on the actual curve. The function is calculated at this point and a new parabola is fitted using this last point and the best two previous points. The process is repeated for the new parabola and so on until a stopping criterion is satisfied. Individual linear searches are combined to obtain the global minimum. For cases where more than one transition type appears possible on the basis of experiment, the matching procedure also provides considerable information on the influence of subsets of the data on the overall fit and transition identification. Typically, piezo-spectroscopic experiments are characterised by well defined data at low stress values providing good initial estimates as a basis for fitting the data for all stress directions, but convergence is found to be satisfactory even when these initial estimates are quite poor. We include an analysis of performance with a range of conditions of estimation in the following section.

4 Powell–Shell results and performance assessment

In the following, we consider results obtained using the Powell–Shell for the problem described in Reference 7. The defect involved is observed in the photoluminescence spectrum of iron-doped silicon, when the sample is rapidly cooled from temperatures in the region of 1000°C to room temperature. Such rapid cooling rates prevent the usual precipitation of impurities such as iron and copper into clusters. This and other properties of these metallic impurities in silicon are discussed by Graff [16]. The results obtained by Do Carmo *et al.* [7] indicate that photoluminescence occurs at a defect of trigonal symmetry involving E to A transitions. Particular features of the defect include adjacent interacting E states and the occurrence of a splitting attributed to spin-orbit interactions in one of the E states. This leads to a complex interaction matrix and provides an excellent test of the Powell–Shell procedure.

The parameter values obtained are compared with those of Do Carmo *et al.* [7] in Table 1. Note that the

Table 1: Comparison of results obtained using the Powell–Shell and those described in Reference 7

Parameters	A_1	A_2	$B_1 = -B_2$	B_{12}	E	L
Do Carmo *et al.*	−5.4	−15.7	−12.4	8.7	2.7 meV	0.45 meV
Powell–Shell	−6.4	−14.7	−9.4	6.1	2.7 meV	0.45 meV

All units are in meV/GPa unless otherwise stated.

values of E and L are fixed by the spectroscopic data and that the physical significance of the parameter values is fully discussed in Reference 7. The differences between the two sets of values are not substantial, but a comparison of the closeness of the fit to the experimental data, shown

We present one further example here which consists of new experimental data obtained for a defect related to beryllium impurities in silicon [8]. Fig. 2 shows the fit obtained to these data assuming E to A transitions at a tetragonal defect. Although the form of the secular matrix

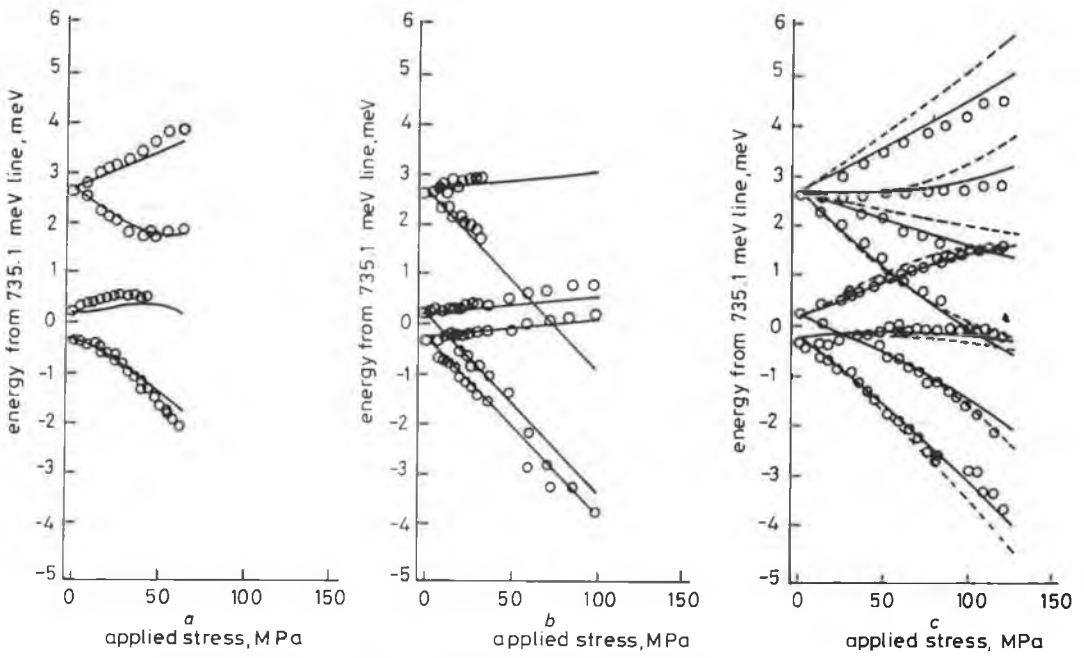


Fig. 1 *Results for stresses in (a) the ⟨0 0 1⟩ direction, (b) the ⟨1 1 1⟩ direction and (c) the ⟨1 1 0⟩ direction*

○ data points
——— fit obtained in this study
– – – – fit obtained with parameter values from Reference 7

in Fig. 1, is revealing. Figs. 1*a* and *b* show the results for ⟨0 0 1⟩ and ⟨1 1 1⟩ stresses; for these cases, the fits cannot be distinguished. For ⟨1 1 0⟩ stresses, shown in Fig. 1*c*, the amount of residual variation has clearly been reduced leading to more precise estimates of the required parameters. The procedure uses as initial estimates the relatively crude values for a subset of the data at low stresses, the fitting procedure is then applied and resulting parameter estimates initialise the fit for the complete data set.

is simpler here, the quality of the fit and parameter estimation is good and satisfactorily supports identification of the assumed transition in this previously untried case.

We also present an overview of the performance of the Powell–Shell, Table 2, in terms of an analysis of times to convergence on alternative choice of PC for different initial conditions. For a given set of initial estimates with a full data set, we obtain a final solution (Table 1) in 1 h 37 min 46 s on a Dell 486 PC running on the microsoft DOS operating system and 1 h 3 min and 35.9 s on a

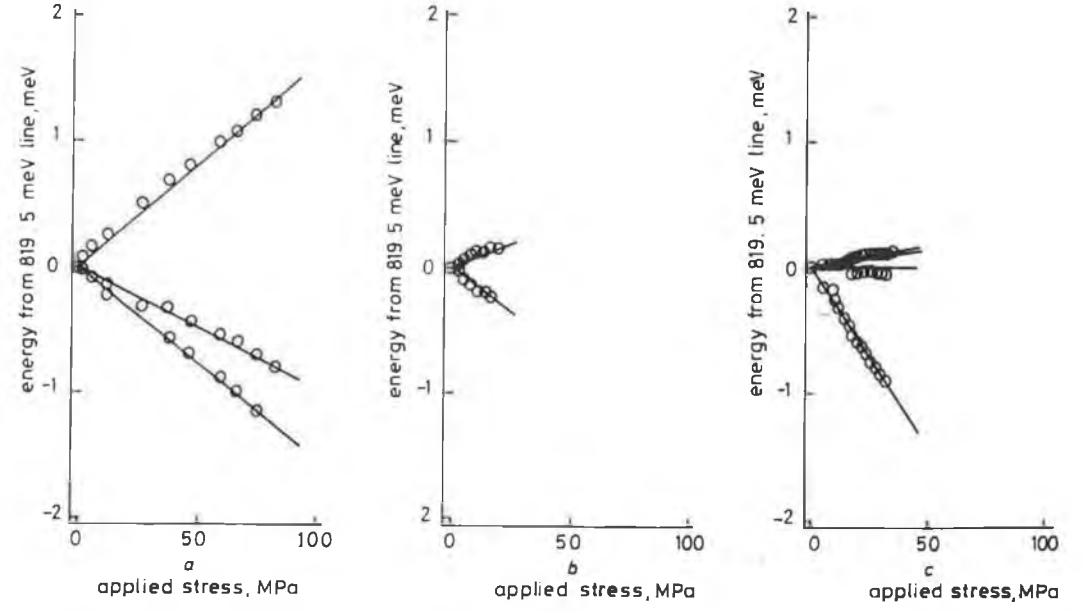


Fig. 2 *Results for stresses in (a) the ⟨0 0 1⟩ direction, (b) the ⟨1 1 1⟩ direction and (c) the ⟨1 1 0⟩ direction*

○ data points from Reference 8
——— fit obtained in this study, identified as a transition between E and A states at a defect with tetragonal (D_{2d}) symmetry